

日本医療研究開発機構 臨床研究等 ICT 基盤構築・人工知能実装研究事業 事後評価報告書

公開

I 基本情報

研究開発課題名: (日本語) 病理診断支援のための人工知能(病理診断支援 AI) 開発と統合的「AI 医療画像知」
の創出
(英 語) Development of artificial intelligence to help pathological diagnosis
and creation of integrated “AI medical image knowledge”

研究開発実施期間: 令和2年9月15日～令和3年3月31日

研究開発代表者 氏名: (日本語) 中島 直樹
(英 語) Naoki Nakashima

研究開発代表者 所属機関・部署・役職:
(日本語) 一般社団法人日本医療情報学会 代表理事
(英 語) Japan Association for Medical Informatics, President

II 研究開発の概要

1. 健康医療情報の商用利用も含めた2次利用のための同意取得の方法の法制度・倫理課題抽出、およびワークフロー整備に関する研究

(1) 研究開発の目的

2017年度の個人情報保護法改正により、要配慮個人情報となった健康医療情報は第三者への提供には個別同意が必要となった。ただし、医学系研究に用いる場合は、「人を対象とする医学系研究に関する倫理指針(以下、倫理指針)」に基づいて倫理審査委員会の承認により、オプトアウトで研究実施が可能な場合もある。倫理指針に基づいて収集した個人情報については、企業による AI 製品の開発等に供する場合は、上記に照らせば新たに患者から個別同意を取得するか、匿名加工化する等の措置をとらなければならない。

また、次世代医療基盤法に則り、認定匿名加工医療情報作成事業者(以下、認定事業者)が匿名加工化する場合には「通知によるオプトアウト(上記のオプトアウトの利用拒否の機会保証を紙の通知を個人に渡すことにより行うこと)」が必要とされる。

包括同意(診療の一環として取得された試料余剰分や情報を、将来実施する様々な研究に利用することを文書で個別同意を得ること)を実施する医療機関もみられるが、その法的根拠は不明確である。さらには、新しい同

意の取得手法として、スマートフォン等のモバイル端末を登録し、電子的な動的同意取得（経時的な説明と同意／同意拒否の確認、異なる目的別の同意取得）を可能とする dynamic consent が期待されている。

そこで本研究では、健康医療情報の商用利用も含めた 2 次利用のための同意取得の方法として「dynamic consent」、「通知によるオプトアウト」、「包括同意」を取り上げ、それらの法制度・倫理課題の抽出、およびワークフロー整備を行う。また、個人情報保護委員会等への意見提出資料を作成する。

（２）研究開発成果の概要

健康医療情報の商用利用も含めた 2 次利用については、医療機関や学会等において既に収集されている健康医療情報（以下、「既に収集された健康医療情報」）を 2 次利用する場合と、医療機関や学会等が今後、収集する予定の健康医療情報（以下、「これから収集する健康医療情報」）を 2 次利用する場合で用いるべき同意手法や課題が異なる。そこで、この区分けを踏まえ、「dynamic consent」、「通知によるオプトアウト」、「包括同意」における法制度・倫理課題の抽出、およびワークフロー整備等を行う。

なお、「通知によるオプトアウト」の検討を行う際、個人情報保護法には同意を得ずに 2 次利用を可能にする匿名加工情報が存在し、これについても比較考察を行った。これについても別途、整理を行っている。

① 「dynamic consent」

スマートフォン等の普及や「人を対象とする生命科学・医学系研究に関する倫理指針」（改定案）において「電磁的方法によるインフォームド・コンセント」が示されたことから、「dynamic consent」の実現性が高まっている。

ただし、「dynamic consent」を行うために、健康医療情報を収集するタイミングで、例えばスマートフォンを登録すること等により、個人とのオンラインコミュニケーション手段を確保する必要がある、「既に収集された健康医療情報」には適さないと考えられる。なお、「既に収集された健康医療情報」の収集組織において、患者との連絡手段を確保している場合はその限りではない。

そこで、主に「これから収集する健康医療情報」を対象として「dynamic consent」を行う場合を対象として検討を行った。健康医療情報を収集するタイミングでは、「dynamic consent」を行うこと自体への同意が必要になる。これを「DC 活用同意」と呼ぶこととする。

したがって、まず DC 活用同意を取得し、その後「dynamic consent」を行う必要がある。さらに、「dynamic consent」には厳密には opt-in で行う「dynamic consent」と opt-out で行う「dynamic consent」が存在する。これらの違いも考慮して、「DC 活用同意+dynamic consent」のワークフローを作成し、検討を行った。

「DC 活用同意+dynamic consent (opt-in)」では、「dynamic consent」において個人に対し情報提供が行われ、それに対して本人が確実に同意を表明するため、商用利用も含めた健康医療情報の 2 次利用が可能になる。「dynamic consent (opt-in)」で取得する同意が商用の 2 次利用の場合、患者への説明内容等について学術研究を対象とした「人を対象とする生命科学・医学系研究に関する倫理指針」（案）に準拠することは求められない。そのため、どのような説明内容とするべきか、同指針と整合を取るべきか検討が必要である。また、DC 活用同意において個人に伝えるべき内容の整理、「dynamic consent」を実施する際の本人確認方法等については、詳細な検討が求められる。

「DC 活用同意+dynamic consent (opt-out)」では、商用利用における健康医療情報の 2 次利用はできない。個人情報保護法では要配慮個人情報の opt-out による第三者提供は禁止されているためである。したがって、「DC 活用同意+dynamic consent (opt-out)」による 2 次利用が可能になるのは学術研究の範囲であり、「人を対象とする生命科学・医学系研究に関する倫理指針」（改定案）で認められている範囲となる。同指針においても、opt-out が認められているのは、個別同意の取得が困難な場合、情報が匿名化されている場合等、限定的であり、DC 活用同意を取得できる状況にあることを考慮すると、学術研究においても「dynamic consent」は opt-in の活用

が基本となると考えられる。なお、ここで言う「dynamic consent (opt-out)」とは、スマートフォン等のモバイル端末により同意の撤回の機会を設けることを想定しているが、これまで医療機関等で実施してきた院内掲示等による通知、説明を行うことを妨げるものではない。「DC 活用同意+dynamic consent (opt-in)」と同様、DC 活用同意において個人に伝えるべき内容の整理、「dynamic consent」を実施する際の本人確認方法等については、詳細な検討が必要である。

