

第1回AI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4年7月26日(火)
15:00～
開催形式 Web会議

<開会>

○事務局（浜岡先端技術評価業務調整役） 定刻となりましたので、第1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、お集まりいただきありがとうございます。

<委員出席状況報告及び配付資料確認等>

まず委員の出欠状況について報告申し上げます。当委員会17名の委員のうち、現在17名に御出席いただいています。全委員の過半数に達していますので、専門部会規程第7条の規定に基づき本委員会の成立を御報告します。

次に、配付資料の確認をさせていただきます。議事次第、資料目録、別紙の資料取扱区分表、今日の資料は資料1～5、参考資料1～6までです。不足等がありましたら事務局までお申し付けいただければと思います。

次に資料取扱区分表を御覧ください。資料は内容に応じて、「取扱い」として「厳重管理」、「取扱注意」、「その他」に分類しまして、それに応じた対応を取ることとしています。本日の配付資料1～4、参考資料1～6はその他に該当しますので、委員各自で適切に保管・管理・廃棄をお願いいたします。資料5については取扱注意のため厳重に保管し、コピー等の複製、第三者への開示は御遠慮くださるようお願いいたします。

それでは佐久間部会長、議事の進行をお願いいたします。

<委員紹介>

○佐久間部会長 東京大学大学院工学系研究科の佐久間と申します。本専門部会の部会長を務めさせていただきます。よろしくをお願いいたします。

議事次第ではこの後委員紹介となっておりますが、私からさせていただきますと思います。東大の佐久間と申します。工学系研究科にいます。専門は生体工学で、手術支援ロボットや不整脈の解析をやっています。一部ユーザーとしてAIを使っています。レギュラトリーサイエンス関係では一時PMDAのお世話になりまして、副審査センター長を非常勤で務めて以来、いろいろ関連の活動をさせていただいています。科学委員会も実は設置された段階から全てに関係してきたということで、今回初めて専門部会長ということでAIのことについて担当させていただきます。よろしくをお願いいたします。

本日は初めての専門部会ですので、先生方の自己紹介をお願いできればと思います。資料1の名簿掲載順をお願いできればと思います。

- 伊藤委員 皆さん、こんにちは。国立がん研究センター東病院副院長で、大腸外科の科長もしています伊藤雅昭と申します。私はAIに関しては、AMEDの事業で日本の内視鏡外科手術の動画のデータベースを作る事業を行ってきました。今現在4,000近い動画のデータベースが我々の手元にあり、それをアカデミアのみならず企業の導出を積極的に行っています。
- その1つとして、AMED事業としてオリンパスの術中の内視鏡手術におけるインストラクション等における、それもAIを活用したデバイスを今、作っています。もう1つは、内視鏡手術の技術評価をAIで自動的に行うということもAMEDで行っています。我々の経験がこの会議に出せればと考えています。よろしくお願いいたします。
- 佐々木委員 東京大学の次世代病理情報連携学講座の佐々木と申します。AIに関しては、恐らく前回のプログラム医療機器の審査基準ワーキングのメンバーからやらせていただいています。またAMEDはJP-AIDという名前ですが、病理学会の病理診断に関するAIの開発の窓口として数年携わっていました。ここで開発したAIですが、今現在、薬事の承認を取りに東京大学の工学部の先生方と共同で行っていきまして、病理学会では私は常任理事、社会保険委員会の委員長も務めています。このプログラムに関して診療報酬で補填できるような方策を現在進めているところです。どうぞよろしくお願いいたします。
- 笹野委員 皆さん、はじめまして、東京医科歯科大学の笹野と申します。私は循環制御内科学の教授をしていまして、循環器内科で特に不整脈を専門にしています。
- AIに関しては、不整脈は心電図で発作時の波形が重要なのですが、発作が起きていないときの波形から発作の発症を予測することに関する研究を行っていきまして、AMEDの支援を受けて企業と機器の開発をしています。今、国の承認などを目指している段階です。その他にウェアラブル機器など、脈波や心弾動図などの生体情報で、その時系列データの解析というところでAIの活用を幾つかやっている関係で、今回選定していただいたものと思っています。どうぞよろしくお願いいたします。
- 澤委員 皆さん、こんにちは。帝京大学の澤と申します。専門は、電子カルテ等の医療情報システムと臨床麻酔科学です。
- 情報システムの方は、例えば電子カルテのデータの標準化や全国ネットワークというところに携わっています。AI、機械学習に関しては、大分昔の話ですが、大学院のthesisがコンピュータサイエンス、AI機械学習で、それ以降20年間ずっとフォローし続けているという形で、現在も

GAFAM 関係の PoC は度々携わらせていただいている状態です。どうぞよろしくお願いたします。

○清水委員 東京農工大学の清水と申します。工学部、工学研究院の出身です。プログラム医療機器の開発に携わった経験があり、最近レギュラトリーサイエンスについても少し携わっています。よろしくお願いたします。

○陣崎委員 慶應義塾大学放射線科学教室の教授をしています。AI、IT、それから医療情報、予防医療関係の副病院長も担っています。内閣府の AI ホスピタルプロジェクトの責任者をしていまして、IT、AI の医療への実装ということ自分のテーマとして担ってきました。

その他 JST の未来社会創造事業に関わったり、放射線医学会で AMED で構築した J-MID という CT 画像を何十万枚と集めてくるようなプロジェクトにも関わっています。特に実装のところをメインにやってきているところです。よろしくお願いたします。

○武田委員 皆さん、はじめまして。こんにちは。大阪大学医学部附属病院の医療情報部の武田と申します。専門は循環器内科で、今は医療情報学という教室で病院の電子カルテの管理も行っています。またいわゆる自然言語処理による情報抽出や、そういったデータを使って画像 AI のようなことも研究として取り組んでいたり、あと Society5.0 のプロジェクトでウェアラブルデバイスなどからの情報収集にも取り組んでいます。どうぞよろしくお願いたします。

○田中委員 皆さん、こんにちは。京都第二赤十字病院の田中と申します。AI に関しては、内視鏡の関連でやらせていただいています。皆様よろしくお願いたします。

○鎮西委員 鎮西と申します。産業技術総合研究所に所属しています。PMDA の科学委員会で 2016 年、2017 年ぐらいに一度機械学習、AI の医療運用というテーマでレポートを出していますが、そのときのメンバーだったということで恐らく呼ばれているのだと思います。よろしくお願いたします。

○殿村委員 弁護士殿村と申します。東京で弁護士をしています。私だけ弁護士なのですが、テクノロジー関連の法務を専門にしていまして、その中でも AI やデータ、若しくは個人情報の利活用を専門の 1 つにしています。その関係ではヘルステックに関与することが多くありまして、プログラム医療機器該当性の判断や医療の分野におけるデータの利活用も扱っております。AMED の方でデータの利活用プロジェクトを去年ぐらいから関わらせていただいています。その関係で御縁を頂いたものと思っています。よろしくお願いたします。

- 中岡委員 国立医薬品食品衛生研究所医療機器部の室長を務めています中岡と申します。皆さん、平素より大変お世話になっております。元々の専門はバイオマテリアルといたしまして、いわゆる医用材料、特に高分子材料に関する生体適合性評価が主なテーマだったのですが、ここにいらっしゃっている先生の大半にお世話になりましたが、次世代医療機器評価指標作成事業というものでAIを使った診断支援システムの評価指標の作成事業の事務局を担当しました。そのときにいろいろ御縁がありまして、いろいろな先生から御指導を受けつつレギュラトリーサイエンスの観点から研究を細々と続けさせていただいていますが、本日それもあってこの委員会に呼ばれたと理解しています。非常に僭越ながら話題提供という形で、そのときの研究の内容も紹介させていただきますので、よろしく願いいたします。
- 中田典生委員 東京慈恵会医科大学の人工知能医学研究部の中田典生です。元々の専門は放射線科医で、画像診断専門領域、それから超音波が専門で、超音波の指導医と専門医をやって、今でも病院で超音波センター長をやっています。現在、主な領域はCNNや勾配ブースティング、アンサンブル学習を使ったAIの開発をしています。直近では骨盤骨折のAIの論文、それから現在の超音波のAIの製品開発を企業とやっています。
- レギュラトリーサイエンス関係では、AMEDの超音波ナショナルデータベースの事業に参加しています。厚生労働省で保健医療分野におけるAI活用推進懇談会の構成員、PMDAでは先ほど中岡先生がおっしゃった医療機器の審査ガイドラインの委員、経産省の開発ガイドラインの委員をやっていました。よろしく願いします。
- 中田はる佳委員 皆様、はじめまして。国立がん研究センターの中田はる佳と申します。専門は研究倫理で、AIに関連するところでは医療機器の研究開発に関する規制や倫理を研究テーマにしたり、データの利活用といったところも研究領域の1つとしています。どうぞよろしく願いいたします。
- 村垣委員 東京女子医科大学の村垣と申します。臨床は脳外科で悪性脳腫瘍を専門としています。研究はAMEDのスマート治療室等に医療機器開発を行っています。メディカルAIセンター長も拝命し、学内全般の研究を見るとともに、自身は例えば遺伝子型を画像等で予測することや白血球減少を予測する等々をやっていますが、現在は手術の意思決定に関わるようなAIを開発しています。このような機会をありがとうございます。よろしく願い申し上げます。
- 森委員 名古屋大学情報学研究科の森と申します。私は画像処理を専門としてい

ます。EndoBRAINの開発等をやってきました。名古屋大学情報基盤センターのセンター長と、それから国立情報学研究所の医療ビッグデータ研究センターのセンター長をやっています、いわゆる大規模な情報基盤を活用した医療情報処理もやっているところです。どうぞよろしくお願いいたします。

