



EXECUTIVE SUMMARY FOR
THE DIGITAL HEATH
ADVISORY COMMITTEE
MEETING

Total Product
Lifecycle
Considerations for
Generative AI-
Enabled Devices

November 20 – 21, 2024

DHAC Executive Summary - GenAI 対応デバイス製品トータルライフサイクルに関する考慮事項
Digital Health Advisory Committee (DHAC)

Note) This document is for purpose of providing reference information for the use of the original document, and does not have the same effect as the original document.

If you have any doubts about this document, please refer to the original Regulatory Authority (FDA) information.

HiroPharmaConsulting assumes no responsibility for any inconvenience caused by the use of this document.

*Only valid the "EXECUTIVE SUMMARY FOR THE DIGITAL HEATH ADVISORY COMMITTEE MEETING
Total Product Lifecycle Considerations for Generative AI Enabled Devices: November 20 – 21, 2024"*



DHAC Executive Summary - GenAI 対応デバイス製品トータルライフサイクルに関する考慮事項

TABLE OF CONTENTS

はじめに	3
GenAI のしくみ	3
ヘルスケアにおける GenAI アプリケーションの可能性	6
AI に対するライフサイクル管理アプローチ	7
FDA による GenAI 対応デバイスの監視	9
GenAI 対応製品に適用されるデジタルヘルスポリシーとプレマーケットパス	11
GenAI 対応機器と TPLC 規制アプローチの課題	12
GenAI 対応デバイスにリスクベースのアプローチを適用する際の考慮事項	13
FDA のデジタルヘルスポリシーへの影響	14
デバイス分類への影響	15
TPLC を介した GenAI 対応デバイスの有効な科学的証拠に関する考慮事項	17
現在の市販前のエビデンスの必要性	18
市販前のエビデンスのニーズに応える新たな方法論	19
市販後の新しい証拠の必要性	19
サマリー	20
パネル質問	21
Appendix	23
Glossary	25
References	30

免責事項:本エグゼクティブサマリーは議論のみを目的としており、ドラフトまたは最終的なガイダンスを示すものではありません。生成型人工知能対応デバイスの規制に関する政策変更を提案または実施することを意図したものではありません。また、ここで引用されている参考文献は情報提供のみを目的としており、推奨として解釈されるべきではありません。

はじめに

ヘルスケア分野における生成型人工知能 (GenAI) ツールへの関心は急速に拡大していますが、FDA の管轄下にある可能性のある医療機器を含む (ただしこれに限定されない) GenAI 対応製品を規制するアプローチについては、未解決の問題が残っています。このドキュメントでは、「GenAI 対応機器 (GenAI-enabled device)」という用語を使用して機器を指します。この用語は、連邦食品医薬品化粧品法 (FD&C 法) のセクション 201 (h) で定義されており、GenAI のメソッドまたはモデルが機器の出力または機能に不可欠である場合です。人工知能 (AI) 対応製品と同様に、GenAI 対応製品の機能は、患者と公衆衛生に独自の利点を提供する可能性があります。FDA が対処するための新たな規制の複雑さももたらします。すべての医療機器と同様に、FDA の規制監督は、デバイスの定義を満たす GenAI 対応製品に適用されます。このような監督は、製品の意図された用途と技術的特性を考慮したリスクベースのもので、さらに、FDA は長い間、AI 対応デバイスを含む医療機器の監督に対するトータルプロダクトライフサイクル (TPLC) アプローチを推進しており、現在の当局を使用してこれらのデバイスの規制アプローチを開発することにコミットしているだけでなく、新しい当局を必要とする可能性のあるオプションを模索しています。

このコミットメントは、デバイスの使用期間中にこれまで以上に迅速かつ頻繁に反復することを意図した技術を組み込んだ医療機器にとって、ますます重要になっています。GenAI 対応製品は、同じ入力に対して可変出力を提供することを意図することができ、迅速かつ頻繁に変更されることを意図したモデルに頻繁に依存する可能性があり、それ自体が医療機器ではないモデルを照会する可能性があります。TPLC アプローチは、将来の安全で効果的な GenAI 対応医療機器の管理にとって重要であり続ける可能性があります。このエグゼクティブサマリーでは、GenAI 対応機器の監督に対する FDA のアプローチに焦点を当てています。これは、AI 対応機器全般に対する FDA のアプローチと多くの類似点を共有しています。このエグゼクティブサマリーでは、GenAI のリスク (その一部は AI に広く適用される可能性があります) と、TPLC 全体での AI および GenAI 対応機器の規制に対する現在の課題についても説明しています。

GenAI のしくみ

GenAI 以前、開発者は医療機器に AI と機械学習 (ML) モデルを組み込んでいました。これらの非生成的 AI/ML モデルは、画像のセグメンテーション、¹ 分類、² バイオマーカーの抽出、^{3, 4} リスク予測⁵ などのタスクを実行するために使用されてきました。非生成的 AI/ML モデルは、完全に説明できない場合や透明性がない場合でも、一般に再現性があると考えられています。

GenAI は、入力データの構造と特性を模倣して派生した合成コンテンツを生成する AI モデルのクラスを指し、画像、ビデオ、オーディオ、テキスト、およびデジタルコンテンツを含めることができます。⁶ 非生成的 AI/ML と同様に、GenAI はデータからパターンを学習します。ただし、非生成的 AI/ML とは異なり、GenAI モデルは一般に、正確な予測を行うためのパターンを特定することを主目的とするのではなく、学習元のデータに似た新しいデータを作成することを意図しています。GenAI モデルは入力データを分析し、トレーニングデータに明示的に表示されていない可能性のあるコンテキストに適した出力を生成できます。

重要なのは、GenAI モデルは、開発中に人間の開発者がデータセットの内容についてすべてを知ることができないほど大規模なデータセットで開発されることが多いということです。他の AI/ML モデルの開発に使用されるデータセットとは対照的に、GenAI モデル開発のデータセットは意図的に広くことができ、最初は特定のタスクに合わせて調整されていない可能性があります。GenAI モデルは、開発データ内の要素間の非常に複雑な関係を派生させます。これは、1 つまたは複数のタスクに適用してさらに最適化することができます。GenAI モデルは、特定の関心領域に適用できる出力を生成する場合がありますが、派生した関係は、開発データに含まれるさまざまな単語、ピクセル、またはその他の言語要素の間にあります。したがって、出力には、派生した関係に関する確率的な予測が反映されません。これは、特定の関心領域に必ずしも関連しない開発データセット内の情報によって通知される場合があります。複数のタスクに広く適用できる GenAI モデルは、多くの場合、基盤モデルと呼ばれます。

現在の基本モデルは、通常、個々の製品に対して作成されていません。また、通常は、連邦食品、医薬品、化粧品法 (FD&C 法) のセクション 201 (h) で定義されているように、デバイスとして使用することを目的としていません。通常、基本モデルの開発者とは異なるエンティティであるアプリケーション開発者は、デバイス定義を満たす可能性のあるアプリケーションなど、さまざまなアプリケーションに単一の基本モデルを適合させることができます。

¹ Seo H, Badiei Khuzani M, Vasudevan V, Huang C, Ren H, Xiao R, Jia X, Xing L. 生体画像セグメンテーションのための機械学習技術: 技術的側面の概要と最先端アプリケーションの紹介。医学物理学。47 (5) :e148-e167. <https://doi.org/10.1002%2Fmp.13649>

² Islam KR, Prithula J, Kumar J, Tan TL, Reaz MBI, Sumon MSI, Chowdhury MEH. 電子カルテを用いた敗血症の機械学習に基づく早期予測: システマティックレビュー。Journal of Clinical Medicine。12(17):5658. <https://doi.org/10.3390/icm12175658>

³ Mansur A, Vrionis A, Charles JP, Hancel K, Panagides JC, Moloudi F, Iqbal S, Daye D. 肝細胞癌のバイオマーカーの検出と実装における人工知能の役割: 展望と機会。癌 (バーゼル)。15(11):2928. <https://doi.org/10.3390%2Fcancers15112928>

⁴ Dillon A, Singh A, Bhalla VK. A Systematic Review on Biomarker Identification for Cancer Diagnosis and Prognosis in Multi-omics: From Computational Needs to Machine Learning and Deep Learning. Archives of Computational Methods in Engineering。30:917-949. <https://doi.org/10.1007/s11831-022-09821-9>

⁵ Alballa N, Al-Turaiki I. COVID-19 diagnosis, mortality, and severity risk prediction における機械学習アプローチ: レビュー。Informatics in Medicine Unlocked。24:100564. <https://doi.org/10.1016/j.imu.2021.100564>

⁶ 人工知能の安全、安全、信頼できる開発と使用。2023 年 10 月 30 日の大統領令 14110、セクション 3 (p)。88 FR 75191。

特定のタスクのために基盤モデルをトレーニングするために使用される方法は、研究者がこの分野に焦点を当てるにつれて、急速に出現し、変化しています。一部の研究者は、セグメンテーションや分類など、現在非生成的な AI/ML モデルで達成される可能性のあるタスクと同様のタスクに GenAI モデルを使用することを模索しています。一方、他の研究者は、広範な画像分析や広範な分野にわたる臨床診断の提供など、新しい機能のために AI または GenAI モデルを使用することを模索しています。AI または GenAI モデルが医療情報を使用して特別にトレーニングされている場合でも、その情報は医療部門内の多くの医療分野やソースを横断する可能性があります。したがって、モデルは特定の種類の出力を生成するだけでなく、いくつかの臨床分野にわたってそれを行うことができる可能性があります。たとえば、生検画像で腫瘍組織を特定するようにトレーニングされたモデルは、いくつかの異なる組織の種類や画像処理方法に使用できる可能性があります。

GenAI モデルを含む AI モデルはまた、ハルシネーション（Hallucinations）、すなわち、ユーザーのプロンプトを満たすというプログラムされた目的を満たすために誤った、または偽のコンテンツを生成することを示す可能性があります。⁷ GenAI モデルに典型的なモデルアーキテクチャやデータの大規模なコーパスを含むモデルの複雑さは、このようなハルシネーション（Hallucinations）を引き起こす要因となり得ます。⁸ したがって、GenAI の潜在的な注目すべき利点は、さまざまな異なるデータタイプから特定の関心領域に適用可能な出力を生成できること、または広範なタスクに関連する出力を生成できることである一方で、GenAI は、他の技術に当てはまるように、アプリケーションごとに異なるレベルのリスク管理を必要とする潜在的なリスクを提示する可能性もあります。たとえば、ハルシネーション（Hallucinations）、特にユーザーにとって本物の出力であるように見えるハルシネーション（Hallucinations）は、非常に正確で真実な情報が重要な特定のヘルスケアアプリケーションで大きな課題となる可能性があります。