次に、DC 活用同意と「dynamic consent」の関係性について整理する。DC 活用同意においても、「dynamic consent」と同様に、同意した内容を撤回することが可能である。ただし、DC 活用同意と「dynamic consent」は連動しているわけではなく、DC 活用同意の撤回が「dynamic consent」の撤回にはならないことについてあらかじめ周知することが必要と考えられる。

「dynamic consent」については、ある研究に対する健康医療情報の 2 次利用の継続性について同意を取得する手段としても有効と考えられ、同意内容の階層化（ある研究に関連した追加同意等）についても、どのように対応すべきか検討する必要がある。

なお、「dynamic consent」については、上記の内容について整理した上で、個人情報保護委員会等に提出する資料についても作成を行った。

② 「通知によるオプトアウト」

「通知によるオプトアウト」、すなわち次世代医療基盤法に基づく認定事業者を用いた健康医療情報の 2 次利用については、「既に収集された健康医療情報」、「これから収集する健康医療情報」、双方において活用することが可能と考えられる。それぞれワークフローを作成し、検討を行った。

「既に収集された健康医療情報」において次世代医療基盤法を用いる場合には、収集組織（例えば、臨床学会等の一次研究の主体者）から認定事業者へ健康医療情報を提供するため、医療情報取扱事業者である医療機関から収集組織へ提供行為に関する委託契約等が新たに必要になる。また、医療機関側では、患者へ「通知によるオプトアウト」、来院していない患者、死亡している患者への対応等の課題がある。収集組織が代行することも想定されるが、本人確認郵便等の費用と手間を要する。なお、その際、当該医療機関の代理で収集機関が実施していること、オプトアウトの受け付けは当該医療機関で行うこと等を確実に患者に伝えることが必要である。オプトアウトを受け付けた当該医療機関は、受け付けた旨を書面等で患者に対して通知する必要がある。当然ながら、通知ができなかった患者のデータおよびオプトアウトを受け付けた患者のデータは、認定事業者が次世代医療基盤法に従って利用することはできない。また、当該医療機関では、主務大臣への届出を行うことが必要であり、これについても収集組織に委託することが可能である。収集組織から認定事業者へ提供された健康医療情報については、民間企業等の申請に基づき、認定事業者が匿名加工医療情報の利用目的の妥当性等についての審査を行い、当該利用目的に沿った匿名加工が可能であることを検討した上で匿名加工医療情報を作成し、契約等に基づき適切に提供し 2 次利用されることになる。

「これから収集する健康医療情報」において次世代医療基盤法を用いる場合、健康医療情報の収集先の医療機関等において、前述したように「通知によるオプトアウト」、主務大臣への届出を行った上で、認定事業者が健康医療情報を収集することになる。収集された健康医療情報については、民間企業等の申請に基づき、認定事業者が匿名加工医療情報の利用目的の妥当性等についての審査を行い、当該利用目的に沿った匿名加工が可能であることを検討した上で匿名加工医療情報を作成し、契約等に基づき、適切に提供、2 次利用されることになる。

③ 個人情報保護法に基づく匿名加工情報

次世代医療基盤法の運用とは異なるものの、個人情報保護法には同意を得ずに 2 次利用を可能にする匿名加工情報が規定されており、これについても比較考察を行った。匿名加工情報についても「既に収集された健康医療情報」、「これから収集する健康医療情報」、双方において活用することが可能と考えられる。それぞれワークフ

ローを作成し、検討を行った。なお、次世代医療基盤法に基づき匿名加工医療情報を作成する場合は、収集組織から認定事業者健康医療情報を提供した時点で認定事業者が健康医療情報の 2 次利用に対して責任を負うことになる一方で、個人情報保護法に基づき匿名加工情報を作成する場合は、収集組織が健康医療情報の 2 次利用に対して責任を負わなければならない。

「既に収集された健康医療情報」、「これから収集する健康医療情報」いずれの場合であっても、学術研究目的で収集した情報から匿名加工情報を作成する際、個人情報保護法第 76 条の適用除外から外れ、個人情報保護法の適用を受けることになる。したがって、収集組織には個人情報取扱事業者としての義務が生じることに留意が必要である。

「既に収集された健康医療情報」に対して個人情報保護法の匿名加工情報を適用する場合、仮に収集組織が学術研究目的で収集していたとしても、収集元となる医療機関において匿名加工を行う必要はなく、収集組織において個人情報保護法に基づいて匿名加工を行うことが可能である。ただし、この匿名加工を行うことを前提として、個人情報保護法の学術研究目的による適用除外を用いることは倫理的に問題があり、実施すべきではない。また、当該研究が倫理審査を受ける際に作成された研究開発計画書において、医療情報の 2 次利用を行わないことが明記されている場合も倫理的に問題がある。研究開発計画書に示した医療情報の取扱いについての内容に抵触しない範囲内で、匿名加工を行えることに留意する必要がある。

収集組織において匿名加工を行う場合、収集組織は個人情報保護法に基づき、匿名加工情報の適正な加工、匿名加工情報の作成や第三者提供に関する公表、匿名加工方法等の安全管理措置等について実施することが必要となる。さらに、収集組織において匿名加工情報を作成して民間企業等に商用のために提供する際は、再識別の禁止、匿名加工情報の安全管理措置等について当該企業等と適切に契約を締結しなければならない。なお、匿名加工情報の作成については、外部に委託することが可能である。

「これから収集する健康医療情報」において個人情報保護法の匿名加工情報を用いる場合、各医療機関等において匿名加工を行うことになる。個人情報保護法に基づき、匿名加工情報の適正な加工、匿名加工情報の作成や第三者提供に関する公表、匿名加工方法等の安全管理措置等について実施することが必要となる。また、収集組織において匿名加工情報を作成して民間企業等に商用のために提供する際は、再識別の禁止、匿名加工情報の安全管理措置等について適切に契約を締結しなければならない。なお、匿名加工情報の作成については、外部に委託することが可能である。なお、匿名加工を行うことを前提として、個人情報保護法の学術研究目的による適用除外を用いて収集することに倫理的な問題があることは、「既に収集された健康医療情報」と同様である。