○横井委員 香川大学の横井です。はじめまして、よろしくお願いいたします。専門は医療情報学をされていて、15年ほど前にPMDAにいた関係もあり、その後医療機器の審査安全対策等に関わってきました。先ほどお話がありました中岡先生や鎮西先生とはAIのガイドライン等で御一緒させていただいて、その後現在は薬事のプログラム医療機器調査会の委員等もさせていただいています。どうぞよろしくお願いいたします。

○佐久間部会長 先生方、ありがとうございました。

<副部会長の指名>

○佐久間部会長 それでは最初に、本部会の副部会長を決める必要があります。PMDAの科学委員会専門部会規程に基づき、部会長の私が指名するという事になっていますので、私から指名させていただきたいと思います。まず中岡先生、お願いできますか。

○中岡副部会長 私のような者であれば、慎んでお受けいたします。皆様のお力を借りることになるとは思いますが、よろしくお願いいたします。

○佐久間部会長 中岡先生、ありがとうございます。中岡先生は先ほどお話がありましたとおりレギュラトリーサイエンスのところでこの議論を進めていて、情報がまとまっていたりするので、その観点からお願いさせていただきたいと思っています。

それから専門部会規程5条の5に基づき、複数の副部会長を指名することができるということで、中岡先生はエンジニアリングというかレギュラトリーサイエンスの立場ということで、臨床的な観点から副部会長をもう一方指名させていただきたいと思います。国立がん研究センター東病院の伊藤先生にお願いしたいのですが、お願いできますか。

○伊藤副会長 私のような者であれば、協力させていただきたいと思います。微力ですが、よろしくお願いいたします。

○佐久間部会長 ありがとうございます。専門部会で報告書を作成した後、これを広く海外にも出すということで、英文のホームページへの報告書や、場合によっては専門部会として論文を作成する、先ほどの鎮西先生の前回の検討結果は、実は英文論文として出しています。そういう形で副部会長の先

生方にも御協力いただきながら、うまくまとめていきたいと思っておりますので、よろしくお願いいたします。

<科学委員会について>

○佐久間部会長 本日は科学委員会に参加されるのが初めての先生方もいらっしゃいますので、科学委員会について事務局から簡単に御説明いただこうと思っております。それでは事務局、よろしくお願いいたします。

○事務局（澁岡先端技術評価業務調整役） まずPMDAの御紹介から始めます。PMDAは平成16年に設立された独立行政法人です。医薬品の副作用等による健康被害の救済、医薬品・医療機器・再生医療等製品などの品質、有効性、安全性についての承認審査。それらの市販後の安全対策などを行っています。私ども、これら3つの業務によって国民を守るシステムをセイフティ・トライアングルと呼んでいます。

こちらはPMDAの理念です。PMDAの業務を行うに当たっては、この理念の下、行っているところです。科学委員会に関連があるのは3番目の所で、科学的視点での確かな判断を行うに当たって、先生方のお力を借り、テーマごとに検討をお願いしているということです。

科学委員会ですが、平成24年に設置されたものになります。当時、先端の研究内容を理解した審査などが特に再生医療などの分野で求められていまして、加速する技術革新をキャッチアップして、最先端の技術の実用化に貢献できる審査員の育成のためには、アカデミアとの密接な連携を図っていくことが必要だろうということで設置されたものです。当時の中期計画にも、下のような形で反映されています。

科学委員会はアカデミアや医療現場との連携を強化して、先端科学技術を応用した製品へのよりの確かな対応を図ることを目的として設置されたというのですが、審査業務等の科学的な側面に関する事項を審議するという位置付けになっています。具体的な役割としては、各審査項目の科学的評価に当たっての留意事項の取りまとめなどが期待されていまして、個別品目の承認審査に関わるというものではありません。委員については、医歯薬工など外部の専門家から構成され、議論を行う上で個別品目に係る資料を用いて議論することがあるので、会議は非公開。ただし、議事録は原則公開という形になっています。

科学委員会の親委員会と本専門部会というような形で成り立っています。それぞれの役割ですが、親委員会の役割は、議論すべきテーマを決定して、そのトピックごとに専門部会を設置し、専門部会委員の選定、それ

から専門部会において作成された報告書(案)について、確認をしていただくということになっています。一方、専門部会の役割ですが、親委員会で決定されたトピックについて議論を行い、報告書(案)をまとめるということになっています。

こちらは第6期の科学委員会のテーマです。科学委員会は2年ごとの任期になっていまして、それが6期目ということですが、その6期の検討テーマです。現在は2つのテーマで検討が進められているところで、それぞれ専門部会が設置されています。専門部会では検討開始からおおむね1年程度で報告書をまとめていただくようなスケジュールで、検討を行っているところです。この1番に書いてあるエクソソームの専門部会については、昨年度より設置されていまして、本年中の報告書取りまとめに向けて検討していただいているところです。AI 専門部会は今日からということ、こちらは資料では「副部会長は調整中」とありますが、先ほど中岡先生と伊藤先生が部会長より指名されたということです。

これは、これまでの科学委員会の成果として作られた報告の一覧です。これらは PMDA のホームページに掲載されるだけではなく、先ほど部会長からのお話がありましたように、内容によっては論文としてまとめていただいて、科学委員会の成果を国際的にアピールするということも行っています。

科学委員会でまとめられた報告書は、PMDA の Web サイトで公開されています。現在、これまでにまとめられた報告書、先ほど御覧いただいた成果については、このようなイメージで掲載されています。

次に科学委員会において検討するテーマについては、ホライズンスキヤニングという取組で選ぶということがありますので、それについて紹介したいと思います。まず概略ですが、ホライズンスキヤニングというのは将来登場するであろう技術を予測してみようという方法の1つで、大量の情報の中から未来に起こるかもしれない変化を多様に検討しているというものです。同定のステップで、新たに出現した技術の端緒をつかむために、学術論文やステイクホルダーなどから広く情報を収集し、ごく僅かな兆しを見いだします。非常に多くの情報を取り扱う必要があるため、これを効率的に取り込むためにデジタルツールの活用について私どもでは検討を進めているところです。同定で見いだされた新技術について、PMDA にとって対応の必要性があるかどうかというところで優先順位付けをし、評価をして、規制への影響であるとか、この技術に対して対処が必要なのかを検討し、必要な対応を取っているということになります。

す。こちらの必要な対応では、例えばガイドラインの作成などということに取り組んでいくこととなります。

こういったことに取り組むことについては、PMDA の中期計画にも定めるところで、このような書き方がされているところです。

こちらは PMDA 内の作業のイメージです。取組の流れを図で示したのですが、様々な情報源から事務局でスクリーニングをしまして、その中からふるい分けをし、PMDA で取り組むべきものを最終的に見いだしてきて、例えば科学委員会や PMDA 内部のいろいろな取組の中で対応するというようにしています。

ホライズンスキヤニングの取組に、デジタルツールを使って解析をしているという話をさせていただきましたが、トピック選定の手掛かり等、デジタルツールを使ってトピック選定をしたという所を論文化したものです。2 報を国際誌に登稿したところ、採択されまして情報発信したということで、情報提供させていただきます。以上、科学委員会などについて紹介させていただきました。ありがとうございます。

○佐久間部会長 ありがとうございます。ただいまの説明について、何か御質問はありますか。

1 点追加ですが、先ほど多くの先生方から、この AI の分野の中で評価指標の作成や、いろいろなガイダンスや WG などに参加されていると伺っています。この科学委員会の性格ですが、個別のものの審査について扱うものではないと言っています。その意味では実はガイドライン、いわゆる評価指標を作るものと何が違うのかという議論があるかと思えます。後で中岡先生のご講演でそこでの議論が出てくると思いますが、なかなか科学的に非常に難しい、どう考えたらよいか整理がされていないような点について議論をして、その論点を整理するという。論点整理の中には、こういう方向でよいというような考え方を出すこともありますし、これについてはまだ現状の学術の最先端をしてもはっきり分からないところがある場合にはそれを今後の課題としてそういう形で示しておくことがあるかと思えます。

後ほど私からこの AI の部門のところで、どういう課題があるかについて論点の例を出させていただきますが、ここに関して正に科学的な視点から考えるときに、科学的な視点だけではなくて今回実際に AI の研究をされている方にも入っていただいています。それから、それを応用してデータを集めてデータシェアリングをすると、そこには倫理上の課題もあるかもしれません。そのような議論ができるように今回委員の先生方

の構成ができていると理解していますので、これらの観点から議論をしていただければと思います。何か先生方から御質問はありますか。

<専門部会の設置の目的、進め方等について

AIを活用したプログラム医療機器専門部会について（佐久間部会長）>

○佐久間部会長 それでは議題に移りたいと思います。今まさに事務局から説明がありましたとおり、科学委員会ではまず親委員会で議論すべきトピックを決めトピックごとに専門部会を立ち上げて議論するというので、このAIを活用したプログラム領域に関する専門部会は昨年12月17日に開催された親委員会において設置が決まって私が部会長とさせていただくことになりました。本日は先生方と目線を合わせて、いわゆる問題の取扱い方についての共通認識を合わせることを押さえて、本専門部会のゴールとタイムラインを共有したいと考えています。

まず本部会の背景、目的、検討のイメージ、専門部会の進め方について、私の方で考えてきましたので御説明したいと思います。

それでは資料3を御覧ください。今回入っていただいた先生方の構成を考える上で、第3期、AIに関しての科学委員会のレポートとの差分をどのようにするのかを考えました。前回レポートについては先ほど鎮西委員が言及されましたが。そのときはAIの深層学習などが出てきて、その技術的な課題についてかなり集中的に議論をして、その論点の整理をしたと思います。