医療機器の場合、GenAI を使用して、患者や医療専門家などの複雑な医療データから、可能性のある臨床診断、治療ソリューション、および新しい関連性を特定するのに役立つコンテンツを生成することは、ヘルスケアに大きなメリットをもたらす可能性があります。しかし、多様で新しい複雑なタスクに取り組む GenAI のこの同じ能力が、デバイスの出力の限界に関する不確実性の一因となる場合があります。制御が不十分な場合、この不確実性は、デバイスの使用目的の範囲を確認することが困難になる可能性があり、FDA による GenAI 対応デバイスの規制に課題をもたらす可能性があります。例えば、デバイスの使用目的が明確に定義されていない場合、FDA がリスクベースのアプローチを適用して GenAI 対応デバイスを分類し、適用される規制要件を決定することは困難になる可能性があります。さらに、TPLC 全体で GenAI 対応デバイスの安全性と有効性を確保するために、適切な評価方法とリスクコントロールが利用可能であることが重要になります。AI 対応機器の安全性と有効性を確保するために利用可能な評価方法やリスク管理の一部は、GenAI 対応機器にも利用できるかもしれないが、GenAI 対応機器への関心とニーズは、新しい評価方法やリスク管理の開発を加速させる可能性もあります。

⁷ National Institute of Standards and Technology. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. <https://airc.nist.gov/docs/NIST.AI.600-1.GenAI-Profile.ipd.pdf>

⁸ Reddy GP, Pavan Kumar YV, and Prakash KP. Hallucinations in Large Language Models (LLMs). 2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream). 2024: 1-6. <https://doi.org/10.1109/eStream61684.2024.10542617>

すべてのデバイスと同様に、FDA は、安全性と有効性を合理的に保証するために、製品の使用目的と技術的特性を考慮したリスクベースのアプローチに従います。

このエグゼクティブサマリーでは、「GenAI 対応デバイス」という用語を使用して、その用語が FD&C 法のセクション 201 (h) で定義されているデバイスを参照しています。GenAI メソッドまたはモデルは、デバイスの出力または機能に不可欠です。別の言い方をすると、GenAI メソッドまたはモデルは、デバイスの主要な機能において重要な役割を果たすか、またはその出力を直接可能にします。

ヘルスケアにおける GenAI アプリケーションの可能性

ヘルスケアにおける GenAI の現在および潜在的な実装には、幅広いものがあります。これらの実装の中には、医療機器である製品を形成するものもあれば、FDA の管轄外のものもあります。これらの実装には、臨床文書の作成や保険の事前承認の取得などの管理機能が含まれる可能性があります。⁹ GenAI モデルは、医療訓練やシミュレーションを容易にするためにも使用される可能性があります。これらのモデルは、リアルタイムで適応する現実的な患者シミュレーションを作成する能力を備えており、医療専門家が制御された環境で診療を行うのに役立ちます。¹⁰ 研究では、臨床診断を生成する際の GenAI モデルの有用性や、デバイスの定義を満たす可能性のあるチャットボットの使用などによる精神衛生状態の診断と治療の可能性についても説明しています。¹¹ 多くの GenAI モデルの基礎は、複数のタスクに幅広く適用できる基礎モデルであるため、生検画像上の腫瘍組織を特定するように訓練された GenAI モデルは、そのサンプルの画像と病理報告について議論するチャットボット機能を提供するように訓練された GenAI モデルと同じ基礎モデルを使用できます。GenAI モデルは、非生成的な AI/ML モデルで一般的に使用される構造化データ形式ではなく、ユーザーからの自由なテキストまたはマルチモーダル入力を使用することもできます。この機能により、GenAI 対応製品は、その使用のために、より多くの可能な入力を受け入れることができます。たとえば、GenAI 対応チャットボットは、個別の検査データとバイタルサイン、患者固有の医療データ、患者のシナリオの説明を記述するフリーテキストまたはマルチモーダル入力を使用して、開発中には遭遇しなかった可能性のある入力の一意の組み合わせから鑑別診断を生成することができます。

以下で詳しく説明するように、GenAI への関心の急速な高まりは、FDA の法律や規制に課題をもたらす可能性があります。将来的には、AI および GenAI 対応デバイスの評価に対する特定のセーフガードと新しいアプローチが開発される可能性があり、FDA がこれらのデバイスの安全性と有効性を確保するために役立つ可能性があります。さらに、FDA は長い間、デバイスの安全性と有効性を保証する TPLC アプローチを提唱してきましたが、AI および GenAI 対応デバイスは、そのようなアプローチの必要性を前面に押し出し、ライフサイクルの考慮事項を効果的に実装および評価する方法について重要な問題を提起しています。これらの考慮事項は、患者を保護し、公衆衛生に利益をもたらすという共通の目標を持つ規制当局と開発者の両方にとって重要です。

⁹ Mesko B, Topol EJ. 医療における大規模言語モデル (または生成型 AI) の規制監督の必要性. NPJ Digital Medicine. 6, 120. <https://doi.org/10.1038/s41746-023-00873-0>

¹⁰ Sardesai N, Russo P, Martin J, Sardesai A. 生成型会話型人工知能を利用した模擬患者遭遇の作成: 麻酔訓練のためのパイロット研究. Postgraduate Medical Journal. 100 (1882) :237-241. <https://doi.org/10.1093/postmi/ggad137>

¹¹ Sezgin E, McKay I. 行動的健康と生成 AI: 治療法と患者ケアの将来の展望. NPJ メンタルヘルス. 3(25). <https://doi.org/10.1038/s44184-024-00067-w>



AI に対するライフサイクル管理アプローチ

1960年代以来、ライフサイクル管理 (LCM) はソフトウェアを管理するための構造化されたプロセスであり、使用期間中に頻繁に更新されることが意図されている GenAI 対応ソフトウェアを含む AI 対応ソフトウェアにとって、今日でも特に重要です。最新のソフトウェア開発ライフサイクル (SDLC) は LCM の原則を具体化し、ソフトウェアの計画、設計、実装、テスト、統合、展開、保守、最終的には廃止するための構造化されたプロセスを提供します。GenAI 対応および AI 対応ソフトウェアの広範なこの概念に基づいて、AI ライフサイクルに合わせて調整されたプロセスでは、データの収集と準備、トレーニングと検証、評価とテスト、展開と統合、AI モデルの監視など、AI ソフトウェア開発に固有のソフトウェア開発の側面を考慮することができます (図 1)。以下で説明するように、これらのプロセスステップの一部(例えば、評価、テスト、モニタリング)は、GenAI モデルを含む AI モデルにとって特に重要な場合があります。AI に対するライフサイクル管理アプローチを検討することは、パフォーマンス中心の考慮事項(正確性や堅牢性など)だけでなく、GenAI 対応製品を含む AI 対応製品にとって非常に重要な可能性のあるより広範な考慮事項(倫理、公平性、プライバシーへの配慮など)にも役立ちます。

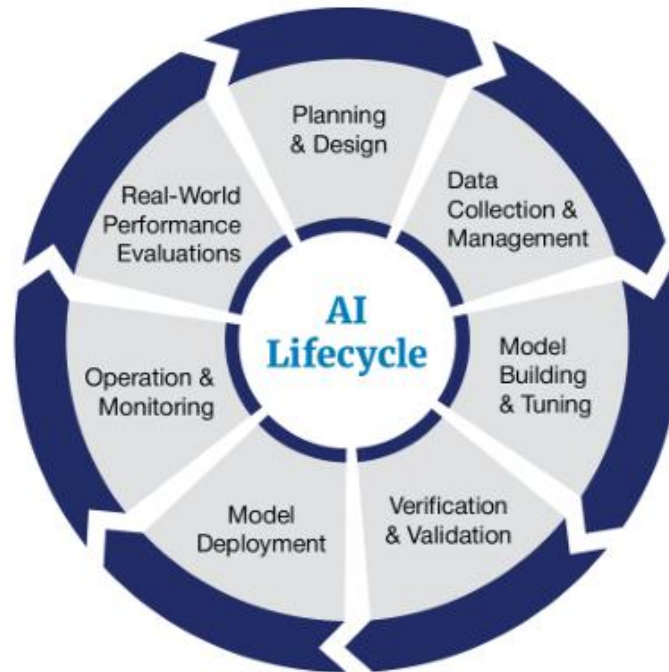


Figure 1. AI Lifecycle

図 1 に示す一般的な AI ライフサイクルの例は、次のように説明できます。

- **計画と設計**- このフェーズでは、通常、AI プロジェクトの目標、スコープ、および要件を定義し、使用するアーキテクチャとアルゴリズムを設計します。
- **データの収集と管理**- このフェーズでは、通常、AI モデルのトレーニングとテストに必要なデータを収集して整理し、品質と関連性を確保します。
- **モデルの構築とチューニング**- このフェーズでは、通常、適切なアルゴリズムの選択、特徴エンジニアリング、およびパフォーマンスの最適化など、収集されたデータを使用して AI モデルを開発および改良します。
- **検証と検証**- このフェーズでは、通常、AI モデルの精度と信頼性をテストと検証のプロセスを通じて評価し、目的の標準と目的を満たしていることを確認します。
- **モデルの展開**- このフェーズでは、通常、トレーニング済みの AI モデルを目的の環境またはアプリケーションに統合し、実際の使用にアクセスできるようにします。
- **運用と監視**- このフェーズでは、通常、AI 対応製品をテストして、展開後の運用環境で円滑かつ確実に動作することを確認します。これには、通常、監視などのタスクが含まれます。

トラブルシューティング、バグ修正、システムの運用期間中に発生する問題に対処するために必要な更新または機能強化を行います。このフェーズには、通常、発生した問題への対処を含む、モデルの機能のメンテナンスも含まれます。

- **実際のパフォーマンス評価-** このフェーズでは、通常、AI 対応製品のパフォーマンスが実際の条件で評価され、意図された目的とユーザー要件をどの程度満たしているかを判断します。これには、通常、正確性、効率性、スケーラビリティ、堅牢性、ユーザー満足度など、システムのパフォーマンスのさまざまな側面の評価が含まれます。

一般に、GenAI 対応デバイス、および広く AI 対応デバイスの AI ライフサイクルを考慮することは、製造業者が TPLC 全体でデバイスを管理するための重要な方法の 1 つです。さらに、AI ライフサイクルは、TPLC 全体でこれらの新しいテクノロジーを適切に管理するために、新しい技術、アプローチ、または標準が必要になる可能性がある場所を特定するための有用なモデルとして使用できます。以下でさらに説明するように、これは、これらのデバイスのライフサイクルを通じて継続的な安全性と有効性を確保するために必要な有効な科学的証拠を理解し、これらのニーズが他のデバイスと異なる可能性があるかどうかを理解するために重要です。たとえば、AI ライフサイクルの実際のパフォーマンス評価フェーズの一部として含まれる考慮事項は、製造業者が AI 対応デバイスの市販後のパフォーマンスを効果的に監視するのに役立ちます。これは、GenAI 対応デバイスにとって特に重要です。GenAI 対応デバイスは、新しいコンテンツを生成するという本質的な性質を考慮すると、長期にわたって継続的な安全性とパフォーマンスを確保するための監視計画から恩恵を受けるか、または監視計画を必要とする可能性があります。そのため、AI ライフサイクルとその管理を検討することは、GenAI 対応デバイス (GenAI 対応デバイスを含む) の管理に関する現在の課題の一部に対処するのに役立つ可能性があります。

