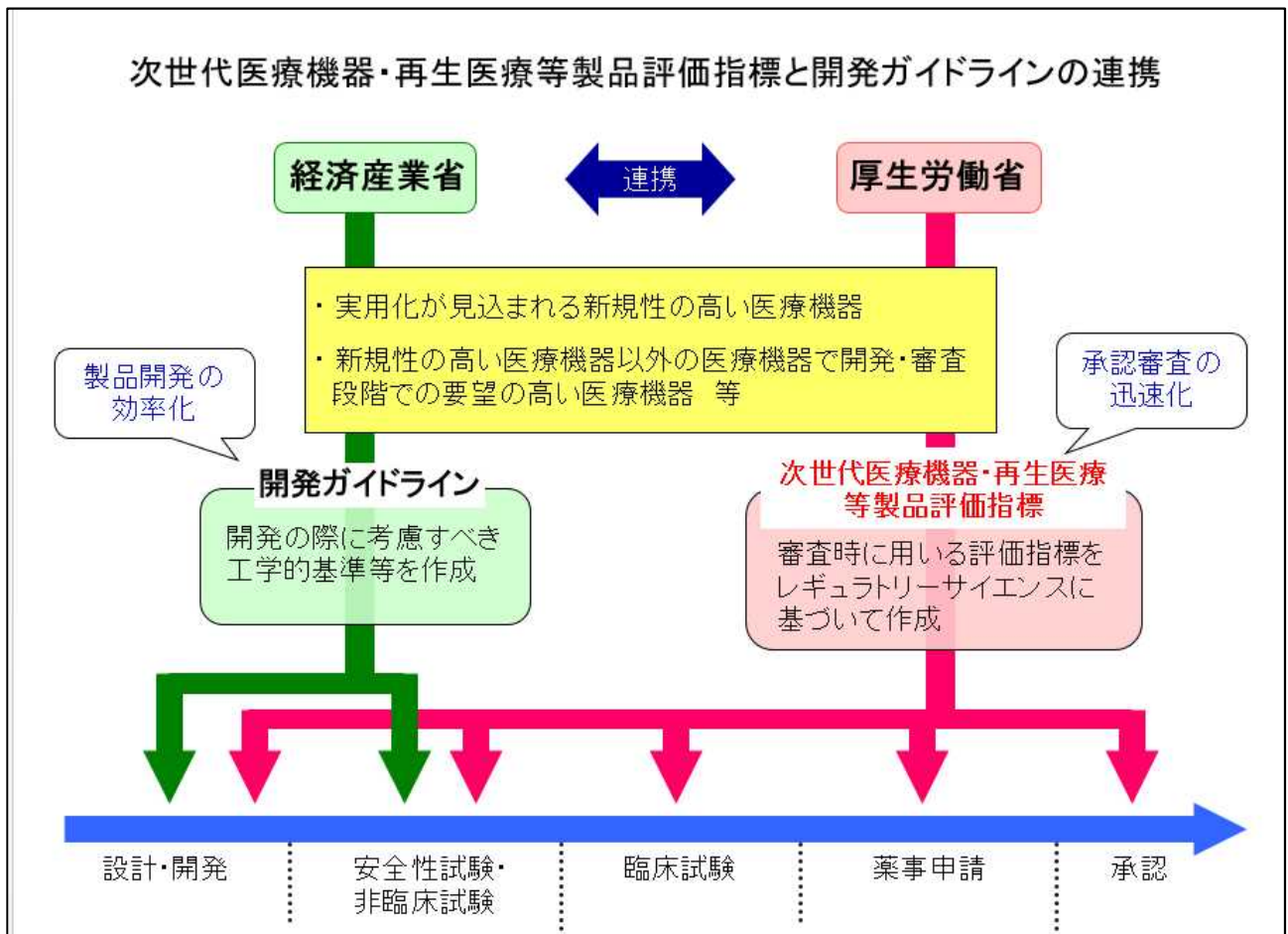


(参考) 評価指標ガイドライン作成事業について

医療ニーズが高く実用可能性のある医療機器及び再生医療等製品について、審査時に用いる技術的な評価指標等をあらかじめ作成・公表することにより、製品開発の効率化及び承認審査の迅速化を図ることを目的として、平成 17 年度からこれらの評価指標を検討し、公表しているところ。

評価指標は、申請資料の収集や承認審査の迅速化等の観点から、製品の評価において着目すべき事項を示すものであり、技術開発の著しい次世代の医療機器・再生医療等製品を対象として現時点で考えられる事項について示したものである。なお、評価指標は、法的な基準という位置付けではなく、製品開発の際に参考とされる事項をまとめたものである。