④ 「包括同意」

「包括同意」は、目的の明確な研究への試料・情報提供の同意を取得する「個別同意」と対をなす概念であり、同意を取得する範囲をまったく限定しない「blanket consent」と、ある程度範囲を限定する「broad consent」に大別される。個別同意と異なり、未確定な研究に対しても、将来の研究を考えて患者から情報の提供を求めて同意を取るため、患者が同意や拒否の判断をするための具体的で明確な個別研究についての説明を同意取得時にはすることができないという特徴がある。健康医療情報については、研究内容が多岐に渡り、将来的な研究を見据えて「包括同意」を取ることがあるが、商用利用も含む 2 次利用まで「包括同意」を取得することは困難であると考えられる。このようなことから、健康医療情報を直接取得する収集組織における適正な情報管理を前提として「包括同意」を有効とすることが重要となる。例えば、収集組織内における倫理審査委員会において、2 次利用について審査、承認する体制、プロセス等を整備し、これを適切に運用することが求められる。また、学術研究利用、商用利用等の利用目的、学術研究機関、企業等の利用組織等についても可能な範囲で明示して、選択的に同意を取得することも考えられる。

また、ヘルシンキ宣言において、同意撤回権の保障がインフォームド・コンセントの標準的要件の一つとされていることを受け、「包括同意」を認める要件として、同意撤回権を挙げる意見もある。そのため、従来の opt-

out や前述した「dynamic consent」における opt-out（この場合、同意としての opt-out ではなく、同意撤回を意味する）等を用いて同意撤回に対応することが考えられる。同意撤回の機会を患者に提供することは、患者に対して包括同意で不足した判断材料となる情報の提供や、それに基づく同意の有効性の向上という面で望ましいと考えられる。

包括同意において同意撤回のために「dynamic consent (opt-out)」を活用する場合には、DC 活用同意を合わせて取得することになるため、患者に対して伝えるべき内容の整理が必要となる。また、「dynamic consent (opt-out)」における本人確認等については、「dynamic consent」と同様に今後の検討課題である。

2. 医用画像データの匿名加工化による 2 次利用に関する実証研究

(1) 研究開発の目的

医療分野の Deep Neural Network (DNN) の技術を応用した人工知能の導入は活発に進められている。特に画像の診断補助の分野では研究ならびに実用化が進められている。基本的に DNN では大量の事例学習が必要になる。評価の定まった教師データである場合は、学習用データはそれほど多くなくても学習が可能になる場合もあるが、評価の確度が様々なリアル・ワールド・データを用いて学習する場合はかなりの量のデータが必要になる。DNN の学習データに臨床情報を提供した場合、当該患者に臨床的な直接の利益はなく、診療目的の利用とは言えない。したがって「医療介護事業者における個人情報の適切な取扱いに関するガイダンス」で明記されている黙示の同意で利用することはできず、明に同意を取得するか、匿名加工を行わなければならない。次世代医療基盤法に基づき通知によるオプトアウト処理でオプトアウトされなければ利用可能ではあるが、この場合も認定匿名加工医療情報作成事業者において匿名加工医療情報に加工する必要がある。

明に同意を得ることができれば問題ないが、臨床現場で学習対象となる情報が発生した時点で、利用目的が明確に定まっているとは限らない。むしろ定まっていない場合が普通であることが推測される。

したがって、個人情報保護法制における匿名加工および次世代医療基盤法における匿名加工医療情報への加工は極めて重要で、本報告ではこれを論ずる。

(2) 研究開発成果の概要

① 学習対象情報

臨床医学に DNN による人工知能を応用する場合、様々な種類の情報が対象となり得るが、本報告では画像情報とそれに付随する画像以外の診療情報と定義する。画像情報にも様々な種別があるが、臨床現場でよく扱われる画像情報として、DICOM 形式の画像および動画、RAW 画像情報、JPEG 画像情報、MPEG-4 動画、あるいはそれらの画像情報を含む PDF 形式を取り扱う。画像以外の診療情報も様々な形式がありうるが、単純なテキスト形式、HL7-CDA に準拠した XML 形式、HL7-FHIR で用いられる JSON 形式を対象とする。

(ア) DICOM

DICOM は米国放射線学会 (ACR) と米国電気工業会 (NEMA) が共同で策定した Digital Imaging and Communications in Medicine の略である。元々は米国の民間規格であるが、医用画像を扱う上での事実上の標準であり、日本では厚生労働省標準規格に指定されている。放射線科で扱う画像情報はほぼすべて DICOM 規格に準拠している。放射線科関連以外でも病理や内視鏡等、多くの画像情報が DICOM 準拠となっているが、まだ完全ではない。多彩な画像情報を扱う規格であるが、そのためにかなり柔軟で、拡張性に富んでいる。画像そのものの情報以外に属性情報が付加されており、タグと呼ばれている。対象となる患者情報や撮影機器 (モダリティ) や画像情報の形式を指定する必須タグは比較的明確に定義されているが、それ以外に情報を交換するもの同士が

定義可能なタグも導入可能となっている。必須タグの患者情報を示す部分に関しては DICOM 規格自体に匿名化に関する規則も存在する。画像自体の匿名化に関する規定はない。

(イ) JPEG

国際標準化団体である ISO/IEC の JTC 1 Joint Photographic Experts Group の作成した画像形式で、一般的に非可逆圧縮で多くのデジタルカメラで採用されている。画像形式以外に付加情報を添付することが許されていて、これを利用した Exif という付加情報規約が日本電子工業振興協会（JEIDA、後に一般社団法人電子情報技術産業協会：JEITA）で定められ、国際的に多くのデジタルカメラに採用されている。Exif は JPEG 以外に TIFF 等の画像規約にも適応可能で、撮影日や設営場所の GPS 情報等、多くの属性情報が格納されていることが多い。JPEG は DICOM でも画像形式として採用されている。

(https://ja.wikipedia.org/wiki/Joint_Photographic_Experts_Group)