それから数年経ちまして、まさに今いろいろな先生方からその実用化に向けて進めているという状況となり、それをどう適切に安全性、有効性を評価してあまねくこの日本の医療の中で使ってもらうために審査をする中で、いろいろな課題が出てきておりそこについて議論するということです。その観点から、これまでは画像というものが主だったかと思いますが、今回実際に病理、内視鏡の画像、放射線、画像診断学関係の画像、バイタルサイン、そういうところの信号、そういうものについてもかなり対象とするものが広がってきています。そういうところで実際に取り組まれている先生方にも入っていただいて議論するという体制にさせていただきました。

専門部会設置の背景ですが、これは実際にAIを使ったということで承認されたプログラム医療機器をリストアップした表です。No.1のEndoBRAINは森先生が開発に携わられたと伺っていますが、それ以降この右に書いてあるような企業から続々と主に医用画像関係のものが出てき

ていると思います。最近では COVID の関係等のものが出てきています。

一方、右側に AMED の研究事業として進められていることがあります。皮膚がんの問題、それから三次元のフローサイトメトリー、これも画像ですがちょっと見え方が違ってきます。先ほど話がありましたが超音波、スマートフォンを使ったものでウェアラブルという言葉が出ましたがそういうことが出てきていたり、そこに示したような様々なものが研究開発されています。最後の AI 技術を活用した統合失調症の早期診断医療機器プログラムの開発、これはたぶん行動の様式やビデオ画像などを見ながらされていると伺っていますが、そのようなものが出てきている、続々と出てきそうだという状況になってきています。

そこで、開発・評価の考え方。これは 2017 年の科学委員会で出したもので先ほど私が言及したのですが、そこではこのスライドに書いたような章立てで報告がされています。技術の現状について話をすることと、AI 医療システムのレギュラトリーサイエンスということで、そこに AI 医療システムの特徴など、これはやはり臨床的位置付けをどうするかによってもかなりリスクが変わってきますし意味が変わってきますので、そういうところを明らかにしようと議論がまとめられています。それからデータセットとしての特性と信頼性をどうするのかといったような問題。リスクの問題です。これは、医療がどのようにこの AI を使うかということによっても変わるといっています。それから市販後、市販前。実は現在の AI の医療機器で市販後に性能が変わるというものは、今までないと伺っています。ただ AI の特徴として、データを蓄積することによって性能が変化することやどうやって扱うのかというのはかなり大きな課題ですが、そこについての基本的なフレームワークを考えていたということです。

あとは倫理の問題を書いていたということでもあります。その後実際にはいろいろな評価指標の通知が、その後出ましたいろいろな研究がされていることから、今日の科学的知見の進展に合わせる必要がある。1 点目にデータの再利用のあり方。実はこの委員会の中にも、データベースを集めている先生方もいらっしゃいます。これを活用する方向になってくると思いますが、1 回それでトレーニングしたものを、もう 1 回それでトレーニングする、評価をするということがあってよいのかという課題です。2 点目として評価データに求められる条件とは何なのか。これはアメリカの NIST でもバイアスという問題が出てきていると思いますが、そのようなところがあります。3 点目に市販後の性能変化があることを

意図するいわゆる Adaptive AI のあり方についてどう考えるべきなのか。これは最新の研究の知見など、ここに対してある種の論点整理、情報のインプットができればと思っています。この間、左下にあるとおり、FDA などいろいろなところでの議論が出てきているかと思っています。いろいろな指針を発表していますので、こういうものもしっかり参照しながら見ていかなければいけないと考えています。

本部会の目的は、先ほど言った現状の学問的なレビューに基づくのですが、そのレビューにとどまらず先ほども示した①から③に関する提言をまとめていければと考えています。これを誰が見るのかに関しては、開発者に対して適切に開発をして評価をすることを考えるということで、それを使ってもらいたいということです。それから今日お話がありますが、審査側からも実は要望を頂いています。やはりこういうものが出てきたときにどう考えるべきかという考え方について資することが目的になるかと思っています。

検討のイメージですが、まず AI を活用したプログラム医療機器の範囲ですが、実はここについて議論がありました。プログラム医療機器と言うと、プログラム単体で動くものです。一方で医療機器の中にこういうものがあるのではないか、これは不十分ではないかという議論があったのですが、ここはいろいろな配慮があります。実は社会的にも、プログラム医療機器はやはり推進すべきだという声があります。そういうところで、ここを少ししっかり出すということでプログラム医療機器としました。しかしながらここで考える AI を使った医療機器ということに関していえば、それはハードウェアも絡んだ中でソフトウェアが果たす役割が同類のものがありますので、そこについての考え方については、これは外挿して使えるだろうということがありますので、こういう形にしたいと考えています。ここも①から③を中心にまとめることについては、繰り返しになります。

スコープですが、今回ちょっとゲノムデータの AI 解析ということについては、除外しました。これはある意味で診断にも関連はするのですが、比較的これはいろいろな情報を統合していくというところで、先ほど来お話があった画像診断をするなど、心電図を解析して何かを知るということで、ちょっと特性が違うこともありますので、そこに関しては除外させていただくこととしました。それだけ御配慮いただければと思います。最後に書いていますが、既存の医療機器の中に組み込まれる AI を用いたソフトウェア機能の評価にも関連すると御理解いただければと思

ます。

進め方ですが、これは案です。ここで案と言いつつ今日実はもう既にやっております、後でまた進め方について御意見をいただきたいと思いますが、専門の先生方から検討項目に関する内容について御講演を頂いて検討を行うことにします。それである程度論点が出てくると思います。論点については、先ほど言った3つ。繰り返しになりますが、データ再利用のあり方、評価データに求められる条件、市販後性能変化にフォーカスしてやるということが審査側から見ても喫緊の課題になっていますので、それをやりたいと思います。

それで出てきたものについて、少しWGを持ちながらその整理をして報告書にどういう形のものを作っていか議論しながら、WGの結果についてまたこの専門部会で意見交換を行うということで、2回ほどさせていただきます。そうした上で、第3回以降に大体の報告書の骨子案を作りたいと思っています。それで実際には、執筆依頼をしていくことになります。執筆をお願いして、第4回くらいで素案の検討を行って、場合によっては追加的に専門家に御講演いただくこともあるかと思っています。先ほどちょっと話をしましたが、この科学委員会はメーカーの方は入らない形になっていますが、産業界の声も重要になってくるのであれば産業界の参考人という形で御講演をいただくことは可能です。そういうことも含めて検討できればと思っています。専門部会で報告書案をやり、第6回くらいまでにまとめて、最終的に親委員会に報告するという形を考えています。

次ですが、第1回ですのでPMDA審査部からどういう課題があるかをプレゼンしていただき、中岡先生からAMEDの研究成果を発表していただきます。8月に第1回のWGで、その3つの点についてどういう問題があるかについて案の作成を行います。そのことについて、専門部会で業界での議論を聞く場とする。参考人として聞く、話をするなどをして、WGで作成した案について、またこの場で議論していただくことにしたいと思っています。そこで報告書執筆等の割り振りを行えればと思っています。

実は親委員会から、AI技術だけではなく医療現場での使われ方まで考えるとより大きな問題になるのではないかといただいています。これは今日中田先生が御存知かもしれませんが、実はWHOからethics and governance of AI for healthというものが出ているので、この辺りについてもちろん参照しながら進めてくださいと親委員会から言われています。

それで、予定をざっくりこのような形で考えています。既に先生方には候補日についてお伝えして、スケジュールを押さえていただいています。第1回、第2回、第3回という形で、全6回を考えています。その間にWGが5回入るという形で、第1、2回は講演という形が主になってきますが、第3回ぐらいから報告書の検討が出てきます。報告書案が来年の冬ぐらいです。ですから、秋から冬にかけて執筆をお願いするという形になります。それについて検討してまたそれを改訂する作業をして、だいたい報告書ができたところで親委員会に部会長報告をする。途中経過になると思いますが、それを受けて親委員会からまたフィードバックを頂きながら、またWGを行って部会長報告をする。このような予定を考えています。

その後どうなるかという、実際に専門部会ということで1年後ぐらい、6月19日でまとめて、これをPMDAと厚生労働省にも確認をしていただきます。その中で全体的に、行政的なことから考えてコメントが入る可能性があるかと思っています。それで報告書案を修正できれば報告書としてPMDAのホームページに出します。翻訳を外注しますから、オープンになってしまうと論文にできませんので、実は専門部会である程度できたところから論文作成作業を進めなければいけないという話になるかと思っています。

どんなことが出ているかという、これは実は前回のものでも鎮西先生に頑張ってもらってまとめたものですが、最後にいらっしゃる光石先生がそのときの専門部会長です。私の名前も入っていますし多くの先生方に入ってもらっていますが、こういう形でレビューペーパーとして専門部会から報告しました。このような形で進めるということです。

一応部会長としてはこのような形で進めていきたいと思いますが、先生方何か御意見はありませんか。何か質問等があれば。私ばかり説明してしまいましたが、スコープはいかがでしょうか。たぶんこれは一番悩まれるところだと思うので。

- 伊藤副部会長 評価書案で1点、WGのメンバー作成はどうやって決められるのか、そのプロセスを教えてください。
- 佐久間部会長 WGについてはこの中で少しお願いをしながら進めていくということで、実は第1回WGはすでにお願ひさせていただいて準備を始めています。適宜お願いしながらという形になるかと思っています。
- 伊藤副部会長 ありがとうございます。
- 鎮西委員 Adaptive AIの件が③で入っているのですけれども、これは何を想定す

るかによって大分変わってくるような気がするのです。本当に勝手に性能変換していくものを考えるのか、それともどこかで製造販売者がコントロールをするのか。当然コントロールはすると思うのですが。確か既にIDATENの適用例が出ていますよね。ただAIではないという話は聞いたのです。その辺りにフォーカスするのか、それとも少し先のことを見るのか、どうしますか。