※ これまでに公表した評価指標（39 評価指標）

分野	評価指標名称	通知
再生医療	重症心不全細胞治療用細胞シート	平成22年1月18日付薬食機発0118第1号
	角膜上皮細胞シート	平成22年1月18日付薬食機発0118第1号
	角膜内皮細胞シート	平成22年5月28日付薬食機発0528第1号
	関節軟骨再生	平成22年12月15日付薬食機発1215第1号
	歯周組織治療用細胞シート	平成23年12月7日付薬食機発1207第1号
	自己iPS細胞由来網膜色素上皮細胞	平成25年5月29日付薬食機発0529第1号
	同種iPS(様)細胞由来網膜色素上皮細胞	平成26年9月12日付薬食機参発0912第2号
	鼻軟骨再生	平成27年9月25日付薬食機参発0925第1号
	ヒト軟骨細胞又は体性幹細胞加工製品を用いた関節軟骨再生	平成28年6月30日付薬生機参発0630第1号
	ヒト(同種)iPS(様)細胞加工製品を用いた関節軟骨再生	平成28年6月30日付薬生機参発0630第1号
	ヒト(自己)表皮(皮膚)再生	平成30年3月20日付薬生機参発0320第1号
	ヒト(同種)表皮(皮膚)再生	平成30年7月25日付薬生機参発0725第1号
	ヒト(同種)iPS(様)細胞加工製品を用いた亜急性期脊髄損傷(外傷性)の治療	令和3年2月26日付薬生機参発0226第1号
	体内埋め込み型能動型機器	次世代型高機能人工心臓の臨床評価
重症下肢虚血	重症下肢虚血疾患治療用医療機器の臨床評価	平成25年5月29日付薬食機発0529第1号
テーラーメイド医療用診断機器	DNAチップを用いた遺伝子型判定用診断薬	平成20年4月4日付薬食機発第0404002号
	RNAプロファイリングに基づく診断装置	平成24年11月20日付薬食機発1120第5号
ナビゲーション医療	骨折整復支援装置	平成22年1月18日付薬食機発0118第1号
	関節手術支援装置	平成22年1月18日付薬食機発0118第1号
	軟組織に適用するコンピュータ支援手術装置	平成22年5月28日付薬食機発0528第1号
	患者の画像データを用いた三次元積層造形技術によるカスタムメイド整形外科用インプラント等	平成27年9月25日付薬食機参発0925第1号
体内埋め込み型材料	整形外科用骨接合材料カスタムメイドインプラント	平成22年12月15日付薬食機発1215第1号
	整形外科用カスタムメイド人工股関節	平成23年12月7日付薬食機発1207第1号
	整形外科用カスタムメイド人工膝関節	平成24年11月20日付薬食機発1120第5号
	可動性及び安定性を維持する脊椎インプラント	平成26年9月12日付薬食機参発0912第2号
	三次元積層技術を活用した整形外科用インプラント	平成26年9月12日付薬食機参発0912第2号
	生体吸収性血管ステント	平成28年6月30日付薬生機参発0630第1号
神経機能修飾装置／活動機能回復装置	神経機能修飾装置	平成22年12月15日付薬食機発1215第1号
	活動機能回復装置	平成25年5月29日付薬食機発0529第1号
コンピュータ診断支援装置	コンピュータ診断支援装置	平成23年12月7日付薬食機発1207第1号
心臓カテーテルアブレーション	心臓カテーテルアブレーション装置	平成27年9月25日付薬食機参発0925第1号
微量診断装置	マイクロ流体チップを利用した診断装置に関する評価指標	令和元年5月23日付薬生機参発0523第2号
生体由来材料	生体由来材料を利用した新規機能を有する医療機器に関する評価指標	令和元年5月23日付薬生機参発0523第2号
血流解析シミュレーションソフトウェア	血流シミュレーションソフトウェアに関する評価指標	令和元年5月23日付薬生機参発0523第2号
人工知能	人工知能技術を利用した医用画像診断支援システムに関する評価指標	令和元年5月23日付薬生機参発0523第2号
ホウ素中性子補足療法(BNCT)	ホウ素中性子捕捉療法用加速器型中性子照射装置システムに関する評価指標	令和元年5月23日付薬生機参発0523第2号
在宅医療機器	在宅医療機器に関する評価指標	令和2年9月25日付薬生機参発0925第1号
難治性創傷治療機器	難治性創傷治療機器に関する評価指標	令和2年9月25日付薬生機参発0925第1号
麻酔支援装置	クローズドループ制御システムを有する医療支援装置に関する評価指標	令和3年6月30日付薬生機参発0630第4号

乳がん診断支援装置に関する評価指標(案)

1. はじめに

乳がんは、女性が罹患する「がん」の中で最も多く、30代から64歳までの女性における「がん死亡率」の第1位を占めている。発症時期は、働き盛りの40代後半に一つの大きなピークがある。ステージIの早期に治療を開始した場合、100%に近い5年生存率が得られるが、がんが進行したステージIVでは37.1%に低下することから、定期的な乳がん検診による早期発見・早期治療の重要性が提唱されている。しかし、我が国における乳がん検診の受診率は41%であり、欧米(75%)と比較して低い状況にある。受診率が低い要因としては、様々な問題が指摘されているが、診断装置自体が抱える課題も挙げられる。我が国の乳がん検診は、2000年初頭まで視触診により行われてきたが、死亡率の減少効果が不十分であると共に、精度管理が難しいことから、検診指針が変更された。現在、視触診は必須でなく、乳がん死亡率の減少効果が唯一証明されているX線マンモグラフィによる検査が一般的に行われている。しかし、X線マンモグラフィは被ばくや、乳房圧迫に伴う被験者への身体的負担を生じる問題がある。若年層に多く観察される、脂肪が少なく乳腺密度が高い「高濃度乳房」の場合、X線マンモグラフィでは、がんの腫瘤影が乳腺自体の白い陰影に覆われてしまい、がんを検出できない(偽陰性)可能性があることも知られている。我が国において、乳房超音波検査(エコー)は乳がん検診の検査方法として利用されてきた。エコーは高濃度乳房の多い40歳代女性の早期浸潤癌の発見率向上が認められている。しかし、エコーによる乳がん死亡率減少効果は示されていないため、その有効性の証拠は不十分であり、死亡率減少効果の代替指標として、累積進行乳がん罹患率減少効果が検討されている。エコーの課題としては偽陽性増加への懸念、画像の再現性や記録が検査担当者の技量に依存することやその精度管理が難しいこと等が挙げられる。