FDA による GenAI 対応デバイスの監視

すべての医療機器と同様に、FDA の規制上の監視は、デバイスの定義を満たす GenAI 対応製品に適用され、製品の意図された用途と技術的特性を考慮したリスクベースのアプローチに従います。FDA の法律と規制のいくつかの側面の GenAI 対応製品への適用は簡単ですが、FDA の現在の法律と規制の適用がより困難な側面もあります。たとえば、多くの GenAI 対応製品の性質が進化しているため、その意図された用途が FDA の現在のデジタルヘルスポリシーの範囲内でどのように調整されるかを判断するのは困難な場合があります。以下では、これらのトピックを調査して、GenAI 対応デバイスの監視に対する FDA の潜在的なアプローチと、未解決の不確実性のいくつかについて説明します。

「装置 (device)」という用語は、FD&C 法第 201 条 (h) において、器具、装置、器具、機械、装置、インプラント、インビトロ試薬、又はその他の類似若しくは関連物品 (構成部品、部品又は付属品を含む) として定義されており、とりわけ、疾患若しくはその他の状態の診断、又は疾患の治療、緩和、治療若しくは予防における使用を意図しています。¹²

¹² 連邦食品医薬品化粧品法のセクション 201 (h) を参照。

特定のソフトウェア機能は、FD&C 法のセクション 520 (o) によってデバイスの定義から明確に除外されています。これには、たとえば、医療施設の管理サポートを目的とした特定のソフトウェア機能や、健康的なライフスタイルを維持または奨励することを目的としたソフトウェア機能が含まれます。また、病気や状態とは無関係です。その他のソフトウェア機能も、FD&C 法のセクション 520 (o) によって明確に除外されていない場合でも、デバイスの定義を満たしていないため、デバイスではありません。製品がデバイスの定義を満たしているかどうかを判断するには、まず製品の使用目的を特定することが重要です。¹³ 使用目的は、製品またはその機能の一般的な目的として定義されます。これは、製品の表示に法的に責任を負う者（またはその代理人）の客観的な意図であり、表現、製品のデザインまたは構成、または製品の配布を取り巻く状況によって示される場合があります。¹⁴

前述のように、GenAI には固有の特性があり、適切なリスク管理が行われていない製品設計の一部として、製品の出力に不確実性をもたらし、製品の意図された用途の範囲を判断することを困難にする可能性があります。したがって、それがデバイスの定義を満たし、FDA のデバイス規制の監視の焦点となっているかどうかを判断することが困難になります。GenAI の固有の特性は、FDA の既存のポリシーが GenAI 対応製品にどのように適用されるかにも影響を与え、不確実性をもたらす可能性があります。例えば、FDA のガイダンス「[Changes to Existing Medical Software Policies Resulting from Section 3060 of the 21st Century Cures Act](#)」および「[Policy for Device Software Functions and Mobile Medical Applications](#)」に概説されているポリシーは、デバイスの定義から除外されている特定のソフトウェア機能、または医療機器であり、デバイスが意図したとおりに機能しない場合に患者の安全にリスクをもたらす可能性のあるソフトウェア機能に対する FDA のアプローチを説明しています。現在の形では、これらのポリシーは、FDA が低リスクであると理解している特定の意図された用途および製品を対象としており、GenAI 対応製品がこれらのポリシーの範囲内にあるかどうかは不確実性があります。したがって、あらゆる新しい技術と同様に、そのような用途に GenAI を実装することの利点とリスクを考慮することが重要です。

FDA の規制上の監視の対象となる GenAI 対応機器については、これらの機器の固有の特性が、その分類や、FDA の審査官がこれらの機器の安全性と有効性を合理的に評価するために市販申請で必要とする可能性のある証拠の種類にも影響を与える可能性があります。すべての機器と同様に、新しい意図された用途は、既存の規制上のアプローチ（例えば、デノボ：De Novo または市販前承認経路）の対象となります。しかし、医療の広範な領域にわたる診断、新しい情報と画像の生成、高度な自動化の利用、医療専門家の役割が縮小された、または医療専門家からの監視がない患者への医療サービスの直接提供など、新しい意図された用途を導入する GenAI 対応機器は、その評価のための新しい種類の有効な科学的証拠またはアプローチの必要性を生み出し、FDA は医療機器に適用される GenAI を広く理解するための新しい評価戦略を必要とします。

¹³ [製品が医療機器であるかどうかを判断する方法](#)に関する FDA の Web サイトを参照。

¹⁴ 21 CFR 801.4 を参照。

GenAI 対応製品に適用されるデジタルヘルスポリシーとプレマーケットパス

デバイスソフトウェア機能およびモバイル医療アプリケーションに関する FDA のガイダンスポリシーに記載されているように、FDA は、医療機器であり、デバイスが意図したとおりに機能しない場合に患者の安全性にリスクをもたらす可能性のあるソフトウェア機能に規制上の監視を適用することを意図しています。大まかに言えば、この同じアプローチを GenAI 対応製品に適用するのが妥当かもしれません。たとえば、すべてのソフトウェアと同様に、GenAI 対応製品は、デバイスではなく FDA の規制範囲内にはないものから、デバイスであり FDA によって規制されているものまでの範囲に存在するようです。この範囲の GenAI 対応製品には、次のようなものが含まれる可能性があります。

- デバイスの定義を満たさない可能性のある特定の GenAI 対応製品;
- デバイスの定義を満たす可能性があるが、FDA が執行裁量権を行使しようとしている製品である特定の GenAI 対応製品;および
- デバイスの定義を満たし、FDA が規制監督に注力しようとしているデバイスである特定の GenAI 対応製品。

特定の GenAI 対応製品は、デバイスの定義を満たしていない可能性があります。たとえば、FDA のデバイスソフトウェア機能およびモバイル医療アプリケーションに関するガイダンスポリシーの例と同様に、医療従事者が医療トレーニングの教育ツールとして使用したり、以前に受けたトレーニングを強化したりすることを目的とした特定の GenAI 対応製品は、デバイスではない可能性があります。その他の GenAI 対応製品は、FD&C 法のセクション 520 (o) に従ってデバイスの定義から除外されている場合、デバイスの定義を満たしていない可能性があります。たとえば、医療施設の管理サポートのみを目的とした特定の GenAI 対応製品や、健康的なライフスタイルを維持または奨励することを目的とした製品で、疾患や状態とは無関係のものは、デバイスではない可能性があります。

その他の GenAI 対応製品はデバイスの定義を満たしている可能性があります。FDA は、公衆に対するリスクが低い場合、これらのデバイスに対して執行裁量権を行使することを意図している可能性があります(FDA は、現時点では、FD&C 法の下での要件を強制するつもりはないことを意味します)。たとえば、医療従事者のための単純なタスクを自動化する特定の GenAI 対応製品や、慢性疾患の患者のための健康に関するリマインダーや追跡ツールを生成する特定の GenAI 対応製品は、デバイスである可能性があります。FDA はこれらのデバイスに対して執行裁量権を行使することを意図している可能性があります。

最後に、多くの GenAI 対応製品は、デバイスとして FDA の規制監督の対象となる可能性があります。これは、ソフトウェアがデバイスの定義を満たしており、ソフトウェアが意図したとおりに機能しない場合、その機能が患者の安全性にリスクをもたらす可能性があるためです。例えば、FDA の「デバイスソフトウェア機能およびモバイル医療アプリケーションに関するガイダンスポリシー」の例と同様に、患者固有の分析を実行し、疾患や状態の診断、治療、緩和、治癒、または予防に使用するために特定の出力または指示を医療従事者に提供する GenAI 対応製品、または患者固有の分析を実行し、患者、介護者、または医療従事者ではない他のユーザーに患者固有の診断または治療の推奨事項を提供する GenAI 対応製品は、FDA の規制監督の対象となるデバイスです。

FDA のデバイスの監督はリスクベースです。つまり、安全性と有効性の合理的な保証を実証するために必要な規制管理のレベルは、通常、デバイスのリスクのレベルと一致しています。各デバイスは、クラス I、クラス II、またはクラス III の 3 つの規制クラスのいずれかに割り当てられます。各クラスには、デバイスの安全性と有効性の合理的な保証を提供するために必要な規制管理のレベルが高くなっています。リスクを軽減できる範囲は、特定の製品の規制管理のレベルに影響を与える可能性があります。

一般に、GenAI を組み込んだものを含む医療機器は、機器の分類とリスクの程度に基づいて、以下のいずれかの規制経路の下で市販前審査の対象となる可能性があります¹⁵。

- 市販前承認 (PMA) : 機器が高リスクの場合
- De Novo 分類要求: 機器が低リスクから中リスクであり、合法的に市販されている述語機器がない場合
- 市販前通知 [510 (k)] : 機器が低リスクから中リスクであり、合法的に市販されている述語機器がある場合

市販前経路の種類 (PMA、De Novo、または 510 (k)) にかかわらず、安全性と有効性の原則は、FDA によるすべての医療機器の審査の基礎となります。最もリスクの低い機器は、(上記の規制経路を通じて審査される他の機器と同様に) 一般的な管理の対象となりますが、一般的には市販前審査の対象外となります。

前述のように、FDA は GenAI 対応機器を含む機器に対してリスクベースのアプローチを適用し続けることが重要です。しかし、GenAI 対応機器に関連するいくつかの特性は、他のソフトウェア対応機器と比較して固有のリスクをもたらす可能性があり、GenAI 対応機器の規制に対する FDA の理解とアプローチに影響を与える可能性があります。

GenAI 対応機器と TPLC 規制アプローチの課題

医療機器への GenAI 導入に伴う特有のリスクを考慮する際、FDA は 2 つの規制上の課題に直面しています。1 つ目は、リスクベースのアプローチを分類に適用し、GenAI 対応機器の規制要件を決定することに関連する課題です。2 つ目は、FDA の規制監督を必要とする GenAI 対応機器について、FDA が TPLC 全体で GenAI 対応機器の安全性と有効性を評価するための有効な科学的証拠 (「エビデンス」)¹⁶ の種類を決定することに関連する課題です。