○佐久間部会長 これはこの後、PMDAからお話していただくところが多分芯になるかと思えます。その辺りで大体のスコopが決まってくるかと思えます。

○村垣委員 こちらのなかで評価データに求められる条件というのは分かるのですけれども、一方でPMDA的には評価方法をすでに確立しているのかというところが課題としてあるのではないかと思います。いわゆる正答率が何パーセント以上だったらよいとかがもう解決している問題なのか。議論が始まったところで、評価データに求められる条件とともに実際の評価方法というのは、まだまだいろいろ議論のあるところなのかなと思えますが、いかがでしょうか。

○佐久間部会長 これもこの後に岡崎さんからお話があると思えます。私が知っている範囲から言うと、評価というのはたぶんそれが何を標榜するかによって決まってきます。ですからこういう評価方法にしないかというのは、たぶん標榜する機能によって決まってくるのです。村垣先生、AI特有の課題ということですか。

○村垣委員 具体的に言うと、通常は人が行うのとAIが行うのと有意差という形で評価すると思えます。一方で、AIが単アームでやったときに正答率が何%であればよいのか？、例えばそれが2分割問題であれば今は9割ぐらいとする意見が多いですけれども3分割問題であればどうなのか、実際まだ誰も決めてないところが、レギュラトリーサイエンスとして興味があるかと思って、話題提供としてお話ししました。

○佐久間部会長 分かりました。多分そういう視点もあると思うので、評価データに求められるのはそういうことも想定してということになると思うので、その辺りも含めて頭に入れてという御指摘かと思えますので、入れさせていただいてその中で議論をさせていただければと思います。

<専門部会の名称について>

○佐久間部会長 それでは次に本専門部会の名称について検討していますけれども、この名称を「AIを活用したプログラム医療機器に関する専門部会」とした背景ですが、先ほど私が少し触れたのですが、規制改革会議と各所でプロ

プログラム医療機器が注目を集めている現状を踏まえて、プログラム医療機器を名称スコープに含めるのが望ましい状況だという判断になりました。そういう意味で、本部会では検討内容はプログラム医療機器に限らず、AIが組み込まれた有体物の医療機器についても展開が可能と考えるという形にしておりますので、以上からこの案としております。最後にちょっと言いましたが、これはこの点という理解で、そういう形の進め方によろしいでしょうか。では、それで進めさせていただきます。

それでは今回「AIを活用したプログラム医療機器に関する専門部会」とさせていただきますが、この英語についてです。これは専門部会の進捗成果物である報告書の英訳などを掲載するというので、英訳としては「Subcommittee on Software as a Medical Device Utilizing AI」、又は「Artificial Intelligence based Medical Software」など考えられますが、英語の名称として何か良いアイデアはありますか。特に御専門の先生方から見て違和感のないものにしておきたいと思っております。

「Subcommittee on」として、「Software as a Medical Device Utilizing AI」という言葉を使ってみました。そうすると少し広い意味になるかと思っただけです。

- 鎮西委員 「Medical Software」という言い方が、最近では逆にあまりしないような気がするのです。
- 佐久間部会長 そうですね。
- 鎮西委員 やはり前者のほうが、私は何となく好みです。
- 佐久間部会長 「Software as a Medical Device」というような言い方をして広く取らせることを意図したのです。それで「Utilizing AI」ということで、AIを使ったということを入れて広くすると。ソフトウェアだから別にプログラムソフト医療機器に限るわけではないので、ある種のフレキシビリティを入れることにしたのです。
- 鎮西委員 今、好みと申し上げましたが、この名前が例えばFDAなどの人たちから見えるものになるということですね。
- 佐久間部会長 そうです。
- 鎮西委員 そうだとしたら、やはり「SaMD」とした方がよいかと思っております。
- 佐久間部会長 そういう意味では「SaMD」という言葉が、今は比較的一般的ですね。
- 鎮西委員 はい。あるいは「Software drill Medical Devices and Utilizing AI」に、「SIG for」を最初に付けて、「Special Interest Group」ですね。そこは多分PMDAで統一がされていると思うので。
- 佐久間部会長 PMDAでは「Subcommittee」ですね。

- 鎮西委員 「Subcommittee」になっているのですね。
- 佐久間部会長 だからこれはやはりプログラム医療機器という話です。「Subcommittee on Software as a Medical Device Utilizing AI」。
- 中岡副部会長 やはり「SaMD」は入れた方が良くと思うのです。先ほどの議論の中では結局 SiMD は一応対象外という言い方は変ですけども、Software in Medical Device の方は今回は取りあえず入れた方が良く思うのです。
- 佐久間部会長 いや、入れる。
- 中岡副部会長 入れるのですけれども、今回の主たる討議は SaMD のほうですよ。
- 佐久間部会長 そうです。
- 中岡副部会長 それで出来上がったものに関して、より SiMD の方にもということを防げるものではないという感じですよ。そういう意味では SaMD を前面に出しておいたほうが良いかと思いました。
- 佐久間部会長 はい。御意見はどうでしょうか。
- 中田典生委員 米国 FDA がわざわざ「AI and Machine Learning」を「AI/ML」としているのは、本当は Machine Learning をしたかったのです。AI と言うと素人受けはするのですけれども、実際は Machine Learning が主体で、Machine Learning 以外の AI というのはもうほとんどルールベースの AI でそれはもう使わないということなのです。だから「Artificial Intelligence and Machine Learning」とわざわざ銘打っているのは、Machine Learning を入れたいけれども政治家や一般国民には AI を入れた方が良く、こういう名前になっているのです。その後に SaMD を付けて、FDA や国際的な文書は全部「AI and Machine Learning in SaMD」としてあるので、これが一番国際的には格好が付く。要するに我々は本当は Machine Learning が中心だけでもというのが分かるので、良いと思います。
- 佐久間部会長 「SaMD」は入れたほうが良い。そうすると資料 3 の 5 の FDA の 2021 年のあれですけども、例えば「Subcommittee on Software as a Medical Device」のどちらを先にするかです。AI と Machine Learning を最初に出すのかということです。
- 中岡副部会長 最近の文章でよく聞かれるのは「AI/ML」と書くのです。IMDRF の方では、どちらかと言うと ML SaMD になってしまっていますから。もちろん一般受けする意味で言えば。
- 佐久間部会長 一般受けすることは入れておいたほうが良いと思うのです。
- 中岡副部会長 Machine Learning という言葉は確かにあれですけども、やはり中田先生が言われたように対象は ML なので、どこかでそのキーワードは

入れたほうが良いという気がしました。

○佐久間部会長 では今、案を作ります。「Subcommittee on Software as a Medical Device Utilizing AI and Machine Learning」、あるいはFDAに倣って「Subcommittee on AI and Machine Learning in Software as a Medical Device」、今の議論を聞いていると感じとしてはそのぐらいですよ。どうでしょう。

○鎮西委員 議論の中心がAIとMachine Learningということですね。

○佐久間部会長 そちらになってしまうけれども、実はそれでよいかという問題もある。今回はやはり評価ということから言うと、そこまでいくのはつらい、あれかなと思うと。

○鎮西委員 やはりAIの使用がメインではないですかね。

○佐久間部会長 そうですね。そうすると今の議論で英語名は1行目のものでいかがでしょうか。よろしいですか。では、それでいきたいと思います。

<AIを活用したプログラム医療機器に関するご講演と意見交換>

①「機械学習を活用する医療機器の審査について」（プログラム医療機器審査室）>

○佐久間部会長 それでは次の議題に移りたいと思います。先ほどお話がありましたけれども、PMDA 審査部から審査の具体例も交えて、本専門部会で検討する論点に至った背景を御説明いただきたいと思っております。そうすることでプログラム医療機器審査室の岡崎室長に「機械学習を活用する医療機器の審査について」というタイトルで御講演を頂きたいと思っております。皆さんもかなり御関心のあるところではないかと思っております。では岡崎さん、よろしく願いいたします。

○岡崎室長 「機械学習を活用する医療機器の審査について」ということで、少しばかりお時間を頂き話題提供をさせていただきます。よろしく願いいたします。

こちらのスライドについては、先ほど佐久間部会長から御説明のあったものです。専門部会の目的ですけれども、誰がというところは開発者への情報提供、PMDA における審査が関心事ということで、それに資することを目的としています。論点についても重複してしまいますけれども、こちらに書いてある3つが今回の関心事項という認識です。

本邦での医療機器の規制については既に御案内のとおりかとは思いますが、改めて御説明させていただきます。医療機器のクラスというのが設定されており、リスク分類に応じたものになっております。クラス分類についてはそれに応じた許認可を得ていないと本邦での製造販

売ができないということになっており、比較的高いリスクが高い医療機器についてはPMDAに承認申請をしていただいで厚労大臣が承認するという立て付けになっております。

実際の承認審査がどういう考え方で行われているかが、こちらのスライドです。薬機法の中でどういった場合に承認を与えないかということが、ここに書かれていることで明示されています。幾つかあるもののうち赤枠で示した所が、今回少し御説明すべき点と考えております。開発されている品目が審査の結果、次のいずれかに該当するときということで①②③と書かれています。例えば効果又は性能を有すると認められないときとか、効果又は性能に比して著しく有害な作用を有することにより使用価値がないと認められるときといったものに該当することが審査で判断された場合には、承認を与えないということが法律で定められています。

次に該当するかしないかを実際にどのように判断しているかというのが、こちらのスライドです。実際には開発される方に通知で定められている資料をPMDAに提出していただくことになっております。その資料に基づいて臨床上もたらされる有効性と臨床現場に導入された際に発生し得るリスクを勘案し、そのバランスが適当であると認められたときにはその医療機器の承認申請を認めるというプロセスで審査を実施しています。