近年、これらの課題を解決し得る新たな診断支援装置のモダリティとして、マイクロ波イメージング、超音波CT、光超音波イメージングの開発が進められており、ともに臨床試験を通じて実用化段階にある。マイクロ波イメージングはX線と比較して低エネルギー、非電離性の特長を有しており、誘電コントラストを利用して乳がんの有無を診断する方法である。リアルタイム画像化技術も開発中であると共に、安全でより高頻度なスクリーニング技術として有望視されている。超音波CTは、リングアレイで乳房を取り囲み、超音波散乱像、音速分布像、減衰率分布像を撮像する装置である。MRIに匹敵し得る解像度が得られており、安価で非侵襲な乳がん診断方法として期待されている。光超音波イメージングは、生体にナノ秒幅のパルス光を照射した際に、光音響効果によって生じる超音波を受信する技術である。非侵襲で微細血管の構造情報や、酸素飽和度に関連する指標等の機能情報を高解像度に可視化できる新規イメージング技術であり、乳がん診断への応用も含め実用化に大きな期待が寄せられている。

このような背景を踏まえ、本評価指標では、これら3つの新規診断支援装置の品質、有効性、安全性を科学的根拠に基づいて適正且つ迅速に評価するための留意点を取りまとめた。

2. 評価指標の対象

本評価指標においては、主にマイクロ波イメージング、超音波CT、光超音波イメージングからなる新規モダリティを用いた乳がん診断支援装置を対象とする。

開発する乳がん診断支援装置が本評価指標に該当するか判断が難しい場合には、必要に応じて独立行政法人医薬品医療機器総合機構(PMDA)に相談すること。

3. 評価指標の位置づけ

本評価指標は、技術開発が著しい機器を対象とするものであることを勘案し、現時点で重要と考えられる事項を示したものである。今後の技術革新や知見の集積等を踏まえて改訂されるものであり、承認申請内容に対して拘束力を持つものではない。本評価指標が対象とする製品の評価にあたっては、個別の製品の特性を十分理解した上で、科学的な合理性を背景にして、柔軟に対応する必

47 要がある。本評価指標の他、国内外のその他の関連ガイドラインを参考にすることも考慮すべき
48 である。

49

50 4. 評価に対して留意すべき事項

51 (1) 基本的事項

52 1) 開発の経緯、臨床的位置づけ（乳がんの可能性のある病変の拾い上げを目的とする検診、良悪
53 性鑑別等の診断精査等）、開発品目の仕様、開発品目及び類似品の国内外での使用状況、装置
54 の設計とシステムの原理（アルゴリズムを含む）、標準的な使用方法等を明確に示すこと。