¹⁵ [医療機器の安全性および 510 \(k\) クリアランスプロセス](#)に関する FDA のウェブサイトを参照。

GenAI 対応デバイスにリスクベースのアプローチを適用する際の考慮事項

前述のように、GenAI の独自の特性は、GenAI モデルが"新しい情報"を生成できることであり、一部の実装では、より"創造的な"応答を生成するために、この能力を意図的に活用することさえあります。ただし、これにより、GenAI モデルが特定または説明が困難なハルシネーション (hallucinations) を起こしやすくなる可能性もあります。FDA とデバイスユーザーにとって、GenAI 対応デバイスによって生成されるハルシネーション (hallucinations) は、デバイスの動作に不確実性をもたらす可能性があります。これは、デバイスの使用目的の特定の範囲を理解することが困難になる可能性があります。たとえば、患者と医療専門家との対話を要約することを意図した GenAI 対応製品の場合、その製品のハルシネーション (hallucinations) の可能性によって、医療専門家と患者との議論を要約することと、対話中に提起されなかった新しい診断を提供することとの間に違いが生じる可能性があります。

さらに、このエグゼクティブサマリーで既に説明したように、多くの場合、基盤モデルやその他の GenAI ツールは、特定のタスクやユースケースに合わせて調整されることなく、幅広いタスクを実行するように設計されています。そのため、これらの基盤モデルやその他の GenAI ツールは、一般に、特定の医療機器の用途を想定していません。ただし、デバイスメーカーが、デバイス定義を満たす特定の用途を持つ製品の一部として基盤モデルやその他の GenAI ツールを使用している場合、基盤モデルを活用している製品は、FDA のデバイス規制の監視の対象となる可能性があります。このような状況では、基盤モデルやその他の GenAI ツールを GenAI 対応デバイスに組み込むデバイスメーカーが、基盤モデルに関する詳細情報 (モデルの属性、アーキテクチャ、トレーニング方法論、データセットなど) を取得することは困難な場合があります。これらのモデルは、基盤モデル開発者がモデルを展開し、ユーザーにアクセスを提供する方法に応じて"ロック"または"ロック解除"¹⁷される可能性があるためです。

さらに、基盤モデルには一般に、適切に制御されていない可能性のある大量のデータが提供されます。そのため、基盤モデルはバイアスの影響を受けやすく、個々の製品開発者が結果として得られる GenAI 対応製品を特定または軽減することが特に困難な場合があります。医療機器に組み込まれた既製品 (OTS) ソフトウェアの可視性が限られていることは、すでに製造業者にとって潜在的なトレードオフとなっています。OTS ソフトウェアを使用すると、デバイスの製造業者はデバイス固有の機能を実行するために必要なものに集中できますが、OTS ソフトウェアは医療機器の特定の用途に常に適しているとは限りません。

¹⁶ このエグゼクティブ・サマリーでは、「有効な科学的証拠」と「証拠」という用語は同じ意味で使用されます。

¹⁷ 「アンロックモデル」、「継続的機械学習」、およびその他の用語は同じ意味で使用できます。

ガイダンスに記載されているように、[医療機器における市販ソフトウェアの使用](#): "OTS ソフトウェアを使用する医療機器メーカーは、一般にソフトウェアのライフサイクル管理を放棄しますが、医療機器の継続的な安全で効果的なパフォーマンスに対して責任を負います"。GenAI 対応デバイスのコンテキストでは、組み込まれた基盤モデルに対するソフトウェアライフサイクル管理の類似の欠如によって、特定の課題が発生する可能性があります。

たとえば、デバイスメーカーが基盤モデルを非常に限定的に制御できるシナリオでは、基盤モデルの全体的なデータ構成の変更が、GenAI 対応デバイスの使用目的に影響を与えるか、パフォーマンスバイアスにつながる可能性があります。これにより、デバイスの動作がさらに不確実になる可能性があります。重要なことに、FDA はロック解除された AI 対応デバイスをまだ承認していません。したがって、サードパーティの基盤モデルを適用すると、モデルの開発に対するメーカー自身の可視性が限られているために、デバイスの動作が不確実になる可能性があるだけでなく、サードパーティの基盤モデルが FDA の審査中にさらに新しい重要な問題を提起する可能性があります。

要約すると、すべてのデバイスと同様に、FDA が規制監督を適用し、デバイスの安全性と有効性を確保する方法を決定できるように、GenAI 対応デバイスの使用目的と完全な機能を理解することが重要です。さらに、これらのデバイスが FDA の審査に新たな課題を提示する可能性があるとしても、規制監督に対する FDA のリスクベースのアプローチをサポートするために、そのような意図された使用と機能の範囲を適切に実証する必要があります。

FDA のデジタルヘルスポリシーへの影響

FDA のデジタルヘルスおよびソフトウェアに関するさまざまなガイダンスは、GenAI 対応製品を含む製品が FDA のデバイス規制監督の対象であるかどうかをメーカーが判断するのに役立ちます。ただし、GenAI に関する前述の特性と関連する不確実性を考慮すると、製品が FDA のデジタルヘルスポリシーの範囲内にあるかどうかを理解することは困難な場合があります¹⁸。

たとえば、GenAI 非対応製品であるモバイルアプリケーションを考えてみましょう。このアプリケーションは、同時に服用したいハーブと薬をユーザーに入力するように促し、相互作用が文献で確認されているかどうかに関する情報と、報告された相互作用の種類の概要を提供することを目的としています。FDA のガイダンス「[デバイスソフトウェア機能およびモバイル医療アプリケーションのポリシー](#)」で説明されているように、このようなソフトウェアは、一般に FDA が執行裁量権を行使することを意図したソフトウェアです。このガイダンスでは、この例と、ガイダンスのこのセクションの他の例は、医療機器の定義を満たす可能性があるが、公衆に対するリスクが低い場合、FDA が執行裁量権を行使することを意図したソフトウェア機能の例であることも言及しています。代わりに、同じ製品が GenAI 対応だった場合を考えてみましょう。GenAI 対応製品は同じ使用を意図しているかもしれませんが、十分に管理されていない場合、1) 新たに明らかになった医療機器の機能との併用(例えば、照会された薬草/薬物の組み合わせに関して確認された相互作用のために、代替の薬草/薬物の組み合わせを用いた治療に患者を誘導すること)、または 2) 照会された薬草/薬物の組み合わせの相互作用に関する誤った又は虚偽の内容。後者の場合、そのような GenAI 対応製品はまだデバイスの定義を満たしていないかもしれないが、意図された用途が既知の情報へのアクセスを患者に提供することである製品において、新しいコンテンツ及び情報を生成する GenAI の実装が潜在的な不確実性及び/又はリスクをもたらすかどうかを考慮することが重要です。

¹⁸ [デジタルヘルスコンテンツに関する](#) FDA のガイダンスを参照してください。

この概念をより一般的に適用するには、GenAI がなければ低リスクとみなされ、FDA の規制監督の対象とならなかった製品に、GenAI の実装が潜在的な不確実性やリスクをもたらすかどうかを検討することが重要です。多くの場合、製品に GenAI が搭載されているかどうかによって、FDA の監督の考慮事項が変わることはありません。しかし、特定の場、例えば、GenAI モデルが十分に制御されていない場合、GenAI 搭載製品は、FDA の機器規制監督の対象となるような方法で機能するために、明示された使用目的の外で動作する可能性があります。場合によっては、製造業者や開発者が、誤った内容や虚偽の内容を提供したり、誤った情報を拡散したりする可能性がある場合には、非 GenAI の実装と比較して、製品の GenAI 実装が公衆衛生に有益ではない可能性があることを考慮することが役立つことがあります。したがって、製造業者や開発者は、特定の使用目的に特定の GenAI モデルを実装することの潜在的な不確実性やリスクを考慮すること、言い換えれば、GenAI が特定の使用目的に最適な技術である場合とそうでない場合を考慮することが役立ちます。ヘルスケアやその他のリスクの高い分野に適用される他の技術と同様に、各タスクに「適切なツール」を適用することを追求し続けることが特に重要です。製造業者が GenAI 対応デバイスを開発することに利点があると判断した場合、製造業者は、デバイスの安全性と有効性の合理的な保証を FDA が評価するために、そのような実装のために生成する必要がある可能性のある証拠を検討することが重要です。GenAI 対応デバイスに関するそのような証拠の生成は困難な場合があります。例えば、以下の装置の正確な説明または特性評価を行うことが困難な場合があります。

基本的な基盤モデルについてほとんど知られていない場合。また、市販前および市販後のパフォーマンスデータを生成することも困難な場合があります。本エグゼクティブサマリーの後半では、FDA が GenAI 対応装置の安全性と有効性を評価するためのエビデンスの生成に関連して、FDA が予測するいくつかの課題について説明します。

デバイス分類への影響

FDA のデバイスの監視はリスクベースであり、デバイスの定義を満たす多くの GenAI 対応製品が FDA のデバイス規制監視の焦点となります。しかし、上述したように、GenAI 対応製品の分類と規制要件の決定に FDA の現在のリスクベースのアプローチを適用することには課題があるかもしれません。このような製品は、制約が困難であったり、意図的に広い用途を有していたり、不確実な性能を示す可能性があります。

前述したように、GenAI は一般的に、情報を想起したり予測したりするために特別に訓練されるのではなく、学習したデータに似た新しいデータを出力することを意図しています。さらに、可変出力を生成するように設計することもできます。GenAI のこれらの側面はどちらもハルシネーション (hallucinations) につながる可能性があり、GenAI ベースの製品が明確に制約された使用目的を持ち、正確で一貫性のある信頼性の高い出力を提供することを実証するのは困難です。FDA は、規制上の監視をどのように適用するか、具体的には、デバイスの安全性と有効性を確保するためにデバイスの分類にリスクベースのアプローチをどのように適用するかを決定するために、デバイスの使用目的と技術的特性を理解する必要があります。GenAI 対応デバイスは、広く AI 対応デバイスと同様に、FDA の現在のデバイス分類スキームにも挑戦する可能性があります。これは、GenAI の技術的特性が、特定の GenAI 対応製品に対して新しいまたは異なるリスクをもたらす可能性があるためです。これにより、デバイスの分類または市販前経路に影響を与える可能性のある安全性と有効性、およびそのようなデバイスの安全性と有効性を確保するために必要となる可能性のある規制管理の種類に関する新しい問題が提起されます。



デバイスの分類または規制経路への影響に関しては、FDA の最大の市販前プログラムである 510 (k) プログラムで最も明らかになっています。FD&C 法のセクション 513 (i) に記載されているように、510 (k) 要件の対象となるデバイスについては、実質的な同等性の決定には、他の要件の中でも、述語デバイスと対象デバイスの意図された用途と技術的特性の比較が含まれます。GenAI の技術的特性と導入された不確実性に基づいて、非 GenAI 対応デバイス (述語デバイス) が GenAI 対応デバイス (対象デバイス) と同じ意図された用途を持つ状況を想定すると、GenAI 対応デバイスは述語デバイスとは異なる技術的特性を持つことが判明する可能性があり、これらの異なる技術的特性によって安全性と有効性に関する異なる問題が提起される可能性があります。