(ウ) MPEG-4

動画を交換するための規格で ISO/IEC JT1 で作成保守されている。かなり広範な規格で、ある程度限定された形式である mp4 のようなサブ規格が使われることが多い。mp4 であっても、様々な情報が格納可能で、音声や静止画、テキストを格納できる。テキストの内容は任意である。MPEG-4 の内、フルハイビジョン規格の一部である MPEG-4 AVC/H.264 High Profile Level 4.1 は DICOM 規格に含まれている。

(エ) PDF

Adobe 社が提唱した印刷可能なドキュメント形式で、現在では ISO 標準となっている。画像情報は JPEG 等を採用しているが、文書情報にはフォント情報も含まれており、印刷あるいは画面表示イメージは一定程度の再現性がある。印刷可能な画像あるいは文書情報以外にかなりの属性情報を XMP 規格※に従って埋め込むことができ、アプリケーションさえ対応すれば、どのような情報でも埋め込み可能である。文書の電子署名は標準で対応されている。

※XMP (Extensible Metadata Platform)：Adobe 社によって提唱されたメタデータ埋め込み規格で ISO 標準となっている。主に PDF ファイルで使われるが、その他の画像情報にも適用可能である。

(オ) HL7-CDA

HL7 International が定めた XML 形式で、HL7 version3 Referential Information Model に基づいたもので、臨床現場で使われる文書形式のいくつかは対応しており、広く使われている。我が国では放射線等の診断レポート、診療情報提供書、特定健診情報等が対応している。文書構造が明確であり、URI による外部参照情報やテキスト変換された画像等のバイナリで埋め込まれた情報を除けば、不測な属性情報が入り込むことはない。

(カ) JSON

Javascript Object Notation の略で、IETF で規格化されている (RFC8259)。XML と表現の自由度はやや劣るものの、構文が単純でメッセージも小さく実装が容易である。WEB アプリケーションや IoT 機器で多用されており、HL7-FHIR でも CDA とともにサポートされている。匿名化を考える上では CDA と同様で、URI による外部参照情報やテキスト変換された画像等のバイナリで埋め込まれた情報を除けば、不測な属性情報が入り込むことはない。

② 画像情報自体の個人情報保護法に基づく匿名加工

個人情報保護法の匿名加工に関するガイドラインには画像情報に特別な言及はない。したがって個人情報保護に関する法律の第 2 条 第 9 項

この法律において「匿名加工情報」とは、次の各号に掲げる個人情報の区分に応じて当該各号に定める措置を講じて特定の個人を識別することができないよう個人情報を加工して得られる個人に関する情報であって、当該個人情報を復元することができないようにしたものをいう。

- A) 第 1 項第 1 号に該当する個人情報 当該個人情報に含まれる記述等の一部を削除すること（当該一部の記述等を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。
- B) 第 1 項第 2 号に該当する個人情報 当該個人情報に含まれる個人識別符号の全部を削除すること（当該個人識別符号を復元することのできる規則性を有しない方法により他の記述等に置き換えることを含む。）。

に準拠すれば良いことになる。

B)は画像情報では考えなくて良いので、A)を考えれば良いが、法第 2 条第 1 項第 1 号に該当する「当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別できるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）」である個人情報は医用画像情報の場合、外見上識別可能な画像情報にあたると考えられ、かなり広範囲が映り込んだ皮膚等の体表面画像、頭部の精緻な断層画像で立体再構成により、顔貌が識別可能な精度で再現される場合、外見上かなり特異な変化を移し込んだ、X 線撮影（高度な姿勢異常を来した脊柱に X 線画像等）あるいは特異な歯列が映り込んだ口腔パノラマ撮影画像等が考慮の対象になる。これらは一定程度識別性があると考えられる場合、一般には匿名加工が難しい。ただし、断層画像の場合は、表面近くの情報をあらかじめマスクできれば立体再構築しても外見上の識別可能性にはつながらず、画像情報としては匿名加工と言える。

一般に内視鏡画像、病理組織画像、眼底画像は、画像そのものとしては個人識別性はないと考えられる。

ただし、前述のように画像情報の形式には属性情報を付加できる形式が多く、このような情報の危険性は考慮しなければならない。これは次節で述べる。

③ 画像情報の次世代医療基盤法に基づく匿名加工医療情報

次世代医療基盤法の匿名加工医療情報に関するガイドラインには医用画像情報の留意点が記載されている。前述したような画像自体の個人識別性の検討に加えて、属性情報への配慮が規定されている。JPEG 等の Exif は基本的に削除すればよく、必要な属性があれば CDA や JSON 等で明示的に記載すればよい。問題は DICOM で患者情報が必須タグとして多くのモダリティで自動的に付加される。

ただし、DICOM 規格では DICOM 規格書 15 巻付属書 E では、匿名化を行う際にどのように処理するかを、以下の 6 つのアクションコードで示し、匿名化対象のデータ要素ごとに表に示している。

D - ゼロでない長さのダミーの値に置換する。値の形式は VR と一致させる。

Z - 長さをゼロにして値をセットしない、あるいは、ゼロでない長さのダミーの値と置換する。値の形式は VR と一致させる。

X - データ要素を削除する。

K - 保持する。（シーケンスでない要素は変更なし。シーケンス要素は消去する）

C - 消去する。識別情報を含まない値に置き換える。値は VR と一致させる。

U - インスタンスとして一貫性のある UID に置き換える。ここでの一貫性とは、例えば同じ検査のデータの場合、Study Instance UID が同じに保つことである。利活用目的によっては一貫性が必須のケースがあるため、注意が必要である。

DICOM 規格書 15 巻付属書 E の表 E. 1-1 から、主な属性データ要素の処理方法を表 1 に抜粋した。

表 1 DICOM 規格書 15 巻付属書 E の表 E. 1-1 から主な属性データ要素の処理方法の抜粋

属性	タグ値	リタイア	処理方法
患者の ID	(0010, 0020)		Z
患者の名前	(0010, 0010)		Z
患者の生年月日	(0010, 0030)		Z
患者の年齢	(0010, 1010)		X
患者の性別	(0010, 0040)		Z
検査日付	(0008, 0020)		Z
受付番号	(0008, 0050)		Z
検査 ID	(0020, 0010)		Z
検査内容	(0008, 1030)		X
患者コメント	(0010, 4000)		X
検査コメント	(0032, 4000)		X
収集コメント	(0018, 4000)	Y	X
オーバーレイデータ	(60xx, 3000)		X
施設名	(0008, 0080)		X/Z/D(Type3/Type2/Type1)
装置名前	(0008, 1010)		X/Z/D(Type3/Type2/Type1)
操作者名	(0008, 1070)		X/Z/D(Type3/Type2/Type1)
検査インスタンス UID	(0020, 000D)		U
SOP インスタンス UID	(0008, 0018)		U
プライベート属性	グループ番号が奇数		X