次に、医用画像診断支援ソフトウェアの評価の一例を御提示しております。これはソフトウェアを取り上げておりますけれども、開発のコンセプト、設計のコンセプトが満たされているかどうかというのを審査で確認するという考え方については、実はどの医療機器でも共通のことになっております。今回はCADを一例に考えますと、開発のコンセプトについては例えば放射線科医の見落としを減らすということが考えられます。一方、設計のコンセプトについては放射線診断専門医相当の検出能力を持つことが考えられます。

これらを実際の視点に落とし込んでいきますと、一番右が基本的な性能が満たされているか、真ん中が臨床性能が満たされているか、一番左が臨床的な有用性が満たされているかということで、これらの切り口から評価をしていくことになるかと思えます。それぞれの例については一番下の枠に書かせていただいておりますけれども、先ほど村垣委員から評価方法についても問題提起がありました。一例として臨床的有用性ということになりますと、CADを使った場合の読影とCADを使

わない場合の読影を比較してその成績が向上するということが言えれば、この品目の臨床的な有用性が満たされることになるのではないかと思います。先ほど佐久間部会長から御説明がありましたけれども、これは品目の開発のコンセプトや設計のコンセプトに依存し、どういった評価方法を採用すべきかが決まってくると考えております。

こちら先ほどと重複いたしますけれども、今年3月末現在で既に承認された、機械学習を活用する医療機器の一例となっております。承認上AIが使われていることが確認できるものとしては、既に20品目が認められています。

一方、機械学習を活用する医療機器の定義という視点になりますが、現在は薬機法上、人工知能や機械学習に関する定義はありません。一方で国際的な枠組みであるIMDRFにおける定義は決まっており、こちらに書かれているようなものが定義されているという状況です。

次に、機械学習を活用する医療機器における臨床データの種別という切り口で御説明差し上げたいと思います。まず上の方ですが、臨床データに対する性能試験というものが実施されます。その下の方に、臨床上的有用性を評価する性能試験という枠組みとなっております。まず上のほうです。臨床データに対する性能試験という枠組みの中で臨床データというものが存在するわけですが、その臨床データの中には、学習データ、バリデーションデータ、最終モデルの性能を評価するために使用されるデータの3つが存在するものと思っております。

どういった切り口で評価をしていくのかという中で、先ほど臨床上的有用性という切り口を御紹介申し上げました。臨床上的有用性を評価する性能試験についてはC)、評価データ、最終モデルの性能を評価するために使用されるデータというのが、非常に重要になってくるという位置付けです。実際にどのように審査をしているかという視点で考えますと、我々の審査では学習データやバリデーションデータがどのようなもので開発されたかという確認はいたしますけれども、最終的に妥当性を問うデータというのはこのデータの中では評価データということになります。どのような評価データに対してどのような結果が得られたのかということが、最終的に標榜できる性能ということになりますので非常に重要になってきます。一方でこの評価データについては実臨床を反映したものであることが望ましいのですけれども、どの程度のバリエーションが必要なのかという点については非常に難しい論点・話題であると認識しております。

可塑性の有無という切り口から見た機械学習を活用する医療機器の種類を考えますと、ここに書いてある Locked type と Adaptive type の2種類が存在しているのではないかと考えております。上の方の Locked type については、機能を変更しないようにアルゴリズムが固定されているもので、こちらは通常の審査と基本的なコンセプトは同じかと考えております。一方の Adaptive type については、新たなデータに基づいて繰り返しその機能などを変更していきます。初回の性能の評価については当然通常の審査と同じになってきますけれども、2回目以降はどうするかということが話題になります。こちらの Adaptive type の評価・開発に資する制度として、先ほども少し御紹介のあった IDATEN という制度が既に施行されていて、一般にはこの IDATEN という制度を利用していくことが期待されています。

こちらが IDATEN の制度の概略を説明した図です。これまでの通常の手続を真ん中に示しております。臨床データ等を集めて必要なデータが充足したら、申請をしていただいて審査を行います。それで問題がないという判断になったら承認をするわけです。例えば性能の変更をするまでに、新たなデータ等を収集して、必要なデータが満たされたら変更の申請をして、改めて一連の申請をしていただき、審査が終了したらリリースできるというのがこれまでの流れでした。

しかし、下の方の変更計画を用いた手続を見ていただくと、それとの差分、2回目の変更の申請後の審査の部分が少し短くなっていることが見て取れると思います。通常の手続であれば変更の申請をしていただいた後に試験方法の妥当性や結果を確認するというフェーズが走っていたのですが、今度の新しい IDATEN の手続は試験方法や変更の計画自体は既に最初に行政側で確認しておくという点が通常の手続と異なっております。その部分を前倒しすることによって届け出た内容を確認するだけで済むこととなりますので、IDATEN の制度については変更が繰り返し行われるような医療機器についてはその活用が望まれていることになっております。

ここで今回の専門部会で御議論いただきたい点のそれぞれの論点について、少し御紹介させていただきます。1つ目としてはデータ再利用のあり方ということで、市販後の製品の改良をする場合に初回の承認申請で用いたデータを再利用するということが想定される一方で、改めてその都度評価データを用意するというのは非常にコストが掛かることが想定されますから、再利用するニーズは非常に高いと思われれます。

ただ一方で、評価データを再利用してしまうと単に評価データに対する最適化がなされてしまって、汎化性能が損なわれていても気付けない可能性があるのではないかと思っております。こういったことを防ぐための対応例としてはこちらに書かれているようなことが考えられますけれども、本当にこれだけなのか、これらが適切なのかといったことで、データの再利用の可否とか、再利用するための要件を御議論いただきたいと考えているところです。

こちらが実際にそのことが問題になった事例です。Nodoca という新医療機器があります。こちらは審査報告書の一部を抜粋したものです。黒塗りになっている部分が多くて大変恐縮ですけれども、実際に審査で確認したことは、11枚目と12枚目のスライドにわたって、審査報告書で説明させていただいています。要するに治験の結果から意図的な性能の向上が行われていないことを審査の中で確認させていただきました。もう1点が、治験のデータのみに対して製品が最適化されていないのかということも審査の中で確認させていただきました。具体的な事例としてこういったものがありましたので、紹介させていただきました。11枚目、12枚目のスライドについての説明は以上です。

次に論点②、13枚目のスライドです。冒頭でも少し申し上げましたが、機械学習を活用する医療機器についてはブラックボックスの特性を有していますので、開発者が認知し得ない弱点が存在し得ると考えております。その検証をする場合には様々なバリエーションのデータを網羅的に収集して評価する必要があると考えておりますが、その一方で網羅的であることを客観的に説明するというのは非常に困難であると思っております。また先ほども申し上げましたけれども、際限なくデータを収集し続けるということはコスト等の観点からも現実的ではないと考えています。

評価データのバリエーションの例として、こちらの枠囲みに書いているようなことを想定しております。後ほど中岡先生からも御紹介いただけたところでいろいろな研究がなされていますが、評価データのバリエーションの考え方や十分性、評価データが存在しないパターンに対する対策の考え方等について、考え方を少し整理したいと考えています。

最後の論点の③、14ページをお願いします。こちらについては、市販後の学習として医療機関ごとにバージョンアップがなされることも想定し得ると思っております。現実的にそういった開発がなされるかど

うかはまだ予想し得ないところがありますけれども、そういった製品が開発されることも想定されている中で製造販売業者と同様の品質マネジメントを臨床現場に求める工夫等が必要になってくると考えられます。実際に何を求めればきちんと品質が満たされるかといったことについて、具体的な事例を御検討いただきたいと思います。

具体的な対応例としては、ここに書かせていただきました。当初設定した性能があったときにバージョンアップをしてしまうことによって、臨床上許容できる性能を下回ってしまうときには初期値に戻すといった機能とか、臨床的な位置付けが変わらないように事前に性能を固定する機能の実装といったことが、具体的な対応例として考えられます。

2つ目として、市販後の学習を通じて市販前に使用していた機械学習のアルゴリズムの変更とか、同一の機械学習モデルにおけるハイパーパラメータの変更も想定されるわけですが、こちらはより具体的な事例について御検討いただきたいと思います。なお、一番下の所に赤字で書かせていただいています。医療機関ごとにバージョンアップを行うことが許されるのかどうか、そういったものが1品目の範囲なのかどうかというのは、そもそもの立て付けの問題として論点にはなるのですが、その論点についてはこの科学委員会での議論ではないという認識をしています。もしそういったことが許容されることになったときにどういった対応が考えられるのかということについては、審査部的には関心事だと考えているところです。以上です。どうもありがとうございました。

○佐久間部会長 岡崎室長、ありがとうございました。委員の先生方、ただいまの御講演に対して御質問、御意見等がありますか。

ある意味で具体的に審査側からの困り事というか、こういう課題はどうしますかということで。一方で今日データベースを扱っていらっしゃる先生方がたくさんいらっしゃいますけれども、こういうデータを本当に集めるのはとても大変だということは理解しています。そうすると評価データを再利用するということが出てくるけれども、やはり一方で機械学習の観点からいくとちょっと危険も存在するかもしれないといったことがある。その辺りをどう整理したらよいかということも含まれているのですが、先生方、何かありますか。

○中田典生委員 この議論はコロナの前からずっとやっていたものの延長線上ですが、残念ながらコロナの間にもっとAIが進んでしまって、画像の深層学習だけではなくてそれ以前のハイパーパラメータを今まで手で決

めていたのを、ハードウェアによって膨大なパラメータを順繰りに端からやっていくようなことができるようになってしまっているのです。そうすると Adaptive AI、要するに市販後の性能が変わるというのは、今後もっと激しくなると考えられます。