GenAI 対応デバイスの安全性と有効性を確保するために役立つ規制またはリスク管理には、多くの種類があります。製造業者は、可能な限り、基になる基盤モデルが制御できない状況であっても、GenAI 対応デバイスの基になる GenAI モデルを制御できるかどうか、およびどのように制御できるかを検討することが役立ちます。さらに、GenAI モデルの適切なガバナンス、デバイスの安全性に関する適切なフィードバックメカニズムの有用性、および実際のパフォーマンス評価などの他のリスク管理は、GenAI 対応デバイスが実際の使用において安全かつ有効に機能していることを保証するのに役立ちます。すべての医療機器は、免除されない限り、一般的な管理の対象となります。これには、医療機器の報告 (21 CFR Part 803)、訂正および削除の報告 (21 CFR Part 806)、施設登録およびデバイスリスト (21 CFR Part 807)、および品質システム規制 (21 CFR Part 820) が含まれます。¹⁹ 一般的な管理以外にも、特定のクラス II デバイスについて、FDA は、デバイスの安全性と有効性を合理的に保証するために必要な場合に、GenAI 対応デバイスに固有の特別な管理を要求することがあります。このような特別な管理には、デバイスの安全性とパフォーマンスの市販後モニタリングの要件、または GenAI 対応デバイスが意図したとおりに機能していない場合の通知要件が含まれる場合があります。

FDA の現在のリスクベースのアプローチを GenAI 対応製品に適用することには課題があるかもしれませんが、GenAI モデルの設計と制御、その基盤となるモデル、および GenAI 対応製品の意図された用途と技術的特性を適切に理解することで、FDA は規制上の監視を適用し、デバイスの安全性と有効性を確保する方法を決定することができます。さらに、FDA が確立する可能性のある明示的な規制管理の範囲を超えて、これらのデバイスが実際の使用において安全性と有効性を維持するのに役立つ可能性のあるリスク管理戦略を検討することは、メーカーやこれらの製品を使用する事業者にとって引き続き重要です。

GenAI 対応デバイスは、広く AI 対応デバイスと同様に、FDA の現在のデバイス分類スキームにも挑戦する可能性があります。これは、GenAI の技術的特性が、特定の GenAI 対応製品に対して新しいまたは異なるリスクをもたらす可能性があるためです。これにより、デバイスの分類または市販前経路に影響を与える可能性のある安全性と有効性、およびそのようなデバイスの安全性と有効性を確保するために必要となる可能性のある規制管理の種類に関する新しい問題が提起されます。デバイスの分類または規制経路への影響に関しては、FDA の最大の市販前プログラムである 510 (k) プログラムで最も明らかになっています。FD&C 法のセクション 513 (i) に記載されているように、510 (k) 要件の対象となるデバイスについては、実質的な同等性の決定には、他の要件の中でも、述語デバイスと対象デバイスの意図された用途と技術的特性の比較が含まれます。GenAI の技術的特性と導入された不確実性に基づいて、非 GenAI 対応デバイス (述語デバイス) が GenAI 対応デバイス (対象デバイス) と同じ意図された用途を持つ状況を想定すると、GenAI 対応デバイスは述語デバイスとは異なる技術的特性を持つことが判明する可能性があり、これらの異なる技術的特性によって安全性と有効性に関する異なる問題が提起される可能性があります。

GenAI 対応デバイスの安全性と有効性を確保するために役立つ規制またはリスク管理には、多くの種類があります。製造業者は、可能な限り、基になる基盤モデルが制御できない状況であっても、GenAI 対応デバイスの基になる GenAI モデルを制御できるかどうか、およびどのように制御できるかを検討することが役立ちます。さらに、GenAI モデルの適切なガバナンス、デバイスの安全性に関する適切なフィードバックメカニズムの有用性、および実際のパフォーマンス評価などの他のリスク管理は、GenAI 対応デバイスが実際の使用において安全かつ有効に機能していることを保証するのに役立ちます。すべての医療機器は、免除されない限り、一般的な管理の対象となります。これには、医療機器の報告 (21 CFR Part 803)、訂正および削除の報告 (21 CFR Part 806)、施設登録およびデバイスリスト (21 CFR Part 807)、および品質システム規制 (21 CFR Part 820) が含まれます。¹⁹

¹⁹ FDA の [規制管理](#) に関するウェブサイトを参照してください。

一般的な管理以外にも、特定のクラス II デバイスについて、FDA は、デバイスの安全性と有効性を合理的に保証するために必要な場合に、GenAI 対応デバイスに固有の特別な管理を要求することがあります。このような特別な管理には、デバイスの安全性とパフォーマンスの市販後モニタリングの要件、または GenAI 対応デバイスが意図したとおり機能していない場合の通知要件が含まれる場合があります。FDA の現在のリスクベースのアプローチを GenAI 対応製品に適用することには課題があるかもしれませんが、GenAI モデルの設計と制御、その基盤となるモデル、および GenAI 対応製品の意図された用途と技術的特性を適切に理解することで、FDA は規制上の監視を適用し、デバイスの安全性と有効性を確保する方法を決定することができます。さらに、FDA が確立する可能性のある明示的な規制管理の範囲を超えて、これらのデバイスが実際の使用において安全性と有効性を維持するのに役立つ可能性のあるリスク管理戦略を検討することは、メーカーやこれらの製品を使用する事業者にとって引き続き重要です。

デバイスとして FDA の規制監督の焦点となっている GenAI 対応製品については、TPLC を介した GenAI 対応デバイスに関する新しいタイプの有効な科学的証拠の生成とレビューに関する第 2 の課題もあります。

TPLC を介した GenAI 対応デバイスの有効な科学的証拠に関する考慮事項

FDA は、FDA の規制監督を必要とするデバイスの安全性と有効性の合理的な保証を決定する一環として、多くの種類の有効な科学的証拠をレビューします。一般に、AI 対応デバイスに必要な証拠は、GenAI 対応デバイスにも適用される可能性があります。デバイスの安全性と有効性を確保するために、追加または異なる証拠で補足する必要がある場合があります。たとえば、コンピュータ支援トリアージ、検出、および診断デバイスの評価に使用されるような現在の評価アプローチは、追加の裏付け証拠がある場合でも、GenAI 対応デバイスに適用できる可能性があります。ただし、特定の GenAI 実装、特に、他の AI 対応デバイスで使用される一般的な構造化形式とは異なるオープンエンドの入力および出力形式を使用する GenAI 対応デバイスでは、これらの実装に新しい評価考慮事項がある可能性があるため、必要な証拠を決定することは困難な場合があります。

GenAI 対応デバイスに提供される特定の情報は、その使用目的と設計によって管理されますが、一部の GenAI 対応デバイスの広範なアプリケーションは、デバイスの明示された使用目的を制限することが困難になる可能性があります。レビュープロセスに課題が生じます。そのため、特定のデバイスに応じて、基盤となるデバイス設計とパフォーマンステスト要件に関して、GenAI 対応デバイスのマーケティング申請で必要とされる詳細のレベルが高くなる可能性があります。デバイスでの GenAI モデルの実装に関する詳細に加えて、OTS ソフトウェアを医療デバイスに組み込むために採用されている現在のアプローチと同様に、使用されている基盤モデルに関する情報も必要になる場合があります。この情報には、基盤モデルの属性、アーキテクチャ、およびトレーニング方法とデータセットが含まれる可能性があります。これは、メーカーが取得するのに簡単ではない可能性があります。基盤モデルのトレーニング方法とデータセットに関する情報は、特に重要になる可能性があります。これは、前述のように、基盤モデルには一般に、適切に制御されていない可能性のある大量のデータが提供されるためです。さらに、基盤モデルには数十億のモデルパラメーターがあります。パラメーター構造が大きく複雑であるため、入力データを少し変更すると、出力が異なる可能性があります。たとえば、要求の言い回しを少し変更すると、GenAI モデルが要求されたフィードバックに対して異なる応答または予測を生成する可能性があります。GenAI モデルにはこのようなオープンエンドの入力があるため、展開中にモデルが遭遇する可能性のあるすべての考えられる入力を評価するのは無理があります。また、サードパーティの基盤モデルがパラメーターを公開していない場合は、特に実行不可能です。このような透明性の欠如と、突発的または予期しない動作の可能性は、市販前の評価が特に困難になる可能性があります。これらの関連するリスクと市販前の課題を軽減し、TPLC 全体で GenAI 対応デバイスの継続的な安全性と有効性を確保するには、市販前のレビューを効果的に調整された市販後の監視戦略で補完することが特に重要になる場合があります。

現在の市販前のエビデンスの必要性

FDAによるデバイスの審査の一環として、FDAはデバイスとその設計に関する一般的な理解を得る必要があります。²⁰ AI対応デバイス全般について、これには設計仕様、データ管理とモデル開発情報、およびモデルの特性に関する理解が含まれます。特に、GenAI対応デバイスでは、モデルに関連する情報の種類は同じですが、期待される詳細は異なる可能性があります。この情報はGenAI対応デバイスに固有のものではないかもしれませんが、この情報は、特に適切に制限された意図された使用および/またはデバイスのパフォーマンスにおける許容可能な変動性の実証に関連する、GenAI対応デバイスの考慮事項の一部に対処するのに役立つ可能性があります。

GenAI対応デバイスの場合、設計の詳細には、特定の設計仕様(例:注意メカニズム、モデルのマージなど)、モデルパラメーター(例えば、温度、top-K、top-pなど)、組み込みのプロンプト戦略(例えば、コンテキスト内学習、ゼロショットプロンプトなど)、エンドユーザーが使用できるプロンプト機能などの情報が含まれる場合があります。エンドユーザーが使用できるプロンプト機能。GenAI対応デバイスのトレーニングの場合、これには、初期の基盤モデルのデータ管理とモデル開発に関する情報と、特定のGenAI対応デバイスの基になるGenAIモデルのデータ管理と微調整に関する特定の詳細が含まれる場合があります。GenAI対応デバイスの特定の適応につながる微調整の手順は、デバイスの使用目的を制御する可能性があるため、理解しておくことが重要です。

最後に、デバイスの出力特性はすべてのAIベースのデバイスにとって重要ですが、GenAI対応デバイスでは特に重要になる可能性があります。これは、説明したように、設計によって変化する新しい情報を出力することができます。GenAI対応デバイスは、従来のデバイスや一部のAI対応デバイスと比較して複雑さのレベルが高くなる可能性があるため、エンドユーザーに対して透明性を確保することが重要です。²¹ これには、デバイスの設計、デバイスのテスト方法、自律性のレベル、ユーザーがデバイスと対話する方法(例えば、迅速なエンジニアリングによって)など、デバイスの出力に影響を与える可能性のある、GenAI対応デバイスのエンドユーザーに対するさまざまな透明な情報が含まれる場合があります。また、GenAI対応デバイスの自律性のレベルを理解し、ヒューマンインザループを組み込んでエンドユーザーに制御レベルを提供するかどうか、およびその方法を理解することも重要です。