55

56 2) 以下の事項を参考として、システム全体の設置、運用等に必要な項目を規定し、評価するこ
57 と。

58

59 ①ソフトウェア（OS: Operating system 及びアルゴリズムを含む）

60 ②設置条件

61 (ア) 重量（使用場所の床に要求される耐荷重条件）

62 (イ) 寸法（格納時も含めて）

63 (ウ) 転倒防止対策

64 (エ) 水平及びその他必要な事項

65 ③騒音・振動

66 ④保守点検とその内容

67 ⑤トレーニング計画の必要性とその内容

68 ⑥使用者向け操作マニュアル等の文書化とその内容

69 ⑦ソフトウェアライフサイクルプロセス（参考：JIS T 2304）

70 ⑧リスクマネジメント（参考：JIS T 14971）

71 装置に限らず、治療の場における事故等、ワーストケースを想定した場合の対応について検
72 討する必要がある。

73

74 3) 最終製品と試験装置との同等性

75 試験装置を用いて非臨床及び臨床試験を実施する場合は、最終製品との同等性を評価すること。

76

77 (2) 非臨床試験に関する事項

78 以下に示す各事項等を通して、システム全体の安全性及び有効性の評価を適切に行うこと。

79

80 1) 安全性に関する評価

81

82 ①モダリティ共通事項

83 (ア) 電気的安全性及び電磁両立性（参考：JIS T 0601-1、JIS T 0601-1-2）

84 (イ) 機械的安全性（参考：JIS T 0601-1）

85 (ウ) 安全機構の種類、構造及び妥当性

86 ・アラーム（参考：IEC 62642-1、IEC 60601-1-8 等）

87 ・緊急停止対策（参考：ISO 10218-1、ISO 13850、IEC 60204 等）

88 ・緊急停止装置及びその構造

89 ・緊急停止する条件（撮影者の意に反する誤動作、安全機構作動時等）

90 ・停止中の患者及び医療従事者への安全性の確保（装置姿勢保持等）

91 ・緊急停止後装置の再稼働方法

92 ・誤動作予防対策（ユーザーインターフェース）

- 93 ・その他に必要な機構
- 94 (エ) 生物学的安全性 (参考: JIS T 0993-1)
- 95 (オ) 耐久性
- 96 ②マイクロ波イメージング
- 97 回転機構等を有する場合は、その安全性について評価すること。(参考: ISO 10218-1 等)
- 98 ③超音波 CT
- 99 超音波の出力に関して、以下の項目等について評価すること。
- 100 (ア) 超音波の機械的作用 (メカニカルインデックス: MI) *1
- 101 (イ) 超音波の熱的作用 (サーマルインデックス: TI) *1
- 102 撮像対象となる組織によって、TIS (軟部組織)、TIB (骨表面)、TIC (頭蓋骨表面) が
- 103 用いられている。乳がんを対象とする撮像であるため、装置構成に応じて、TIS、もしくは
- 104 は TIS と TIB の両方を考慮する必要がある。
- 105 (ウ) 体表の温度上昇 (参考: JIS T 0601-2-37 等)
- 106 *1: (公社) 日本超音波医学会・機器及び安全に関する委員会編纂の「超音波診断装置の安全性
- 107 に関する資料」等を参考とすること。
- 108 ④光超音波イメージング
- 109 以下の項目について評価すること。なお、当該装置が超音波を出力する場合は、上記③項「超
- 110 音波 CT」の記載内容に準じて評価すること。
- 111 (ア) レーザ光に対する保護 (参考: JIS C 6802、JIS C6802 等) *2
- 112 (イ) レーザクラスに応じた対策*3
- 113 *2: 薬生機審発 0629 第 4 号「レーザ医療機器の承認申請の取扱いについて」の「(別添 1) レ
- 114 ーザ医療機器に共通した性能及び安全性に関する評価指針」を参考とすること。
- 115 *3: 厚労省基発第 0325002 号「レーザ光線による障害の防止対策について」を参考とすること。
- 116
- 117 2) 性能に関する評価
- 118 マイクロ波イメージング、超音波 CT 及び光超音波イメージングの各装置の非臨床試験における
- 119 性能評価に関しては、モダリティ毎に画像化の手法や得られる診断情報が異なることから、その性
- 120 能については以下の事項を参考に個別に評価すること。
- 121
- 122 ①モダリティ共通事項
- 123 (ア) End-to-end 性能評価
- 124 ・システム全体としての性能の評価
- 125 (イ) 外乱ノイズの影響の評価
- 126 (ウ) 装置の動作状況の表示
- 127 (エ) 自己診断機能 (動作精度のバリデーションを含む)
- 128 (オ) 目的に応じた造影剤使用時の描出能
- 129 (カ) 時間分解能、空間分解能、コントラスト分解能
- 130 (キ) 検査時間、画像表示時間
- 131 (ク) 診断部位の網羅性
- 132 (ケ) 保存容量 (圧縮に伴うデータロスの有無等)
- 133 (コ) 再現性
- 134 ②マイクロ波イメージング
- 135 (ア) マイクロ波信号発生及び受信装置の評価
- 136 ・使用周波数帯
- 137 ・実効放射電力
- 138 ・信号発生及び制御方法

- 139 ・アンテナ放射特性及び放射効率
140 ・増幅器の位相及び利得の周波数特性
141 (イ) 送受信素子の設置及び移動
142 ・回転機構等の制御方法
143 ・素子固定治具の電磁気的特性
144 ・RF ケーブルの入出力特性
145 (ウ) 画像診断の方法と精度
146 a) レーダ方式
147 ・画像化精度及び強度情報の信頼性
148 b) トモグラフィ方式
149 ・複素誘電率の再構成精度
150 ③超音波 CT
151 (ア) 散乱像 (参考: JIS T 1501)
152 生体模擬ファントムを用い、以下の項目について評価すること。なお、画像作成プロセスに
153 おいて音速補正等を実施する場合は、その影響の検証も行うことが望ましい。
154 ・ペネトレーション深度 (超音波が到達して画像化が可能な深度)
155 ・スライス厚み評価
156 ・音響作動周波数及び帯域幅
157 ・表示精度 (画像中の距離スケールの精度)
158 ・計測精度 (計測モードにおいて、距離や面積を表示する場合の精度)
159 (イ) 音速再構成像、減衰率再構成像
160 これらの再構成撮像法に関しては、未だ性能評価方法が確立していない。当該手法で評価す
161 る際は、その妥当性及び精度を検証すること。
162 ④光超音波イメージング
163 評価項目は、開発機器の装置構成や特性等に応じて選定する必要がある。ファントムを用いた
164 試験及び動物試験において、血管の 3D 形状、酸素飽和度に関連するパラメータを指標とした性
165 能評価項目の一例を以下に示す。これらの試験で病態を再現することが困難な場合、臨床試験に
166 おける結果から必要な基本性能の数値等を算出し、その特性に基づいて設計されたファントム
167 を用いた試験等で評価すること。
168 (ア) 撮像範囲、深達長と撮像時間
169 (イ) 酸素飽和度に関連する指標 (S-factor 等) の精度
170 (ウ) 距離分解能
171
172 3) *in vivo* 試験 (動物試験)
173 動物試験の実施が必要な場合は、以下の事項に留意して適切に評価すること。
174
175 ①試験動物
176 (ア) 動物の種類とヒトへの外挿可能性 (解剖学的、生理学的特徴等)
177 ②試験プロトコル
178 (ア) 評価項目、評価基準、評価方法、評価期間及び評価者
179 (イ) 計測データ (生理学的、機械的及び電氣的データ、レーザ強度等)
180 (ウ) 例数の設定
181 ③評価にあたって考慮すべき点
182 (ア) 診断の達成状況 (肉眼病理観察や組織病理評価等)
183 (イ) 生体に対する有害事象の程度及び頻度
184 (ウ) 動物試験で確認する項目に関わる機器の不具合