その対策として今流行っているのが Explainable AI です。要するに、この AI は一体何を根拠にしてこれを出したのかというもの、画像以外だと Feature Importance という、Feature(特徴量)の、どれが一番 AI が寄与しているかというのを出してくれるようになります。これは何のために出しているかということ、結局これを使うユーザー、医師と AI の関係ですね。今後は医師がこれを信頼するかどうかを見ながら決めていくというように変わってくると考えられます。

ということは今 PMDA が審査に大変苦勞されているのは非常に分かったのですが、ある程度緩い DMZ ではないけれども厳密にする前の段階のゾーンを作らないと。アメリカの FDA はそうしているわけです。それは何かというと人間、いわゆる医師の責任を問うようなものを設定する必要性が今のところはあって、AI は絶対ではないということです。しかも Deep Learning というのは、ある 1 つの特徴量だけを取り出すのは非常に優れているのですけれども、人間の放射線科医のように全ての所見を拾い上げるのは、多分できないだろうということです。

恐らくこれから出てくるのは画像以外、例えばカルテの音声の文字情報から得られた情報、臨床検査のデータ、画像診断で出てきたある所見の確率、これらを全部ミックスしたものが出てくるはずですが、そうするとそれは結構大変な今後の新しい AI の医療機器の方針になるのです。そういうことで私の言いたいことは、中間ゾーンのようなものを設定する必要があるのではないかというのが、このスライドを見ていての印象です。

○佐久間部会長 他にはいかがでしょうか。

○森委員 13 ページで気になった所が 1 つだけあります。「ブラックボックスの特性により」と書いてあるのがすごく引っ掛かっているのです。機械学習はブラックボックスだからと書くのは、やめておいたほうがよいのではないかと思いました。機械学習を使わないソフトウェアでも、開発者が認知し得ない弱点というのは大量に存在するものですから、機械学習がブラックボックスだから開発者が認知しない弱点があるというわけではないと思いました。

次も 13 ページに書いてある事柄です。これは AI/ML でしか対応、特

化したような問題点ではなくて、実を言うとソフトウェアを使った全ての装置というのはここに書いてある問題点は全て当てはまるのではないかと思います。普通に画像処理をやるような装置でも全然違ったデータが入ってきたらちゃんと動かないとか、いろいろな問題点が、今も画像処理をやっているだけでもたくさん発生しますので、ここに書かれている問題点は、AIだけに限ることではないのではないかと思います。

2つ目は論点の③、市販後に性能変化することを意図するAIの審査のあり方です。ここでどうやって対応すればよいかというのはやはりこの委員会の中で議論をしなければいけないと思うのですが、少なくともここに書いてあるように臨床上許容できる性能が変更されたときに初期の性能が保たれているということをどうやって担保するかというそのやり方はかなり難しい気がしていますので、そこも実際に議論ができるとよいのかなというように思っています。

あと最近、今、中田先生が言っていたように、オートMLといって大量の計算を一気に流すものがあります。スーパーコンピュータレベルのものだと大量にハイパーパラメータ調整をしますから、そのことまでどこかで触れられればよいかなと思います。ちょっと長くなりましてすみません。

○佐久間部会長 ありがとうございます。他はございますか。

○村垣委員 先ほどの議論に似ているのですけれども、臨床的有用性と性能評価の2つあり、承認のときには有用性ということでした。例えば医者との比較で非劣性なのか優越性なのかどちらを見ているか、また、今回IDATENでその後アップデートするときに再度その以前の医者との比較をもう一回やるのか、承認された以前のデータとの比較をするのか、あるいは単に性能評価で劣っていないことを見るのか、がわからなかったです。これまでの事例とか、あるいはPMDAが今後そういった形でその学習が繰り返されて良くなったものを評価するとき、性能が劣らないということでしたが、その辺nodocaの例でもよくわからなかったので教えていただければと思います。どういうデータを再利用するのかと評価方法と、両方からんでくると思いますけれどもいかがでしょうか。

○佐久間部会長 岡崎さん、いかがでしょうか。nodocaの例はやや関連しますね。

○村垣委員 そのアップデートしたものが以前と変わらないということは、結構変数が多いと思うのです。もう一度元の医者と比較させるのかとか。

○岡崎室長 御質問ありがとうございます。nodoca の事例についてはスライド 11、12 枚目になるのですが、まず一旦データを学習して評価が行われて、その後治験を実施したわけなのですが、その後でもう一度学び直して改良を実施したということになっていて、では何に対する性能の向上なのかという御質問については、当初比較したベースとなる性能の基本の値というのがあって、それに対して性能が向上したことが示されたということになります。そういう意味ではベースとなる元のデータというのは変わっていないということです。本件に関してはそういう形になります。

あと既に利用されている IDATEN でどういう評価がなされていたのかということについてはまだ IDATEN の活用事例が多くないので、このように評価されていましたとここで説明してしまうと差し障りもあるので詳細の御説明は避けさせていただきたいと思っておりますけれども、一般論として IDATEN であっても通常の一部変更承認申請であっても、対照群のデータというのは一般的に妥当な対照群のデータを示していただいて、それが対照群のデータとして妥当であるということが説明できればそれは OPC であれ何であれ、利用ができるのではないかと考えているところです。

○村垣委員 簡単にでもよいのですけれども、これまで 20 の機械というのはほとんど医者との比較で承認されているのか、機械の性能自体で承認されているのか、ざっくりどのような感じですか。

○岡崎室長 臨床的有用性の観点かと思いますが、そちらについては両方あるのかと思います。

○村垣委員 了解です。また教えてください。

○佐久間部会長 それでは時間も押していますので、この辺りでよろしいでしょうか。今日出てきた議論をもう一度見直していただいて、今後の議論にお役立ただけであればと思います。資料を渡してありますので、よく御覧いただければと思います。

<AI を活用したプログラム医療機器に関するご講演と意見交換

②「人工知能等の先端技術を利用した医療機器プログラムの薬事規制のあり方に関する研究」から（仮題）（中岡委員）>

○佐久間部会長 次は AMED 医薬品等規制調和・評価研究事業 2019～2021 年度の「人工知能等の先端技術を利用した医療機器プログラムの薬事規制のあり方に関する研究」で得られた成果を専門部会にインプットしていただく

ために、中岡副部長に御講演いただきたいと思います。中岡先生、よろしくお願ひいたします。

○中岡副部長　　今、論点をいろいろ PMDA のほうに御紹介いただいたのですが、それを聞いたときに少し思ったのが、次世代事業で今回参加している多数の先生方に御協力いただきましたけれどもその当時から論点になっていた部分が結構残っているかなという印象と、あとこの事業でもいろいろ議論をさせていただきました。

元々この研究のお題はここにありますように薬事規制のあり方に関する研究ということで、大枠を考えるとというような形の事業でこれは立ち上がりました。今更いろいろ言う必要がないと思いますけれども、一番大きいのはやはり市販後性能変化というところに関しては従来の薬事では対応できないケースが生じうるというところで、薬事のあり方自体を見直したほうがよいのではないかというお題を頂きまして、研究をスタートしました。

目的は、いわゆる研究事業による課題抽出と解決策の検討。それを通じて将来あるべき薬事規制システムを提言するというものでございました。基本的には薬事的課題は産官学で抽出・整理ということでしたけれども、基本的には業界団体の視点で当初は行っておりました。

抽出整理された課題についての解決策ですが、これも業界から一応素案を頂いた上でそちらを有識者で議論すると。そちらのほうを取りまとめて議論を経て作り上げた解決策等を整理して、提言にするということを想定して活動しておりました。

体制としてはそのために私が一応代表を務めさせていただきましたけれども、JIRA の古川さんにも研究開発分担者として参加いただいていわゆる産業界からの意見抽出の取りまとめをしていただきました。実際に JIRA の参画メーカーだけではなく医機連のほうにお願いして、関連業界団体でAIを使ったいわゆる Machine Learning を使った医療機器の開発等を検討している団体が多数おられましたので、そういう方々に参加いただいて WG を設立して、自分たちが今直面している薬事的な課題の抽出をしていただいた上で、解決策を作ると。そちらのほうを九大の名誉教授の橋爪先生に座長をお願いして、今回参加していただいているメンバーですと清水先生、横井先生、森先生とか、そういう方々に御参加いただいて、そちらの解決策に関して妥当性に関していろいろ御検討いただいて、その出来上がったものを提言にするという形で考えておりました。

実際に「これまでに」とは書いたのですが、WGで抽出されたメインの課題がこれぐらいになります。AI医療機器という形で言いますと、結局分類とかがはっきりしない。設計過程のみでAI機能を使用したらどうなるのかと。保守として使用した場合にはどうなるのか。この辺りは結局、薬機法で言うところの標榜です。広告とかの所の問題ということです。こちらのほうは特にマークを付けておりませんが、これは結局個人情報保護法に絡む部分です。薬事の問題というところを超えていますので、その辺りに関して議論は十分できませんでしたけれども、やはりメーカーとしては結構問題視しているというところではあります。

先ほどもいろいろありましたデータに関してです。学習データとか検証データに関しての信頼性であるとか、実際に必要なデータ数、どれぐらいのデータ数が必要なのかというところ。データの特定化、及び動画の扱いをどうしたらよいか。教師付きデータの場合は、資格、誰がやればよいのかというところ。大概、製品化する前はベースになるモデルがあって、それに対して対象となる疾患の学習データを学習させていくということをやるので、その前の学習済モデル、基本モデルの扱いをどう考えたらよいのですか、これは薬事的にどのように考えたらよいのですか。他社開発のモデルの扱いはどうなるのかというところではあります。あと承認申請の実際の手続に関してはいろいろ議論がありました。

赤字の部分は実際に検討会まで討議を経て、成果物を整備したものです。この辺り、最初のほうに関しては、どちらかというところと広告の標榜の問題なので、これは適切に広告してくださいというところではあります。