なお、匿名化を実施した場合、患者識別削除 (0012, 0062) を値“YES”で追加し、匿名化方法コードシーケンス (0012, 0064) に、DICOM 規格書 16 巻の CID7050 で定義されている匿名化方法のコードをセットするか、匿名化方法 (0012, 0063) に匿名化の手法について記述すること。

日本放射線学会から蓄積画像情報の一部の提供を受け、全データの DICOM タグの精査を行った。図 1 に示すように AOL の DICOM Viewer を用いて個々の画像を確認した。図 2 で個々の画像について表示し、Viewer の簡易表示機能で標準的なタグ情報を表示したものが図 3 である。

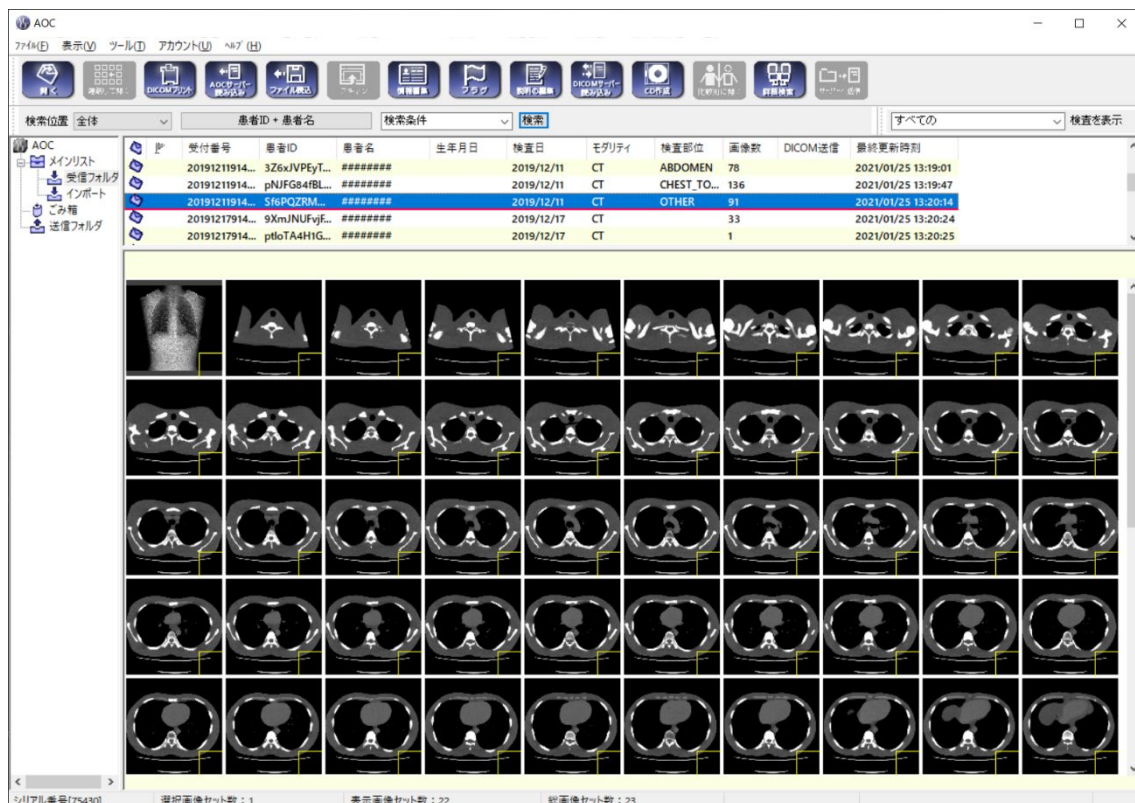


図 1 Dicom Viewer の初期画面 患者氏名はダミー化されている。

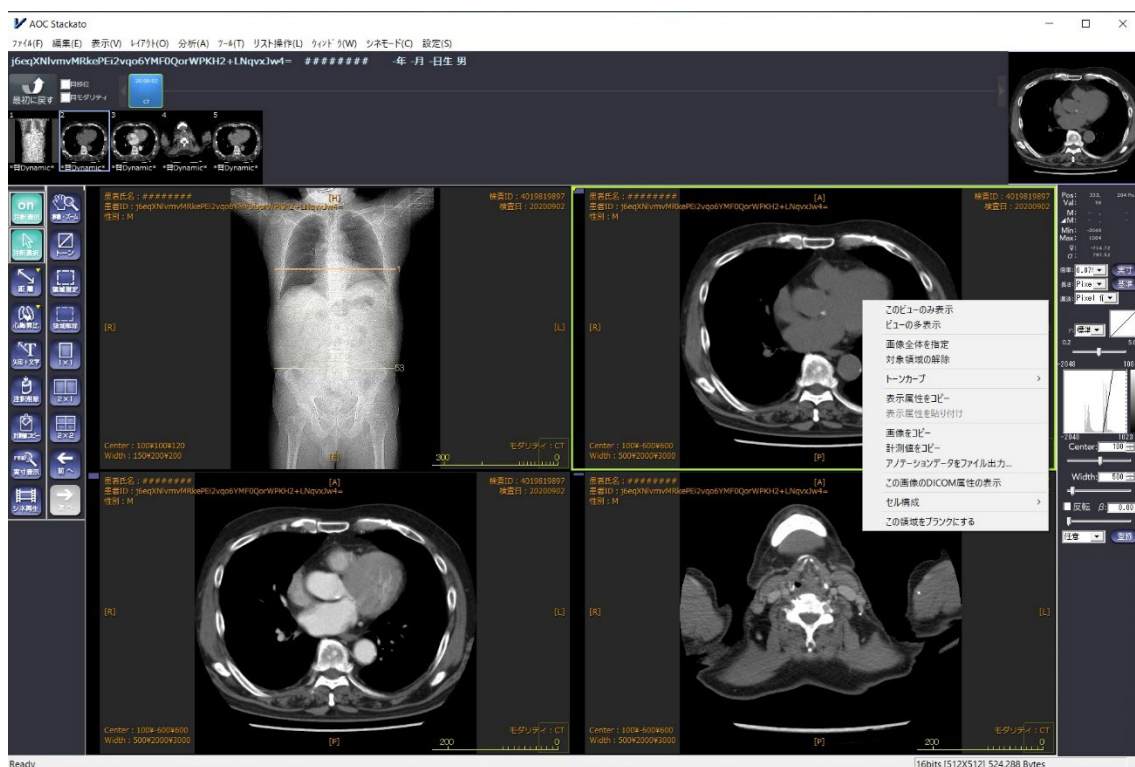


図 2 個別画像表示画面 DICOM タグの一部を表示できる。

×

DICOM属性編集

クリア

患者ID:	U9bvMDNCLMzLXGCeujD79PaML3QM=	☐
患者名:	#####	☐
生年月日:		☐
性別:	☒ 男 ☐ 女 ☐ その他 ☐ 元のまま	☐
検査日:	20191211	☐
検査時刻:	201310	☐
施設名:	4019819897	☐
照会医師名:		☐
検査ID:	4019819897	☐
受付番号:	2019121191405260	☐
検査記述:	JMID CT kyushu univ.011	☐
読影医師名:		☐
モダリティ:	CT	☐
検査部位:	OTHER	☐
患者位置:	HFS	☐
左右:		☐

画像セットから取得

コピー(C)

貼り付け(P)

クリア(L)

全てクリア(A)

☒ 全ての画像に適用する

☒ 検査・シリーズの構成を現在のままにする

☒ SOP Instance UIDを維持する

☐ 検査を1個にまとめ、シリーズは現在の構成のままにする

☐ 検査・シリーズを1個にまとめる

☐ 新しいスタディを生成する ☐ 新しいシリーズを生成する

☐ DICOMサーバーに送信 送信先

☐ 送信が完了した画像セットを、現在のフォルダから削除する

OK

キャンセル

図3 DICOM属性の内、標準的なものの表示と編集

氏名等、直接個人を特定する情報は削除またはダミー値に置き換えられていたが、日付情報等、静的情報および準識別子の一部はそのまま保存されていた。

表 2 は Viewer の機能で検出したすべてのタグ情報であるが、またその多くは個人識別性はないと考えられるが、プライベートタグはほぼすべて保存されていた。

表 2 タグ情報に一例（一部記載を省略）

0002,0000	: [UL] 184 (B8H)	// Group Length
0002,0001	: [OB] 00 01	// File Meta Information Version
0002,0002	: [UI] 1. 2. 840. 10008. 5. 1. 4. 1. 1. 2	// Media Storage SOP Class UID
0002,0003	: [UI] 1. 2. 392. 200036. 9116. 2. 6. 1 . .	// Media Storage SOP Instance UID
0002,0010	: [UI] 1. 2. 840. 10008. 1. 2. 1	// Transfer Syntax UID
0002,0012	: [UI] 1. 2. 840. 114302. 9	// Implementation Class UID
0002,0013	: [SH] 6. 7. 4 b3	// Implementation Version Name
0008,0005	: [CS] ISO_IR 13¥ISO 2022 IR 87	// Specific Character Set
0008,0008	: [CS] ORIGINAL¥PRIMARY¥AXIAL	// Image Type
0008,0016	: [UI] 1. 2. 840. 10008. 5. 1. 4. 1. 1. 2	// SOP Class UID
0008,0018	: [UI] 1. 2. 392. 200036. 9116. 2. 6. 1 . .	// SOP Instance UID