- 185 (エ) 動物試験で得られる結果とシミュレーションとの差異
186 ④マイクロ波イメージング特有の項目
187 動物試験においては、国際非電離放射線防護委員会のガイドラインによる電磁界生体曝露指
188 標を順守すること。
189
190 4) *in silico* 試験
191 生体を用いた試験とは異なり、数学的なシミュレーションモデルを用いた評価試験を、以下 *in*
192 *silico* 試験と呼ぶ。*in silico* 試験により評価可能であることが十分に検証された項目に関しては、
193 妥当性確認や精度検証のために、*in silico* 試験の利用が想定される。*in silico* 試験により評価
194 を行う場合は、以下の事項に関して科学的根拠に基づき適切に評価すること。
195
196 ①機器の一部の評価における、シミュレーションモデルを用いた *in silico* 評価の実施
197 ②評価に用いるシミュレーションモデル
198 ③シミュレーションモデルの使用法、パラメータの選択、設定
199 ④シミュレーション結果の解釈
200
201 (3) 臨床試験（治験）に関する事項
202 1) 試験デザイン
203 試験デザインは、開発品の臨床的位置づけ（集団検診又は精密検査等）及び試験の目的を踏ま
204 えて、適切に計画される必要がある。試験の妥当性については、必要に応じて PMDA に相談する
205 こと。
206 有効性、安全性評価の項目は、撮影者、読影者の技量、ラーニングカーブ等により大きく影響
207 を受けることが考えられる。個々の習熟度のばらつき等による機器の評価への影響を抑えるた
208 め、使用方法や読影方法について事前に検討すると共に、撮影および読影に対する一定のトレー
209 ニングを行う等の配慮が必要である。
210
211 2) 施設数及び症例数
212 臨床試験の目的や主要評価項目等を踏まえ、科学的かつ統計学的な根拠に基づき、当該医療機
213 器の有効性、安全性の評価に適切な施設数及び症例数とする。ただし、解析においては一症例に
214 複数病変の可能性もあるため、病変数をもって統計処理する。また、信頼できる海外データを承
215 認申請の添付資料として使用できることがあるが、そのみで臨床評価を行うことができるかど
216 うかについては十分に検討すること。
217 対象病変は、腫瘍性病変、T1（腫瘍径 2cm まで）サイズの病変を主体とする。最終的には、針生
218 検、吸引式乳房組織生検もしくは手術時に、病理組織学的診断（乳がんにおいては浸潤癌、非浸潤
219 癌の区別）により確認すること。各症例に関しては、以下の項目について明記すること。
220
221 ①乳がん治療の有無
222 ②乳腺腫瘍（良悪性を問わない）摘出術や乳房温存術の有無
223 ③年齢及び体格（体重、身長、乳房サイズ及び胸郭形状）
224 ④授乳の有無
225 ⑤妊娠の可能性
226 ⑥月経情報
227 ⑦ペースメーカー等埋植機器の有無
228 ⑧その他、必要な事項
229
230 3) 評価期間