学習済モデルに関しては、結局、開発する企業の責任をもって、その学習済モデルがいわゆる法を侵していないと。日本で言うところのオプトインで取ったデータ以外のデータを使っていないということに関しては、ちゃんと自分の所が責任を取って確認しなさいと。そういう形で考えた場合には、学習済モデルというのは原材料として扱ってよいのではないかと。いわゆるQMSで言うところの原材料の管理で、きちんとしましょうということを提言すればよいのではないかとということになりました。黄色の所が、実は科学的なデータがないとなかなか難しいというところがありましたので、この議論がなかなか難しいというところがありまして、AMEDと相談して実証研究をやらないかというお話になりました。

実証研究はそれなりの費用が当然必要になりますので、実際できるか

どうか分からないけれどもやってみましょうということで調整費を申請したところ無事採択されまして、学習データ、教師付きデータ、承認申請に関する事項等に関係するような、いわゆるデータの考え方で。特に、先ほどからいろいろ議論に出ていました Adaptive な AI に対して通用するもの。いわゆる市販後学習と我々一応名付けていますけれども、市販後の追加学習によって性能が変わっていく機器を対象にした場合の学習データや教師付きデータ、そういうものに対する要件を明確化するために実証研究をしましょうということで、今回部会にも委員として参加されている清水先生の多大な協力を得まして実際に調整費を使って実証研究をしました。今日はそちらの成果を紹介させていただきますが、基本的には私が統括する形でやりましたけれども清水先生が実際のデータの方、解釈等々、実験の流れ等は当然担当していただきましたので、ここは必要に応じて補足いただければと思います。

調整費が採択された背景としては、先ほどから言われています IDATEN が施行されたことが一番大きな要因です。ここでいわゆる届出の対象となるものということで、市販後のリアルワールドデータを使用した診断の正確性の向上というものが書いてありますけれども、実際にどういうものになるのかということろだったのですが、結論から申しますと法的には臨床なし一変に相当するものに関しては一変ではなく届出にという扱いで IDATEN に載せることができるということになりました。その辺りの区別、判断ですが、実際に診断正確性の向上と標榜を計画する場合にそれが実際に臨床ありなのか、臨床なしの一変になるのかということろの判断の基準というものに関しては、そこは踏み込むのは難しいというのが基本的には検討会のほうでも結論となっておりますので、そちらに関しては基本的に規制当局に任ずということで、あくまでもそうなった場合というものを想定してそのための要件を整理していきましょうかと計画いたしました。

実際にどうしても市販後学習すると、性能が単純に上がるというだけではなくて、当然メーカーのほうは適用範囲を拡大しようとかいろいろなことを考えてきます。臨床のお医者さんは、想定外の使用方法を思いつくことが当然あるかと思えます。アメリカも同じような状況だと思います。実際 2019 年以降、すごいスピードでいろいろなことが動いています。佐久間座長からも御紹介がありましたけれども、GMLP なるシステムを入れて、いってみれば Machine Learning を利用した SaMD

用の QMS、そういうものを作れば実際に性能変化を共有できるのではないかという意見がアメリカの方でも出ていると思いますけれども、いまだにやはりそこがきちんとした形にならないのは、いろいろな問題点があるのだと思います。そういう状況であれば、日本も IDATEN をうまく活用することでいわゆる Adaptive な AI の実用化を進めていくことが仮にできれば、日本が早くそういうものを実装した国になるということもあります。

ただ現実的にリアルタイムに、施設ごとに性能が変わるものを対象とした場合、これは医療機器の品質担保という観点からみると医療機器として規制するのは難しくなってしまうということがありますので、あくまでも想定すべきは段階的とはいえメーカーが市販後学習を実施する場合、メーカーが市販後学習を実施して性能変化をすることができるよう AI 利用医療機器をまず第一段階として社会実装しよう。それを実装するためにどういうことを考えればよいかということを考えました。

考慮すべき点ということでは、先ほど言いましたとおり Adaptive な場合には使用施設ごとに性能が変化してしまいます。ということで、医療機器としての承認時における品質性能を担保をすることは可能かと。この辺りは論点になっているものと同等の問題だと思います。

実際に学習に使用するデータに対して、要求すべき品質が不明と。よく機械学習、深層学習等は学習をさせていった場合に突然性能が劣化してしまう、変化してしまうと聞いております。その辺りは恐らくデータの一貫性であるとか質の問題とかがあるのかと。そう考えた場合には例えば製品開発のデータとどの程度の同等性が必要になるのかとか、異なる装置、パラメータで撮像したデータというのは市販後学習に使用できるのかという疑問点が出てまいります。学習後に変化した性能に対する評価を承認時と同じ方法で実施するのは妥当か、これも論点として出ました。評価データを使い回しても本当によいのかと。テストデータが対象とする疾患の母集団を代表しているのであればよいのでしょうかけれども、データの数が少ない場合とかなれば当然そのデータに特化して性能が上がってしまっていて、一般化した性能は多分必ずしも上がっていない可能性があります。その辺りに関してやはりしっかりと知見を得ることができないかと。そういうものを得ることができれば IDATEN も効率よく回せるだろうと考えて研究をスタートしています。

研究体制としてそれまでは提言を作るだけだったのですけれども、清水先生に御尽力いただきまして、東京大学医学部附属病院、現在千葉大学のほうへ移られました野村先生、実際に annotation とかをお願いした日本メジフィジックス、この方々を新たに実証研究グループとして追加して調整費の取得とともに研究を始めました。

具体的にやっていたということは、次のスライドからになります。まず annotation の影響です。東大病院のグループに担当していただきました。

目的としては、正解入力作業に携わる診断医の能力及び人数による市販後学習後の性能変化検証とその影響です。影響を及ぼす要因候補の明確化というところです。

具体的に実際にばらつきの例を東大病院から提供いただいていますけれども、ここの肺結節の部分の領域をマークしてくださいということになると、2名ともここが患部だと入力しているのですけれども、このように入力範囲が異なっています。こういうばらつきが生じる。あと当然1名の方はここは違いますと言いながら、もう1名はここが疾患だよねという形でマークすると。このようにばらつきが生じてきます。まずは対象のCADソフトウェアは、東大病院でモデル開発済の肺結節検出のソフトウェア、それと頭部MRA画像から脳動脈瘤検出ソフトウェアをベースとして使って、ちょっと煩雑なのですけれども、まず初期開発に関しては3名の放射線科診断専門医による annotation を行っていたいただいたもの。しかも同じ先生方に評価用症例の annotation をしていただきました。そういうものを用意しました。■■初期開発用症例の annotation をしていただいた先生方に、annotation していただいたデータをテストデータとして用いてその性能変化を見て検討していただきました。

結果をもうここで示してしまいます。実際にどれだけのデータを使ったかというのはお手元の資料を御覧ください。これが、実際肺結節の12名の方々のデータを使って再学習をしたときの感度の変化を見たものです。■■実施する方に応じてこれだけ変わるということです。

■■ annotation においてはある一定の専門医がやれば何とかなるのではないかという感じの結果を得ております。

もう1つ、撮像装置及び条件の影響ということで清水先生たちにやっていたいただいた結果の方ですが、こちらに関しては結局施設ごとの学習を模擬してその場合は向上する場合も低下する場合もあるだろうと。

特にただ、その施設で学習してその施設で取得したデータをテストデータとした場合は、もしかしたらその施設に特化した性能になるのではないかと考えられました。ただそういうことが起こるのかということもありますし、データの中の撮像条件とか患者集団というものが影響を及ぼす可能性があるということで、その辺りに関しての検討を骨シンチデータを用いて実験を行いました。

■■結局施設による最適化というのは起こっていないということが示唆されました。

■■

まとめをこのように示しておりますが、結局ここまでのデータにおいては annotation に関しては一定の専門医の方が一貫性を持って annotation すれば恐らく性能が劣化することはないであろうと。■■

あと、これは市販後学習と少し違うのですけれども、初期用モデル構築の段階で実は私どもで購入したマシンと清水先生のほうで購入していただいたマシンとで全く同じデータ、全く同じプロトコールで初期モデルを構築したのですけれども、そこで得られた最適なモデルというものの性能が結構異なっておりました。ということはある程度そのところも対策を練っておかないと、市販後学習をする、施設ごとに仮にすることを容認するとなると、そのプロトコールを適切にメーカーが構築しておかないと、性能が大きくばらつく可能性があることも分かりました。少し駆け足になりましたけれども、以上がこの研究の概要、大枠になります。

今後の展開と書いてあるのですが、実は AMED の先ほどの研究は昨年度で終わったのですが、後継の研究で実は規制の評価に着目して実証実験を続けたいということで二次公募に出しまして採択されました。今回論点として PMDA が提示された案件とか、村垣先生が先ほどコメントされた評価方法の妥当性というものを検討するためにはどういうところに着目すればよいかというところを、RS の観点から清水先生と共に研究を続けていくことになっておりますので一応参考情報としてお知らせいたします。

最後、謝辞。これだけのメンバーにいろいろ御協力いただきました。本当に感謝しております。私からの発表は以上になります。

○佐久間部会長

10分ぐらい意見交換させていただきたいと思います。何か先生方から御質問はございませんでしょうか。施設ごとに学習データが良くなる場合もあれば良くならない場合もあるという結果があるということ

をお示しいただいたのかと思いますが。

○鎮西委員

実は施設ごとということに関しては、開発ガイドラインで2019年から2018年ぐらいに扱ったときも少しディスカッションになりました。最終的には、施設ごとができるかどうかというのは薬機法の制度的な問題というよりは、むしろメーカーがそういうものに対して製造物責任はソフトウェアはないのだけれども民法上のそういう責任をちゃんと負えるかどうかというところで結局決まるのではなかろうかという話になりました。なのでどういうやり方でやるかによりますし、今ディスカッションになったような本当に施設ごとに性能を上げていくことができるのかどうかということと併せて、メーカーが結局それをよしとするかどうかということで決まってくるでしょうというのが、そのときの結論です。