0008,0020	: [DA] 20200902	// Study Date
0008,0021	: [DA] 20200902	// Series Date
0008,0022	: [DA] 20200902	// Acquisition Date
0008,0023	: [DA] 19000101	// Content Date
0008,0030	: [TM] 110543	// Study Time
0008,0031	: [TM] 110709. 485	// Series Time
0008,0032	: [TM] 110720. 114	// Acquisition Time
0008,0033	: [TM] 000000. 000	// Content Time
0008,0050	: [SH] 2019090658926321	// Accession Number
0008,0060	: [CS] CT	// Modality
0008,0070	: [LO] Canon Medical Systems	// Manufacturer
0008,0080	: [LO] 4019819897	// Institution Name
0008,0090	: [PN]	// Referring Physician' s Name
0008,1010	: [SH] AquilionPRISM3	// Station Name
0008,1030	: [LO] 7C1610	// Study Description
0008,1032	: [SQ]	// Procedure Code Sequence
0008,0100	: [SH] 7C1610	// Code Value
0008,1090	: [LO] Aquilion ONE	// Manufacturer' s Model Name
0010,0010	: [PN] #####	// Patient' s Name
0010,0020	: [LO] j6eqXNlvmvMRkePEi2vqo . .	// Patient ID
0010,0030	: [DA]	// Patient' s Birth Date
0010,0040	: [CS] M	// Patient' s Sex
0010,1010	: [AS] 077Y	// Patient' s Age
0010,1030	: [DS] 070. 50	// Patient' s Weight
0018,0022	: [CS] HELICAL_CT	// Scan Options
0018,0050	: [DS] 5. 0	// Slice Thickness
0018,0060	: [DS] 120	// KVP
0018,0090	: [DS] 500. 00	// Data Collection Diameter
0018,1000	: [LO] #####	// Device Serial Number
0018,1010	: [LO] iRad-QA	// Secondary Capture Device ID
0018,1012	: [DA] 20200902	// Date of Secondary Capture
0018,1014	: [TM] 113029	// Time of Secondary Capture
0018,1016	: [LO] INFOCOM CORPORATION	// Secondary Capture Device Manufacturers
0018,1018	: [LO] iRad-QA	// Secondary Capture Device Manufacturer' s Model Name
0018,1019	: [LO] 1. 2. 0	// Secondary Capture Device Software Version(s)
0018,1020	: [LO] V10. 00JR008	// Software Version(s)
0018,1030	: [LO] *腎 Dynamic*	// Protocol Name
0018,1100	: [DS] 320. 312	// Reconstruction Diameter
0018,1110	: [DS] 1072. 55	// Distance Source to Detector
0018,1120	: [DS] +0. 0	// Gantry/Detector Tilt