- 231 介入後、適切な期間を設定し、安全性・有効性を評価すること。
- 232
- 233 4) 有効性に関する評価
- 234 読影は盲検化し、一定水準以上の資質を有する2名以上の読影医による独立判定で行う。結果
- 235 は多施設研究による中央判定が望ましい。
- 236 開発品の特性や臨床的位置付けを踏まえ、ACR BI-RADS[®]や本邦の各種ガイドラインに準じた診
- 237 断基準を設定した上で、以下を参考に適切な評価項目を設定すること。診断の客観性を担保する
- 238 ため、第三者医師を含む評価委員会等を設置することが望ましい。なお、画像診断アルゴリズム
- 239 に人工知能(AI)技術を利用する場合は、令和元年5月23日付け薬生機審発0523第2号：厚生
- 240 労働省医薬・生活衛生局医療機器審査管理課長通知「次世代医療機器・再生医療等製品評価指標
- 241 の公表について」別紙4「人工知能技術を利用した医用画像診断支援システムに関する評価指標」
- 242 を参考とすること。
- 243
- 244 ①モダリティ共通主要評価項目
- 245 (ア) 従来法(X線マンモグラフィ、乳房超音波検査(エコー)装置、CT、MRI、針生検等)と開
- 246 発品との要精査対象病変発見に係る正診率、感度及び特異度等の比較(病変サイズ、がん
- 247 及び良性病変、浸潤癌、非浸潤癌、組織型、肉眼型、サブタイプ、ステージ毎等)
- 248 (イ) 切除標本との対比(肉眼・病理所見)
- 249 (ウ) 良悪性鑑別
- 250 (エ) 乳房構成又は乳腺密度
- 251 ②モダリティ共通副次評価項目
- 252 (ア) 撮像可能範囲
- 253 (イ) 乳がん以外の乳房内病変の撮像能
- 254 ③超音波CT特有の項目
- 255 (ア) 脱気水清浄度及び脱気度(読影に影響する可能性のある場合)
- 256 ④光超音波イメージング特有の項目
- 257 (ア) 酸素飽和度に関連する指標(S-factor等)の精度
- 258 (イ) 血管像の分岐数VBP(Vascular Branching Point)
- 259 (ウ) 血管像の密度
- 260 (エ) 脱気水清浄度及び脱気度(読影に影響する可能性のある場合)
- 261
- 262 5) 安全性に関する評価
- 263 熱的作用を含めた有害事象の種類、発生頻度及びその程度について評価すると共に、アンケート
- 264 により快適性や安全性について評価することが望ましい。モダリティ特有の項目については以
- 265 下の事項を参考に個別に評価すること。
- 266
- 267 ①マイクロ波イメージング
- 268 (ア) 電磁界生体曝露指標(参考：国際非電離放射線防護委員会ガイドライン)
- 269 (イ) 保護シール等の安全性
- 270 ②超音波CT
- 271 (ア) 水槽内の衛生管理
- 272 ③光超音波イメージング
- 273 (ア) 水槽内の衛生管理
- 274 (イ) レーザからの保護

令和3年度次世代医療機器・再生医療等製品評価指標検討会(厚生労働省)／
医療機器開発ガイドライン評価検討委員会(経済産業省)合同検討会

乳がん診断支援装置審査ワーキンググループ 委員名簿

座長：橋爪 誠 社会医療法人北九州病院 北九州古賀病院 院長
九州大学 名誉教授

副座長：椎名 毅 京都大学大学院医学研究科 教授
(兼任：日本生体医工学会推薦専門家)

学会推薦専門家(五十音順)：

植松孝悦 静岡県立静岡がんセンター 乳腺画像診断科 兼 生理検査科 部長
(日本乳癌学会)

中島康雄 画像相談クリニック 院長
聖マリアンナ医科大学 名誉教授 (日本乳癌検診学会)

中村清吾 昭和大学医学部外科学講座乳腺外科学部門 教授
昭和大学病院 ブレストセンター長 (日本超音波医学会)

委員(五十音順)：

木寺正平 電気通信大学 大学院情報理工学研究科 准教授

高尾信太郎 兵庫県立がんセンター 乳腺外科 部長

高木 周 東京大学 大学院工学系研究科 教授

山賀 郁 京都医療センター 乳腺外科 医師

人工知能技術を利用した医用画像診断支援システムに関する評価指標

1. はじめに

コンピュータ診断支援（Computer-Aided Diagnosis:以下 CAD）とは、X線画像に代表される放射線画像をはじめとする医用画像に対して、コンピュータで定量的に解析された結果を「第2の意見」として利用する「医師による診断の支援」である。

当該診断支援を行う医療機器に関しては、「次世代医療機器評価指標の公表について」（平成23年12月7日付け薬食機発1207第1号厚生労働省医薬食品局審査管理課医療機器審査管理室長通知）別添3「コンピュータ診断支援装置に関する評価指標」にて公表しているところであるが、近年の計算機技術の著しい発展により実用化が可能となった学習データのさらなる追加（市販後学習）等により市販後に性能が変化する人工知能技術を利用した診断支援装置については述べられてない。また、人工知能技術の中でも、近年着目されている深層学習では、最終結果を導き出すアルゴリズムがブラックボックス化している点が特徴であると同時に、市販後学習等による性能変化の内容も基本的にはアウトプットでしか評価できないこと等から、その評価や運用に当たって従来にない新たな課題が生じることが指摘されている。当該診断支援装置を用いた際、学習データに含まれていなかったタイプの事例に対する性能をどのように評価するのか（又はそのような事例にどのように対応するのか）、学習データの真正性や偏り等を考慮してどのように学習の品質確保を行うのか、さらに、市販後の性能変化に対する評価方法等に関して、従来とは異なる新しい視点も考慮する必要がある。

このような状況に鑑み、本評価指標は、上記通知も踏まえて、人工知能技術を利用した医用画像診断支援システムの有効性と安全性を評価する際の問題点や留意すべき点について、現時点での考え方をまとめたものである。