もう一点は真ん中辺りに「調整費による追加検討のキーポイント」というスライドがありますけれども、その1つ前の真ん中辺り、学習後に変化した性能に対する評価を承認時と同じ方法で実施するのは妥当か。その下の所でテストデータが対象とする疾患の母集団を代表しているかどうか分からないとしたら、そもそもAIであろうとなかろうとそういうデータは使えないですよ。一応母集団を代表していると仮定してやるものなのではないかと思ったのですけれども、ここはどういう意図でこういう文章になったのでしょうか。

○中岡副部長

表現はちょっと悪かったと思いますがけれども、テストデータが疾患の本来はその母集団を代表とするということの間違いないと思いますけれども、ただ実際に市販後学習していたデータを追加していったときに、そのデータを追加していたことで母集団からずれていく可能性はないのかということです。適応拡大ではないとしても実際その疾患の度合いとかに応じて、かなりテストデータから外れた値を学習に入れている可能性もあるということになってくると、少しその辺りに関してはいわゆる承認申請時に使ったテストデータの使い回しというのは本当に妥当なのかということは議論の余地があるかなという意味合いで書きました。

○鎮西委員

ありがとうございます。先ほどの審査部からの資料の中にもあったのですが、データが再利用で、本当は覚えてほしくないデータも覚えてもらうのが困るのだという話だったと思うのですが、それならテストをした前のバージョンに戻れば済むだけのような気がするのですね。開発ガイドラインの中でも結局そういう話にして、だからこれ

自体それほど大きな問題ではないという言い方をしたような気がするのですけれども、そうではないケースというのもあるのでしょうか。要するに、バージョンを戻せないような状態というのを想定するかどうかだと思っております。

○中岡副部会長　そうですね。そこはだから製販のほうの管理という面だと思います。だからもう1つ言えば、いわゆるテストデータの適切な管理というものも必要になるのかなと。間違っているバリデーションを使ってしまったら、途端にテストデータとしての意味はなくなりますので、その辺りのプロセスが適切かどうか。

○鎮西委員　これは正にQMSの世界だと思います。

○村垣委員　短時間で今回もすばらしい研究だと思いますけれども、学習データなのか評価データなのか分からないのですけれども、そのときに施設間で比較するときには陽性率なのか重症率なのか、その辺をそろえておかないと正しい評価ができないのかなと直感的に思いました。よくランダム化のときにバイアスを外すために層別化をやるのですが、それと同様に何らかの形でイベントに関する割合に関して調整しなければいけないのかなと思って、この問題の難しさを感じました。

もう一点は私の直感で、恐らく臨床試験のときにいわゆる実力テストみたいな隠しデータ、セットを持っておいてそれで取りあえず評価しておく。そしてアップデートしたときもそれは当然試験をやっている人たち、企業とかソフトウェアには分からないわけなので、アップデートとしては同じような実力テストをやらせて変わらなければという方法が一番妥当ではないかなと考えました。

○中岡副部会長　ありがとうございます。多分今の性能評価の縦軸に関して今回フォールスポジティブ・プラス・フォールスネガティブを選んでおられるのはその辺りを考慮されてのことではないかなと思いますけれども、清水先生、いかがですか。

○清水委員　FP+FNというのは単純に誤りの画素、誤りの割合を表しているので、クラス分類における基本データということで使っています。ただ目的に応じて他の指標が良いということは、十分ありうるかと思えます。

○佐久間部会長　清水先生、評価のデータのところでやはり評価のデータを1回使ってしまおうとそれで評価を繰り返すということの危険性があるということ、かつて先生がどこかで指摘されていたような気がしますけれども、それも実はあるのですけれども、先ほど村垣先生がおっしゃっていたこの評価データとして固定しても、結局その問題が出てくるとい

うことがありうるということですか。

○清水委員

同じ実力テストを何度も使って評価すると、やはりそれにバイアスがかかってしまう可能性があります。この実力テストで常に1回だけテストをしてこれで合格ということであればよいのですが、例えば1度実力テストのボーダーを下回ってしまった、アルゴリズムを変えてもう一回（同じ）実力テストを受けてとすると、結局幾つかのモデルの中から実力テストを使って選択するというプロセスになってしまいますので、やはりどうしてもバイアスがかかってしまうリスクがあります。そのバイアスがどれだけの大きさか、それからバイアスをどうやって補正するかが、多分これから大事な問題になってくるのではないかと考えております。

○伊藤副部長

中岡先生、素晴らしい御研究ありがとうございます。ただ施設間で学習することによって結構ばらつきがあるというのは、ある意味当たり前ですよね。多分最初に使った学習とかなり異なった状況なり画像なりでやると、やはりセンシティブティのばらつきがあるというのは、これは致し方ないことのような気がします。世界中の全てのデータを学習で使うことによってどの施設もばらつきがないというのだったらよいのですけれども、こういったタイプのプログラム機器は世の中に医療を前に進めるか否かが一番大きいではないですか。ですのでやはり性能だけでばらつきがあるから駄目というのは、私は問題があるのではないかと思うのです。やはりそこに臨床的なアウトカムにおいて、IDATENにおいても承認後にやはりそれが担保される、あるいは上乘せになるという、先ほどから性能が良くなるはずだということもあると思いますけれども、むしろプラスマイナスのある程度非劣性マージンみたいな、ちょっと悪くなったとしてもそこは許容できますというマージンを取った上である程度性能を評価するという、その辺の度量といいますか、こういったアプリケーションを世の中に広めていくためにはそういう考え方も結構重要なのではないかなと思います。いかがでしょうか。

○中岡副部長

おっしゃるとおりだと思います。理想はもちろんある程度データに一貫性を持たすということであって、日本における放射線とか超音波のデータですけれども、取得のデータにおいては割とアメリカと違って技師さんとかお医者さんのほうの特性というか好みというか、そういうものがあるので、そういう意味ではアメリカなどと比べればデータのばらつきは生じやすいのではないかと思います。

非劣性のマージンというか、ここら辺は多分どちらかという規制側とメーカーとの間の技というか、合意の下、そういうような幅ですね。

○伊藤副部長 そうです。

○中岡副部長 その幅というものをどのように臨床的意義も踏まえて設定するかというところがまず1つの肝になってくるということと、これに関しては規制の方々がいなかった所で業界の方とお話はしたのですが、市販後学習という観点ではなく例えば納品後、その技師さんとか先生方が取ったデータを使った市販後学習を行って、その施設でちゃんと性能が出るようにそういうチューニング目的の市販後学習というものは許容可能なのではないかと。それをやった上で例えばその後、施設ごとに市販後学習をメーカーがやっていくと、より良いものになるのではないかという話はしていました。ですので薬事のあり方として、その辺りも考慮すべき点ではないかなと考えているところです。

○中田典生委員 1つレギュラトリーサイエンス側に知っておいていただきたい開発側の事情として、同じプログラムで同じデータでモデルを作成しても10回やれば10回違う性能のものが出来上がる。業者は其中で一番良いものを使います。なぜ同じプログラムで同じデータを使って学習しても同じ結果にならないかという、大体Deep Learningだとランダムシード、乱数表の指数を固定していないからなのです。結構施設間でも差があったりするだけではなくて、プログラムだって何回もやれば違う結果が出て、業者というのは一番良い性能を求めて作るということをお忘れはいけません。科学研究ではそれを比較するためにわざと乱数シードを固定してみたりして、わざわざ性能が低くなったりすることがあるのですけれども、そういうことがあるので開発サイドとしてこれもっと工夫できるではないですかみたいことは、審査側が知っていないといけないということです。

○中岡副部長 そこは先ほどもお話したように清水先生といろいろ議論になったところであって、結局統計学的な処理をしているだけですから、学習しているのもランダムに取ってきますし、そのプログラムの特性に応じて、当然同じプログラムを使っても性能が、あるところを見れば変わってくる、違ってくる。その中で最適なモデルを選択して製品化するというのは当然のことだと思いますし、そこは規制側も理解されていると思います。ただそれをベースに今度市販後学習にしたときにも同じことが起こりうることは気を付けないといけないという趣旨で入れ

させていただいたと、御理解いただければよいかなと思います。

○佐久間部会長 ありがとうございました。本日の議事は以上ですけれども、今日は中岡先生ありがとうございました。

今日は随分深い議論を、短時間ですけれどもさせていただいたのではないかなと思っています。一方で様々な活動を先生方にしていただいているので、今最後に言われた実際に審査側が知っておくべきこととか、その辺り、この医療機器は是非世の中に出していかなければいけないと思いますので、それを適正に安全かつ有効に社会に出すための論点整理をするというのが目的になると思いますので、是非よろしく願いいたします。

本日の議事は以上ですけれども、事務局から他に何かございますでしょうか。

<その他>

○事務局（澁岡先端技術評価業務調整役） 次回の専門部会につきましては、9月26日（月）の15～17時の開催を予定しております。詳細につきましては、追って御連絡させていただきます。また8月10日（水）の14～16時には、第1回のWGを予定しております、御参加いただく先生方には既に御案内を差し上げておりますので、よろしく願いいたします。

○佐久間部会長 準備の段階で部会長ということで、今回こういう議論になるだろうということで、まず第1次のWGの先生方にお声掛けをしております。まずその結果を出させていただいて、またどう先生にWGに入っていくということも含めて、継続的にやってまいりますので、まず第1回につきましてはそういう形で進めさせていただければと思っています。

<閉会>

○佐久間部会長 本日の専門部会はここまでとさせていただきます。本当に今日は長時間にわたり熱く議論していただき、ありがとうございました。