0018,1130 : [DS] +143.00	// Table Height
0018,1140 : [CS] CW	// Rotation Direction
0018,1150 : [IS] 500	// Exposure Time
0018,1151 : [IS] 315	// X-Ray Tube Current
0018,1152 : [IS] 157	// Exposure
0018,1160 : [SH] LARGE	// Filter Type
0018,1170 : [IS] 37	// Generator Power
0018,1190 : [DS] 0.9#0.8	// Focal Spot(s)
0018,1210 : [SH] FC13	// Convolution Kernel
0018,5100 : [CS] HFS	// Patient Position
0018,7050 : [CS] ALUMINUM#COPPER	// Filter Material
0018,9302 : [CS] SPIRAL	// Acquisition Type
0018,9305 : [FD] 0.5	// Revolution Time
0018,9306 : [FD] 0.5	// Single Collimation Width
0018,9307 : [FD] 40	// Total Collimation Width
0018,9310 : [FD] -32.5	// Table Feed per Rotation
0018,9311 : [FD] 0.813	// Spiral Pitch Factor
0018,9318 : [FD] 0 0 -1066.5	// Reconstruction Target Center (Patient)
0018,9323 : [CS] 3D	// Exposure Modulation Type
0018,9324 : [FD] 4.16	// Estimated Dose Saving
0018,9327 : [FD] -130	// Table Position
0018,9334 : [CS] NO	// Fluoroscopy Flag
0018,9345 : [FD] 9.7	// CTDIvol
0020,000D : [UI] 1.2.392.200069.1.1092. . .	// Study Instance UID
0020,000E : [UI] 1.2.392.200036.9116.2. . .	// Series Instance UID
0020,0010 : [SH] 4019819897	// Study ID
0020,0011 : [IS] 2	// Series Number
0020,0012 : [IS] 2	// Acquisition Number
0020,0013 : [IS] 3	// Instance Number
0020,0020 : [CS] L#P	// Patient Orientation
0020,0032 : [DS] -159.843#-159.8434#-1066.50	// Image Position (Patient)
0020,0037 : [DS] 1.00000#0.00000#0.00000# . .	// Image Orientation (Patient)
0020,0052 : [UI] 1.2.392.200036.9116.2.6. . .	// Frame of Reference UID
0020,1040 : [LO] SN	// Position Reference Indicator
0020,1041 : [DS] -130.00	// Slice Location
0028,0002 : [US] 1(1H)	// Samples per Pixel
0028,0004 : [CS] MONOCHROME2	// Photometric Interpretation
0028,0010 : [US] 512(200H)	// Rows
0028,0011 : [US] 512(200H)	// Columns
0028,0030 : [DS] 0.625#0.625	// Pixel Spacing
0028,0100 : [US] 16(10H)	// Bits Allocated
0028,0101 : [US] 16(10H)	// Bits Stored
0028,0102 : [US] 15(FH)	// High Bit
0028,0103 : [US] 1(1H)	// Pixel Representation
0028,0120 : [SS] 63488(F800H)	// Pixel Padding Value
0028,1050 : [DS] 100#-600#600	// Window Center
0028,1051 : [DS] 500#2000#3000	// Window Width
0028,1052 : [DS] 0	// Rescale Intercept
0028,1053 : [DS] 1	// Rescale Slope
0040,0260 : [SQ]	// Performed Protocol Code Sequence
0008,0100 : [SH] 7C1610	// Code Value

AMED 事業の研究においては同意ベースの利活用であり、直接個人を特定するタグの処理で可と思われるが、個人情報法の匿名加工情報法や次世代医療基盤法の匿名加工医療情報とするには不十分と言わざるをえない。プライベートタグは提供された画像で見た限りは個人識別性はないと思われるが、おそらく 2 次利用に関しては不要と思われる、定義が曖昧な情報は削除すべきである。

後述の付随する臨床情報と合わせて検討する必要があるが、画像情報単体としても DICOM 画像の場合、DICOM 規格 15 巻付属書 E で既定されている匿名化は行うべきと考えられる。

④ 画像情報に付随した画像以外の臨床情報の匿名加工

画像以外の臨床情報は基本的にはテキスト情報で、個人情報保護法制上の匿名加工であっても機微性を考慮すれば次世代医療基盤法の匿名加工医療情報のガイドラインに準拠して匿名加工医療情報を作成することが推奨される。まず対象情報のリスク分析を行い、表 3 に示すように情報を分類し、必要に応じて表 4 に示す手法で加工し、リスク評価し、図 4 に示すように十分な安全性が確保されるまで、これを繰り返すべきである。

表 3 臨床情報の個人識別性の観点での属性分類

分類	概要
識別子	個人に直接紐づく情報（氏名、被保険者番号等）
準識別子	複数を組み合わせることで個人の特定が可能な情報（生年月日、住所、所属組織等） ※医療機関コードは準識別子に該当すると考えられる。
静的属性	不変性が高い情報（身長、血液型、アレルギー、受診日等の日付等） 障害等の外見的に特徴に関する情報 ※不変性の高い慢性疾病情報の取扱いは要検討
半静的属性	一定期間、普遍性がある情報（体重等） 疾病、処置、投薬等の情報はここに該当する想定
動的属性	常に変化する情報（検査値、食事、その他診療に関する情報等）

表 4 匿名加工の方針

分類	匿名加工方法の例
識別子	削除、もしくは非可逆な仮名化
準識別子	k-匿名性を満たすように一般化（生年月日→生年、住所→都道府県等）あるいはマイクロアグリゲーション データ項目削除を実施 医療機関コード等は属性（地理、規模等）を付加して特定できない形にコード変換
静的属性	数値はトップ・ボトムコーディング 一般化あるいはマイクロアグリゲーション 受診日等については、一般化やオフセット
半静的属性	数値はトップ・ボトムコーディング 機微な疾病等については必要ない場合は削除
動的属性	匿名加工不要であるが、必要に応じて数値はトップ・ボトムコーディング 異常値の重要性を考慮し、値の分布を見て上下数%を丸める等の加工を想定