前述のように、これらのエビデンスのニーズの多くはGenAI対応デバイスに特有のものではないが、堅牢な設計仕様と出力特性は、FDAとユーザーが、意図された用途とデバイスの性能の境界に影響を与える可能性のあるGenAI対応デバイス特有の考慮事項をよりよく理解するのに役立つかもしれません。

²⁰ 例えば、21 CFR 807.87、21 CFR 860.220、または21 CFR 814.20を参照してください。

²¹ 「[Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles](#)」も参照してください。

市販前のエビデンスのニーズに応える新たな方法論

一部の GenAI 対応デバイスでは、現在のパフォーマンス評価の方法論が依然として適用される可能性があります。さらに、迅速なエンジニアリング、入力および/または出力の種類に関する品質管理、トピック固有の基盤モデルの使用などの方法論と戦略の使用は、特定の GenAI 対応デバイスを評価するために現在の定量的なパフォーマンス評価方法論を使用するのに役立つ可能性があります。

しかし、安全性と有効性を維持するために、GenAI 対応デバイスのパフォーマンスを広く評価するための新たな方法論も開発する必要があります。現在の定量的評価アプローチは、デバイスのパフォーマンスの完全または完全な評価を提供しない可能性があり、追加のパフォーマンスメトリックで補完する必要があるかもしれません。例えば、GenAI 対応デバイスの定性的パフォーマンス評価のための新たなアプローチは、自律性、透明性、および説明可能性のレベルに関連する GenAI 対応デバイスの基礎となる GenAI モデルを正確に特徴付けるのに役立つかもしれません。さらに、このような新たなアプローチは、医療分野全体で幅広いアプリケーションを持つ GenAI 対応デバイスのパフォーマンスを評価するのに役立つかもしれません。

しかし、必要なパフォーマンス評価方法は、GenAI 対応デバイスの特定の意図された用途と設計によって支配され、その一部は、特定の意図された用途のための新しいパフォーマンスメトリックの策定を必要とするかもしれません。すべてのデバイスと同様に、市販前および市販後の証拠を含む証拠の全体は、TPLC 全体でこれらのデバイスの安全性と有効性の合理的な保証を支持することができます。

市販後の新しい証拠の必要性

新しい方法論を使用して生成される可能性のある市場前の証拠に加えて、市場後の証拠も、GenAI 対応デバイス、および一般的な AI 対応デバイスの安全性と有効性を確保する上で役割を果たす可能性があります。大まかに言えば、AI 対応デバイスは、ローカライズされたライブデータ、ユーザーの操作、および条件の変化に基づいて継続的な調整を受ける可能性があります。前述のように、GenAI 対応デバイスでは、その基盤となる基盤モデルが条件の変化の影響をさらに受けやすい可能性があります。広く AI 対応デバイスを監視および継続的に調整するためのフレームワークと比較して、このような GenAI 対応デバイスに同様の監視と制御を提供することは特に困難である可能性があります。たとえば、一部の GenAI 対応デバイスでは、デバイスメーカーは、デバイスに影響を与える可能性のある基盤モデルのデータを管理および可視化できない場合があります。デバイスの安全性と有効性を維持するためには、GenAI 対応デバイスを含む AI 対応デバイスを市販後に監視および評価することが重要になります。さらに、効果的な評価と監視に関連する課題がどのように進化するかを考慮することが重要です。たとえば、メーカーが、ロックされていないモデルの使用、サードパーティの基盤モデルを利用したデバイス開発、複雑化する多層設計、および AI 対応デバイスにおける部位または患者固有のモデルの使用を提案し始めるとします。

大まかに言えば、製造業者は、AI 対応および GenAI 対応デバイスの安全性とパフォーマンスをプロアクティブな方法で監視するメカニズムとして、堅牢な市販後モニタリング戦略を検討することができます。これは、FDA がデバイスの安全性と有効性を経時的に保証するのにも役立ちます。例えば、FDA のガイダンス「[Consideration of Uncertainty in Making Benefit-Risk Determinations in Making Medical Device Pre-market Approvals, de Novo Classifications, and Humanitarian Device Exemptions](#)」で説明されているように、「FDA はまた、ベネフィット・リスク情報の不確実性に対処するために、リスク軽減と市販後データの収集の適切性を検討しています」。これは、市販前の不確実性を軽減するのに役立ちます。さらに、ガイダンスでは、「医療機器がどのように使用され、どのように機能し、対応する患者のアウトカムについての理解を継続的に改善するためには、学習するヘルスケアシステムの一部として、市販前および市販後の環境全体で継続的で堅牢なエビデンスを生成することが重要です」と述べられています。

したがって、すべてのデバイスと同様に、製造業者は GenAI 対応デバイスの市販後モニタリング戦略を実装することが重要になります。特に、製造業者は、展開後にデバイスの正確性、関連性、信頼性が維持されるように、GenAI 対応デバイスの特定の用途を効果的に評価し、監視する方法を検討することが重要になります。たとえば、各 GenAI 対応デバイスのサイト固有の実装など、GenAI 対応デバイスのさまざまな実装がある場合、市販後モニタリング戦略がより複雑になる可能性があることに注意することが重要です。最後に、製造業者は、GenAI モデルだけでなく、必要に応じて基盤モデルの市販後モニタリング戦略を検討することが重要になる場合があります。基盤モデルと GenAI モデルの両方の条件が変化すると、最終的に GenAI 対応デバイスの安全性とパフォーマンスに影響を与える可能性があるためです。特定の AI 対応デバイスまたは GenAI 対応デバイスについては、デバイスの使用目的とリスクに応じて、FDA が規制管理として市販後モニタリングを要求する場合があります。AI 対応デバイスおよび GenAI 対応デバイスの継続的な安全性と有効性を確保するためには、市場前のエビデンスを、堅牢なデバイスパフォーマンスモニタリングアプローチを通じて収集された市場後のエビデンスで補完することが重要になる場合があります。

要約

FDA の CDRH は、患者と医療提供者が安全で効果的かつ高品質な医療機器をタイムリーかつ継続的に利用できるようにすることに取り組んでいます。この使命の一環として、CDRH は規制科学を推進し、業界に予測可能で一貫性があり、透明性があり、効率的な規制経路を提供し、米国で販売されている医療機器に対する消費者の信頼を保證することで、医療機器のイノベーションを促進します。

GenAI の新しい機能は、患者と公衆衛生に独自の利益をもたらす可能性があります。GenAI の使用と採用には、デバイスの規制に対する FDA のアプローチに挑戦する特定のリスクと複雑さも伴います。特に、FDA は、GenAI 対応デバイスの分類と規制要件の決定にリスクベースのアプローチを適用することに関連する課題に直面しており、FDA の規制監督を必要とする GenAI 対応デバイスについては、FDA が TPLC 全体で GenAI 対応デバイスの安全性と有効性を評価するための有効な科学的証拠の種類を決定することに関連する課題に直面しています。FDA は、AI 対応デバイスを含む医療機器の監督に対する TPLC アプローチを長年推進しており、これらのデバイスの規制アプローチを前進させることにコミットしています。ヘルスケア部門全体で GenAI ツールへの関心が急速に拡大する中、FDA は、FDA の規制管轄下にある可能性のある GenAI 対応製品の規制アプローチには未解決の問題が残っており、FDA が専門家と協力してタイムリーにこれらの問題に対処することが公衆衛生上重要であると指摘しています。

パネル質問

- 1. 市販前性能評価:** 第一世代 AI 対応デバイスの安全性と有効性を評価するために FDA が入手すべき、第一世代 AI 対応デバイスが活用する基礎モデルが時間の経過とともに変化し、これらの事前トレーニング済みの生成モデルに利用されるトレーニングデータに関する情報が限られている可能性があることを考慮して、生成 AI に関連する具体的な情報について説明してください。
 - a. デバイスが生成 AI によって有効化されている場合、デバイスの説明または特性評価の一部としてどのような情報を市販前申請に含める必要がありますか。例えば、人間がループに入ることの意図していない場合、またはデバイスが新しい推奨事項を生成するのではなく、情報を想起することのみを意図している場合。非生成 AI と比較して、生成 AI で有効化されたデバイスの安全性と有効性を評価するために特に価値のある情報は何ですか。
 - b. デバイスが安全で有効であるかどうかを理解するために、FDA は製品ライフサイクル全体のパフォーマンス評価とトレーニングデータの特性に関して、生成 AI 対応デバイスに固有のどのような証拠を市販前評価で考慮する必要がありますか。
 - c. 非生成 AI と比較して、生成 AI によってユーザビリティに関連するどのような新しい固有のリスクが導入される可能性がありますか。医療専門家、患者、および介護者に関連する特定の情報がある場合は、透明性の向上やこれらのリスクの制御に役立つように伝達する必要がありますか。
 - d. これらのテクノロジーの複雑さを考慮すると、これらのテクノロジーに特に適している/最も有益なパフォーマンスメトリックはありますか。入力、出力、またはその両方がマルチモーダルである可能性があるテキスト/画像モデルなど、マルチモーダルにはどのような種類のパフォーマンスメトリックが必要ですか。パフォーマンスメトリックは、通常、デバイスの使用目的によって異なります。議論をサポートする既知のメトリックの例としては、生成テキスト(当惑、参考文献との定量的比較)、生成画像(フレイッシュ開始距離 (FID)、構造類似性指数メジャー (SSIM))、生成オーディオ(対数スペクトル距離、音声品質の知覚評価)などのモダリティ固有のものもあります。また、生成 AI 対応デバイスによって行われるエラーの頻度や種類など、機能ベースのものもあります。
- 2. リスク管理:** 医療デバイスの生成 AI によって、新しい用途や既存の用途での新しいアプリケーションなど、どのような新しい機会が実現されましたか。また、これらの機会を可能にする生成 AI テクノロジーに関連するリスクを軽減するために、どのような新しいコントロールが必要になりますか。たとえば、ガバナンス、トレーニング、フィードバックメカニズム、および実際のパフォーマンス評価に関連するコントロールです。
- 3. 市販後のパフォーマンス監視:** これらのデバイスでは、特に非決定論的であるため、市販後のパフォーマンスモニタリングと評価が重要になる可能性があります。さらに、多くの生成型 AI 対応デバイスでは、ローカライズされたライブデータ、ユーザーの操作、展開後の条件の変化に基づいて継続的な調整が行われます。これらのデバイスの安全性と有効性を維持するために重要となる市販後のモニタリングと評価の側面について説明してください。

- a. 特に新しいデータに適応する場合に、十分な精度、関連性、および信頼性を維持するために、生成型 AI 対応デバイスの市販後のパフォーマンスを効果的に評価および監視するには、どのような具体的なモニタリング機能を考慮する必要がありますか？
- b. 複数のサイトに実装された生成型 AI 対応デバイスのパフォーマンスと精度を監視および管理し、一貫性を確保し、承認されたデバイスと比較して潜在的な地域バイアスやデータの変動に対処するために、どのような具体的な戦略とツールを実装できますか？
- c. 多層アプリケーション設計を使用する生成型 AI 対応デバイスの市場後のパフォーマンスを効果的に監視および評価するには、どのような方法とメトリックを利用できますか？ つまり、デバイスは、それ自体が医療機器ではない外部の消費者向け AI サービスにクエリを実行しますか？