2. 用語の定義・説明

本評価指標における用語の定義等は以下のとおりである。

(1) 医用画像診断支援システム

臨床現場において、医師が様々な画像撮影装置を用いて画像診断（読影）を行う際に参考となる情報を提供するシステム又はソフトウェアであり、CADe や CADx の機能を持つ。支援の方式としては使用者（医用画像診断支援システムを実際に臨床現場で使用する医師等）に対する支援形式に応じ、first reader、second reader、concurrent reader 等が存在する。

1) CADe (Computer-Aided Detection)

画像上で病変の疑いのある部位をコンピュータが自動検出し、その位置をマーキング

する機能を有する単体ソフトウェア又は当該ソフトウェアが組み込まれている装置。コンピュータにより医用画像データのみ又は医用画像データと検査データの両方を処理し、病変又は異常値の検出を支援する。

2) CADx (Computer-Aided Diagnosis)

病変の疑いのある部位の検出に加え、病変候補に関する良悪性鑑別や疾病の進行度等の定量的なデータを数値やグラフ等として出力する機能を有する単体ソフトウェア又は当該ソフトウェアが組み込まれている装置。診断結果の候補やリスク評価に関する情報等の提供等により診断支援を行うものを含む。

注) CADe と CADx の定義は前出の通知より一部修正の上、引用

(2) 人工知能

人の高度な知能によって行われている推論、学習等を模倣するコンピュータシステム又はソフトウェア。本評価指標においては、特に、臨床で使用を開始した後、追加データによる機械学習（例えば、深層学習等）により、その性能が変化する機能（市販後学習機能）を有するものを対象とする。

(3) 機械学習

人工知能技術の一分野として大きく発展した技術であり、人間が行っている学習等の機能をコンピュータ（ソフトウェア）で実現するための手法。アルゴリズムはさまざまであり、判別分析法等の古典的手法から、決定木、ニューラルネットワーク、サポートベクターマシン等、多数の方法が提案されている。代表的な応用としてはクラス分類や回帰があるが、これらは CAD の目的と一致していることから、多くの CAD システム開発時に利用されてきている。

(4) 深層学習

機械学習の一つであり、最近特に注目されている学習方法。多層化したニューラルネットワークを大量のデータを用いて学習させることに特徴がある。ニューラルネットワークには、階層型ネットワークのような確定的モデルと、ボルツマンマシンのような確率的モデルがある。学習アルゴリズムには、教師あり、教師なし、半教師学習以外に強化学習等もある。

2006 年の Hinton らの研究に始まり、囲碁の世界王者を倒した AlphaGo や、各種コンペティションにおいて上位を占めたソフトウェアで使用され、当時の常識を塗り替えるほど突出した成績を示したことから盛んに研究されている。性能の飛躍的向上の一つの理由は、多層のネットワークを学習可能とする新しい技術の登場である。また、深層ニューラルネットワークの学習のための大量のデータと並列分散計算環境の整備も理由のひとつである。

(5) 転移学習

20年以上も前から認識されている機械学習の一分野であり、「新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題」とされている。

近年、深層学習でも盛んに転移学習が用いられるようになったが、その具体的な利用例として以下が挙げられる。

- 1) 大規模なデータセットで学習させたニューラルネットワークを特徴抽出器として利用し、得られた特徴を別のタスクへ転用する。例えば、一般の自然画像で学習済みのニューラルネットワークに医用画像を入力し、中間層の出力を特徴量としてその医用画像の認識・診断支援等に利用する場合は考えられる。
- 2) 別の目的で学習済みのニューラルネットワークを用意し、異なるタスクの学習データを用いて再学習させる。例えば、CT像の認識・診断支援を目的とする場合、1)に例示した自然画像で学習済みのニューラルネットワークにCT像と教師データを与えて再学習させる場合は考えられる。この場合には、一からニューラルネットワークを学習させる場合と比べて、少数のCT像と教師データで性能の高いニューラルネットワークが得られることが報告されている。

3. 本評価指標の対象

本評価指標は、医用画像診断支援システムのうち、人工知能技術を利用して様々な画像撮影装置で得られた臨床画像から、

- ・ 疾患名まで特定せず、病変の疑いがある部位のみ検出する（いわゆる CADe）
- ・ 病変の疑いがある部位の検出に加え、疾患名の候補を提示する（いわゆる CADx）
- ・ 疾患名の候補の提示に加え、それらの重み付け（順位づけ）まで行う（同上）

等の機能を有するシステム全て（CADe、CADx等、以降「支援システム等」と記述する。）を対象とする。

対象とする支援システム等は、人工知能技術の利用の形態や学習アルゴリズムの種類にかかわらず、機械学習等による市販後学習等に伴い診断支援性能が変化することを意図したものとする。なお、対象はあくまで診断を行う医師を支援するシステムであり、仮にシステムが想定外の挙動、誤動作をした場合に、使用者側でそれを検知できることが基本要件となる。

人工知能技術を利用した支援システムのうち、製造販売業者が、市販後学習により生じる性能変化に伴う品質管理を行うことが難しいもの、例えば、使用者が市販後学習させることにより施設ごとに異なる性能変化が生じ得るものに関しては、解決すべき事項が数多く存