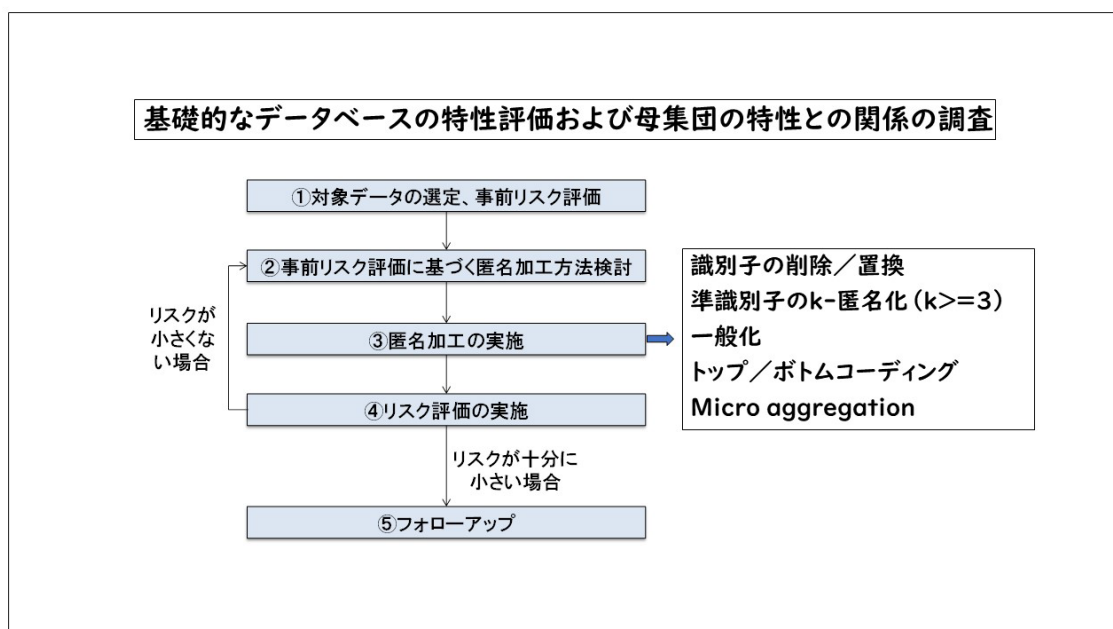


図4 リスク評価に基づいた匿名加工プロセス

⑤ それ以外の留意点

以上は通常の診療に置いて生成された画像情報およびそれに付随する臨床情報について述べた。つまり悪意を持った情報漏洩は考慮されていない。

様々なバイナリ情報で、理論的には steganography と呼ばれる手法で、人には検知しにくい形で、隠れた情報を紛れ込ませることが可能である。悪意ではなくても、転々流通する情報のトレースをするために、埋め込まれることもあるが、この手法で個人情報情報が埋め込まれている場合、その事実を知らない限り対応が難しい。個人情報情報が含まれていない場合でも DNN による学習にも影響を与えることが予想され、避ける必要がある。画像を取り扱う医療機関で意図的に処理しない限り紛れ込むことはないため、関係者間で合意し、必要に応じて契約等で担保することが望まれる。

⑥ 考察

医療分野の Deep Neural Network (DNN) の技術を応用した人工知能の導入は活発に進められており、その学習用情報として、画像情報を中心とした臨床情報の匿名加工について論じた。画像情報自体は身体の外見上の特徴を示す場合以外は大きな問題はなく、特殊な身体の外見的特徴を来す疾患以外では、頭部の CT, MRI 画像の立体再構成による顔貌の識別性や歯列パノラマ撮影画像以外は大きな問題にならない。しかし医用画像情報には一般に多くの属性情報が付加されており、「③」4章で例示したように DICOM 画像では 100 以上のタグ情報が付加されている場合もある。タグ情報で患者氏名等のように直接個人識別性を示す情報が限定的ではあるが、日付情報等、他の情報との連結により、個人識別性を増加させる情報も多い。人工知能の学習データとして用いる場合、これらの情報の多くは不要と考えられるため、必要性を勘案した上で、不要であれば削除、あるいは関連性のないダミー値に置換することが求められる。また悪意のある Steganography 等の混入を防止するためには、契約または覚書の締結も求められる。画像だけで、診断等が学習できるわけではなく、付随する臨床情報がなければ学習データとはならないが、非画像臨床情報に関しては、個人情報保護法やそのガイドラインには具体的な方法や基準の記載はなく、次世代医療基盤法のスキームを用いるか用いないかに関わらず、匿名加工医療情報の関する指針にしたがって、リスクベースで匿名加工を進める必要があると考えられる。

⑦ 結語

画像情報を含む臨床情報の2次利用における匿名加工に関して論じた。同意ベースの学術研究では倫理審査を経た研究手法で処理すれば良いが、学術研究を離れて商用利用等の学習データとして利用する場合は、基本的には次世代医療基盤法の匿名加工医療情報に関する指針に準拠する必要があることを示し、その実例を提示した。

参考文献

1. 個人情報の保護に関する法律, https://www.ppc.go.jp/files/pdf/201212_personal_law.pdf (2021年3月26日確認)
2. 個人情報の保護に関する法律についてのガイドライン(匿名加工情報編), <https://www.ppc.go.jp/files/pdf/guidelines04.pdf> (2021年3月26日確認)
3. 医療分野の研究開発に資するための匿名加工医療情報に関する法律(次世代医療基盤法), <https://elaws.e-gov.go.jp/document?lawid=429AC00000000028>, (2021年3月26日確認)
4. 医療分野の研究開発に資するための匿名加工医療情報に関する法律についてのガイドライン, https://www.kantei.go.jp/jp/singi/kenkouiryou/jisedai_kiban/pdf/h3005_guideline.pdf (2021年3月26日確認)
5. DICOM企画書, https://www.jira-net.or.jp/dicom/dicom_data_02_01.html (2021年3月26日確認)