付録

このエグゼクティブサマリーで簡単に説明されており、GenAI 対応デバイスを含む AI 対応デバイスの TPLC ニーズに関する議論を促進するために提供されている AI ライフサイクルは、図 2 に示すように、各ライフサイクルフェーズに関連する技術的な考慮事項の一覧にさらに拡張できます。このエグゼクティブサマリーでは、これらの"ライフサイクルごとのフェーズ"の一連の技術的な考慮事項は、AI 対応ソフトウェアの広範なライフサイクルフェーズを代表するものであることを意図していますが、GenAI 対応ソフトウェアについても同様です。AI ライフサイクルの考慮事項は、AI および GenAI 対応デバイスの製造元(規制当局、顧客、臨床医などと同様に。)に、デバイスのライフサイクル全体を通じて、AI 開発のベストプラクティス、ツール、適用可能なメトリック、および標準への準拠を開発および評価するための実用的なフレームワークを提供できます。このようなフレームワークは、ライフサイクルフェーズに適した所有者と関係者がデバイスの品質を確保するために適用して文書化できる、コンプライアンスと品質に関する考慮事項への構造化されたアプローチを提供できます。たとえば、GenAI 対応デバイスメーカーは、そのデバイスに組み込まれているサードパーティの基本モデルの設計と実装に関する可視性が限られている場合があります(エグゼクティブサマリーで説明しています)。AI ライフサイクルは、開発の最も早い段階から、GenAI に固有のフェーズと考慮事項(データ管理、モデルのトレーニングとテスト、モデルの検証と検証など。)、および GenAI が独自の課題を提示する可能性があるセキュリティ、リスク管理、運用化の考慮事項などの一般的な考慮事項に注意を向けることで、GenAI 対応デバイスの開発、提供、および運用化に役立つ場合があります。技術的な考慮事項は、網羅的なものではなく、規制対象デバイスの要件や推奨事項として意図されたものでもありません。

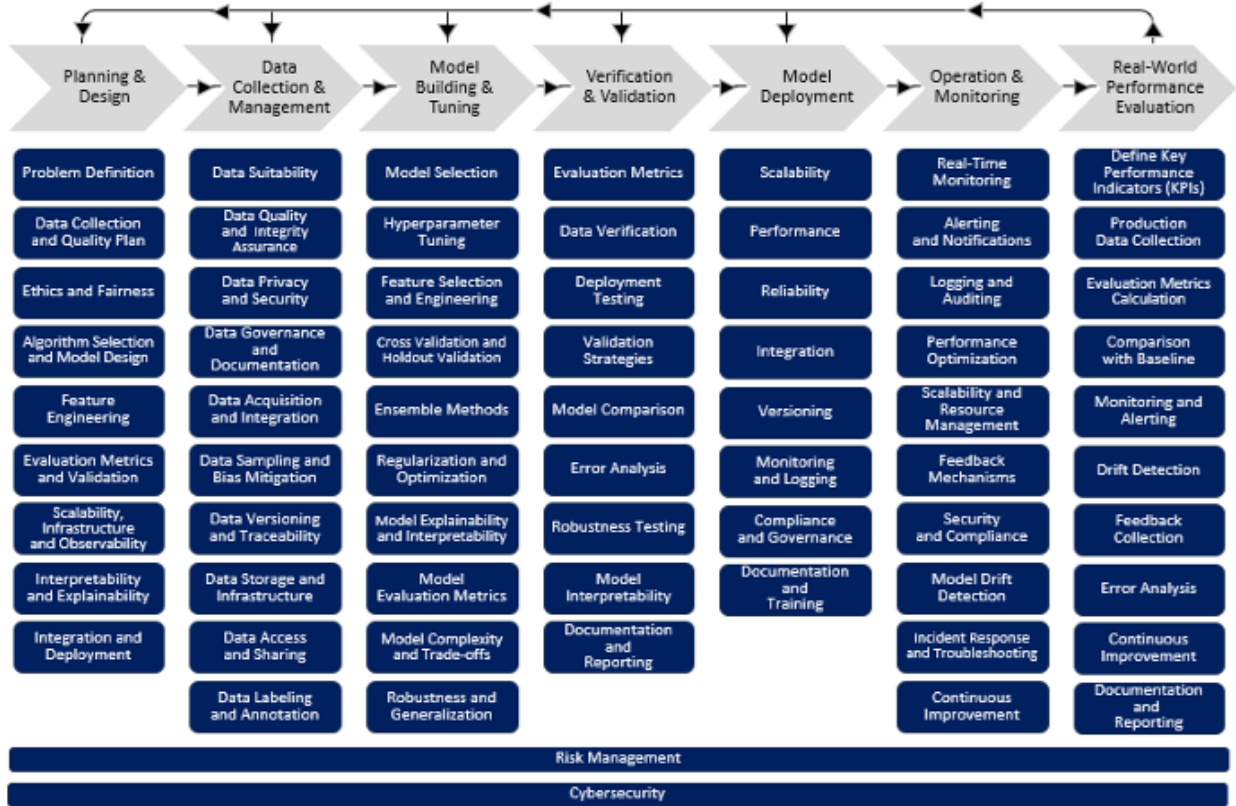


Figure 2: AI Lifecycle Considerations



用語集

用語	定義
バイアス Bias	<p>他と比較した特定の対象、人、またはグループの扱いにおける体系的な違い。</p> <p>注 1:扱いとは、知覚、観察、表現、予測、または決定を含むあらゆる種類の行為です。(ISO/IEC TR 24027:2021)</p> <p>Source: International Medical Device Regulators Forum. (2022). Machine Learning-enabled Medical Devices: Key Terms and Definitions</p>
人工知能 (AI)	<p>人間が定義した特定の目標セットに対して、現実または仮想環境に影響を与える予測、推奨、または決定を行うことができる機械ベースのシステム。人工知能システムは、機械および人間ベースの入力を使用して、現実および仮想環境を認識します。自動化された方法での分析を通じて、そのような認識をモデルに抽象化します。モデル推論を使用して、情報またはアクションのオプションを策定します。</p> <p>Source: Executive Order 14110, (October 30, 2023), Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence</p>
人工知能パフォーマンスモニタリング (AI パフォーマンスモニタリング)	<p>展開された AI システムの使用に関するデータを定期的に収集および分析して、実際の設定で意図されたタスクを達成する際のパフォーマンスを評価するプロセスを指します。AI モデルのパフォーマンスの評価には、特定のアプリケーションに応じたさまざまなパフォーマンスメトリックと基準が含まれます。通常、このモニタリングの目的は、これらの AI システムの実際のパフォーマンスを評価し、パフォーマンスの低下または変更(例えば、データドリフトのため)を検出し、誤用のインスタンスを特定し、安全性またはユーザビリティの問題に対処することです。</p> <p>Source: DH/AI Glossary</p>
人工知能システム (AI システム)	<p>全体または一部が AI を使用して動作するデータシステム、ソフトウェア、ハードウェア、アプリケーション、ツール、またはユーティリティ。</p> <p>Source: Executive Order 14110, (October 30, 2023), Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence</p>
継続的機械学習	<p>以前の知識/情報を保持しながら、時間の経過とともに新しいデータや経験を取り入れることで、モデルのパフォーマンスを適応させる能力。モデルの変更は、特定の入力セットに対して、変更が実装される前と後で出力が異なる可能性があるように実装されます。通常、これらの変更は、新しいデータの分析に基づいてパフォーマンスを向上させることを目的とした、明確に定義されたプロセスを通じて実装および検証されます。ロックされたモデルとは対照的に、継続的機械学習モデルには、その動作を変更するための定義された学習プロセスがあります。</p> <p>Source: DH/AI Glossary</p>



データドリフト	<p>展開されたモデルが時間の経過と共に受け取る入力データ分布の変化を指します。これにより、モデルのパフォーマンスが低下する可能性があります。これは、基になるデータのプロパティが変更された場合に発生します。データドリフトは、予測モデルの精度と信頼性に影響を与える可能性があります。</p> <p>たとえば、医療用 AI に対応した製品では、モデル開発に使用されるデータと臨床運用で使用されるデータとの間に統計的な差異が発生する可能性があります。これは、医療行為やトレーニングと臨床使用の間の使用状況の違いによるものです。また、患者の人口統計、病気の傾向、データ収集方法の時間の経過による変化もあります。</p> <p>Source: <i>DH/AI Glossary</i></p>
デバイス	<p>機器、装置、器具、機械、装置、インプラント、体外試薬、またはその他の類似または関連する物品（構成部品または付属品を含む）。</p> <p>(A) 公式の米国薬局方または米国薬局方で認められています。</p> <p>(B) ヒトまたは他の動物の疾患またはその他の状態の診断、または疾患の治療、緩和、治療、予防に使用することを目的としています。</p> <p>(C) ヒトまたは他の動物の体の構造または機能に影響を与えることを目的としており、ヒトまたは他の動物の体内または体内での化学作用によってその主な目的を達成することはなく、その主な目的を達成するために代謝されることに依存しません。「デバイス」という用語には、セクション 520 (o) に従って除外されるソフトウェア機能は含まれません。</p> <p>Source: <i>Section 201(h)(1) of the Federal Food, Drug, and Cosmetic Act</i></p>
説明可能性	<p>「AI システムの動作の基礎となるメカニズムの表現を指す」。(出典:NIST)</p> <p>説明可能性は、ブラックボックスシステム(すなわち、内部の作業及び意思決定プロセスが透明でない又は容易に理解できないシステム)の不透明性を克服するのに役立つ可能性があります。これらの説明は、フリーテキストの説明、顕著性マップ、SHapley Additive exPlanations (SHAP)、またはデータからの関連する入力例など、さまざまな形式をとることができます。主な目的は、AI システムが特定の決定を行った"なぜ"という質問に答えることです。適切な説明可能 AI (XAI) メソッドを使用すると、より正確で、公平で、解釈可能で、透明性の高い AI システムを開発して、医療における人間の意思決定を安全に強化できる可能性があります。</p> <p>Source: <i>DH/AI Glossary</i></p>
基礎モデル	<p>大規模な、通常はラベル付けされていないデータセットと膨大な計算リソースを使用してトレーニングされた AI モデルは、モデルが特別に開発されてトレーニングされたものではないもの(すなわち、創発的能力)を含む、幅広いコンテキストに適用可能です。これらのモデルは、さらなるモデルを構築し、さらなるトレーニングを通じて特定の用途に適応させるための基盤として機能する(すなわち、微調整)。これらのモデルは、テキスト合成、画像操作、オーディオ生成など、さまざまな一般的なタスクを実行できます。これらのモデルは、トランスフォーマーのようなディープラーニングアーキテクチャに基づいており、ユニモーダルまたはマルチモーダルの入力データを使用できます。</p> <p>Source: <i>DH/AI Glossary</i></p>



<p>生成的人工知能 (Generative AI/ GenAI)</p>	<p>「派生した合成コンテンツを生成するために、入力データの構造と特性をエミュレートする AI モデルのクラス。これには、画像、ビデオ、オーディオ、テキスト、その他のデジタルコンテンツを含めることができます。」。(出典:E.O.14110)</p> <p>これは通常、入力データの統計的分布を近似することによって行われます。例えば、ヘルスケアでは、生成 AI を使用して合成医療データ(画像機能、テキストラベルなど)に注釈を生成し、アルゴリズムのトレーニング用のデータセットを拡張することができます。</p> <p>Source: <i>DH/AI Glossary</i></p>
<p>Confabulation (または Hallucination)</p>	<p>生成 AI システムが、ユーザーのプロンプトを満たすというプログラムされた目的を達成するために、誤ったまたは虚偽のコンテンツを生成し、自信を持って提示する現象を指します。</p> <p>Source: <i>National Institute of Standards and Technology. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile.</i> https://airc.nist.gov/docs/NIST.AI.600-1.GenAI-Profile.ipd.pdf</p>
<p>使用目的</p>	<p>成形品の表示に法的に責任を負う者 (またはその代理人) の客観的な意図を指します。意図は、そのような者の表現、成形品のデザインまたは構成、または成形品の流通を取り巻く状況によって示されることがあります。この客観的な意図は、例えば、そのような者またはその代理人による表示クレーム、広告事項、口頭または書面による陳述によって示されることがあります。客観的な意図は、例えば、そのような者またはその代理人が知りながら、成形品が表示も広告もされていない目的で提供または使用されている状況によって示されることがあります。ただし、企業は、そのような用途のために医療提供者によって処方または使用されているという企業の知識のみに基づいて、承認され、承認され、販売許可が与えられ、または市販前通知の適用を除外されているデバイスの未承認の新規用途を意図しているとはみなされません。成形品の意図された用途は、その製造業者によって州間取引に導入された後に変更されることがあります。例えば、包装業者、流通業者または販売業者が、成形品を受け取った人が意図した用途とは異なる用途を意図している場合、そのような包装業者、流通業者または販売業者は、新しい意図された用途に従って適切な表示を提供する必要があります。</p> <p>Source: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-801/subpart-A/section-801.4</p>



<p>大規模言語モデル (LLM)</p>	<p>大規模なテキストデータセットでトレーニングされた AI モデルの一種で、自然言語の単語間の関係を学習します。これらのモデルは、これらの学習パターンを適用して、受け取ったさまざまな入力やプロンプトに対する自然言語応答を予測および生成し、翻訳、要約、質問応答などのタスクを実行できます。これらのモデルの特徴は、膨大な数のモデルパラメーター(すなわち、訓練されたモデル内の内部学習変数)です。</p> <p>LLM は、基本的な言語パターンを超えて、より包括的な言語理解を開発することによって、基本的な AI モデルを構築します。たとえば、LLM のコンテキストでは、チャットボットは、人と人の会話を模倣する方法で、テキストまたは音声コマンドを介して LLM と人間間のコミュニケーションを可能にするプログラムです。</p> <p>Source: <i>DH/AI Glossary</i></p>
<p>ロックモデル</p>	<p>同じ入力適用されるたびに同じ出力を提供し、パラメーターや構成を更新できないため、使用しても変化しないモデル。AI 対応の医療製品の場合、ロックモデルは一貫したパフォーマンスを確保するのに役立ちます。</p> <p>Source: <i>DH/AI Glossary</i></p>
<p>Machine Learning (ML)</p>	<p>データに基づいてタスクのパフォーマンスを向上させるために AI アルゴリズムをトレーニングするために使用できる一連の手法。</p> <p>Source: Executive Order 14110, (October 30, 2023), Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence</p>
<p>機械学習モデル (ML モデル)</p>	<p>入力データの推論または予測を生成する数学的構成要素。このモデルは、データから学習する ML アルゴリズムの結果です。モデルは、データを処理して結果を導き出すために使用されるステップバイステップの手順であるアルゴリズムによってトレーニングされます。AI システム(例えば AI 対応医療機器)は、意図した目的を達成するために 1 つ以上のモデルを使用します。</p> <p>Source: <i>DH/AI Glossary</i></p>
<p>モデルの重み</p>	<p>入力に応答してモデルの出力を決定するのに役立つ AI モデル内の数値パラメータ。</p> <p>Source: Executive Order 14110, (October 30, 2023), Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence</p>
<p>ニューラルネットワーク</p>	<p>間の脳の構造に着想を得た計算モデル。相互接続されたノード、つまり「ニューロン」で構成され、データを受け取る入力層、データのパターンを処理および識別する 1 つ以上の非表示層、最終的なネットワーク出力を示す出力層の層に編成されています。</p> <p>Source: <i>DH/AI Glossary</i></p>
<p>パフォーマンス・メトリックス</p>	<p>AI のコンテキストでは、特定のタスクに対して望ましい出力を生成するモデルの能力を評価するために使用できる定量的または定性的な測定値。メトリックの選択は、特定のタスクとモデルの目的によって異なります。</p>



	<p>定量的指標の例としては、精度、精度、感度 (想起)、特異度、FL スコア、受信者動作特性曲線下面積 (AUC-ROC) などがあります。定性的指標には、ヒートマップの評価や視覚的な解釈が含まれる場合があります。これらの指標により、モデルの体系的な評価、比較、改良が可能になり、モデルが意図した目的を満たしているかどうかの評価に役立ちます。</p> <p>Source: <i>DH/AI Glossary</i></p>
Total Product Lifecycle (TPLC)	<p>FDA が採用している統合されたデバイスレビュー、追跡、報告、コンプライアンスのスキーム。TPLC アプローチにより、FDA はデバイスの誕生から陳腐化までのすべての規制活動を統合できます。(Source: Infusion Pump: Glossary FDA)</p> <p>このドキュメントでは、TPLC アプローチは、最初の構想から最終的な廃止および廃棄に至るまで、医療デバイスのライフサイクルのすべてのフェーズに対応しています。</p>
トレーニングデータ	<p>これらのデータは、AI システムの製造元が、モデルの重み、接続、コンポーネントの定義など、AI モデルを構築するための手順とトレーニングアルゴリズムで使用されます。</p> <p>Source: <i>DH/AI Glossary</i></p>
透明性	<p>機械学習対応医療機器 (MLMD (使用目的、開発、性能、および利用可能な場合はロジックを含む)) に関する適切な情報が、関連する対象者に明確に伝達される程度を表します。</p> <p>Source: U.S. Food and Drug Administration (2024). Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles. https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices-guiding-principles</p>
ウォーターマーク	<p>AI が作成したアウトプット (写真、ビデオ、オーディオクリップ、テキストなどのアウトプットを含む) に、一般的に削除が困難な情報を埋め込む行為。アウトプットの信頼性、出所、変更、または伝達の身元や特徴を確認する目的で行われます。</p> <p>Source: Executive Order 14110, (October 30, 2023), Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence</p>

References:

(Source): November 20-21, 2024: Digital Health Advisory Committee Meeting Announcement - 11/20/2024 | FDA
<https://www.fda.gov/advisory-committees/advisory-committee-calendar/november-20-21-2024-digital-health-advisory-committee-meeting-announcement-11202024#event-information>

Event Materials

Title	File Type/Size	Source Organization
DHAC Nov. 20, 2024 FDA Presentation (Victor Garcia and Aldo Badano)	pdf (2.10 MB)	
DHAC Nov. 20, 2024 Presentation (Bernardo Bizzo, M.D., PhD)	pdf (1.71 MB)	
DHAC Nov. 20, 2024 Presentation (Danielle Bitterman, M.D.)	pdf (1.84 MB)	
DHAC Nov. 20, 2024 Presentation (Faisal Mahmood, Ph.D.)	pdf (13.29 MB)	
DHAC Nov. 20, 2024 Presentation (GE HealthCare)	pdf (1.21 MB)	
DHAC Nov. 20, 2024 Presentation (Harvard Medical School)	pdf (6.11 MB)	
DHAC Nov. 20, 2024 Presentation (HCA Healthcare)	pdf (553.79 KB)	
DHAC Nov. 20, 2024 Presentation (Keith Dreyer, DO, PHD, FACR, FSIIM)	pdf (2.49 MB)	
DHAC Nov. 20, 2024 Summary Minutes	pdf (371.60 KB)	
DHAC Nov. 20, 2024 Transcript	pdf (876.95 KB)	
DHAC Nov. 20-21, 2024 24 Hour Summary	pdf (288.11 KB)	



Title	File Type/Size	Source Organization
DHAC Nov. 20-21, 2024 Agenda	pdf (268.34 KB)	
DHAC Nov. 20-21, 2024 Roster	pdf (165.55 KB)	
DHAC Nov. 21, 2024 FDA Presentation (Jessica Paulsen)	pdf (1.18 MB)	
DHAC Nov. 21, 2024 Presentation (Christopher A. Longhurst, MD, MS)	pdf (3.91 MB)	
DHAC Nov. 21, 2024 Presentation (Deloitte)	pdf (320.03 KB)	
DHAC Nov. 21, 2024 Presentation (Gabiella Waters)	pdf (2.75 MB)	
DHAC Nov. 21, 2024 Presentation (Google Health)	pdf (754.63 KB)	
DHAC Nov. 21, 2024 Presentation (Keith Dreyer, DO, PHD, FACR, FSIIIM)	pdf (904.46 KB)	
DHAC Nov. 21, 2024 Presentation (Kerri Haresign)	pdf (1.18 MB)	
DHAC Nov. 21, 2024 Presentation (Nina Kottler, MD, MS, FSIIIM)	pdf (714.84 KB)	
DHAC Nov. 21, 2024 Presentation (Yujan Shreshta, MD)	pdf (2.87 MB)	
DHAC Nov. 21, 2024 Summary Minutes	pdf (388.70 KB)	
DHAC Nov. 21, 2024 Transcript	pdf (851.59 KB)	
DHAC Nov.20-21, 2024	pdf (257.26 KB)	

Title	File Type/Size	Source Organization
Discussion Questions		
DHAC Nov.20-21, 2024 FDA Executive Summary	pdf (1.73 MB)	
DHAC Nov.20-21, 2024 Webcast Links	pdf (155.22 KB)	

Showing 1 to 27 of 27 entries

Webcast Information

November 20, 2024

YouTube:

<https://youtu.be/14ZTT-C-JCQ>External Link Disclaimer

November 21, 2024

YouTube:

<https://youtu.be/eBFVzBSR4T0>