在することから、本評価指標における検討の対象とすることは困難である。しかしながら、人工知能技術の著しい発展に伴い、これらの支援システム等が近い将来開発される可能性に鑑み、それらの品質、安全性及び有効性を確保するための基本的な考え方を別添に取りまとめた。

なお、開発する支援システム等が本評価指標に該当するか判断し難い場合は、必要に応じ厚生労働省医薬・生活衛生局医療機器審査管理課に相談すること。

4. 本評価指標の位置付け

近年、支援システム等の開発では、深層学習のようにデータから自動で学習をする機械学習の利用が主流になっている。特に、大量のデータを継続的に収集する仕組みが整備され、学習に用いるコンピュータの性能が飛躍的に向上したことにより、短期間で繰り返し学習をすることが可能となった。また、臨床で得られたデータや支援システム等の結果を自動的に収集し、市販後学習データとして人工知能に与えて性能を変化させることも技術的には実現可能となった。そのため、臨床で使用しながらその性能が段階的又は連続的に変化する支援システム等の実現が期待されている^{注)}。

本評価指標は、対象とする支援システム等において、現時点で考えられ得る問題点、留意すべき事項を示したものであるが、今後の更なる技術革新や知見の集積等を踏まえ改訂が必要なものであり、承認申請内容等に関して拘束力を有するものではない。

支援システム等の評価に関しては、個別の人工知能構築方法、使用目的及び製品特性を十分に理解した上で、科学的な合理性をもって、柔軟に対応することが必要である。なお、本評価指標以外に前出の「コンピュータ診断支援装置に関する評価指標」及び現存する国内外の関連ガイドライン等、例えば、平成28年3月31日付け厚生労働省医薬・生活衛生局医療機器・再生医療等製品担当参事官室事務連絡「医療機器プログラムの承認申請に関するガイダンスの公表について」別添「医療機器プログラムの承認申請に関するガイダンス」等を参考にすることも考慮すべきである。

注) 深層学習の市販後学習の対象には、ネットワークの結線（結合）の重み以外に、ネットワークのハイパーパラメータ（層数、特徴マップ数、畳み込みのフィルタサイズ、ユニット数等）も考えられる。これらのハイパーパラメータも含めた学習とは、例えば、パラメータの組み合わせごとに教師データとネットワークの出力間の誤差（損失）を最小化する学習を網羅的に行い、それらの中から誤差最小のハイパーパラメータを選択すること等が挙げられる。ただし、変更されたハイパーパラメータによって、承認の軽微な変更にあたるか、一部変更承認を要するものかを検討する必要があるため、当該変更の実施については、事前に独立行政法人医薬品医療機器総合機構と相談することを

推奨する。

5. 支援システム等で想定される主な課題と解決の方向性

(1) ブラックボックスとしての性質

機械学習、特に深層学習は、その特性上、ニューラルネットワークによる判断の過程がブラックボックスであり、アウトプットの判断根拠を明確に示すことが困難な場合がある。このため、通常の医療機器ではその性能確保のために示すことが求められる「動作原理」の明示が困難となる。そのため、承認審査においては、診断アルゴリズムを詳細に精査するというよりは、一般的な医療機器プログラムの評価と同様、そのインプットに対して所要のアウトプットが得られているかを確認することに重点を置き、その性能を評価することが適当と考えられる。

また、動作原理及びアルゴリズムの詳細を記述することが困難であるため、性能の確認が行われた事例以外（例えば、開発時の学習データセットに含まれていなかった希少事例）での挙動等が想定しにくいことも考えられる。このようなリスクに鑑み、機械学習、特に深層学習を利用した支援システム等においては、その性能を確保するために、使用方法及び目的に応じて、製造販売業者が対象疾患の検出率、偽陽性率、偽陰性率、検出に要する時間等、性能に係る値を適切に規定し、その性能が常に担保されていることを示すことが必要となる（6（3）項を参照）。

支援システム等が想定外の挙動を示した場合に、使用者に対して当該挙動を通知する仕組みも必要となる。例えば、想定外の挙動が判断可能な使用者のみに使用を許可する等で使用者側に一定の要件を求める、又は臨床での使用状況を逐次フィードバックすることで、製造販売業者側が性能検証可能となる仕組みを構築する等、支援システム等の目的に応じた適切な方策を用意することが望ましい。（6（4）項を参照）。

想定外の挙動、誤判定等を完全に無くすことは極めて難しいが、そのような情報を収集し、対策を講じることができるような運用方法もあらかじめ考慮しておく必要がある。

(2) 性能変化

支援システム等においては、人工知能技術を利用することで市販後学習による診断アルゴリズムの変化を受けてその性能が向上することが期待される反面、学習の内容等によっては意図しない性能変化が生じ、当初規定していた性能を下回る可能性も否定できない。このような利点とリスクとを勘案し、支援システム等には、継続的な性能検証によるほか、学習用データの要件や学習のプロセス等を規定し、学習の品質を確保するための対応策が求められることになる。

1) 継続的な性能の検証方法

臨床使用に伴って市販後学習が行われ、性能が連続的又は相当程度高い頻度で変化する場合には、支援システム等の品質、安全性及び有効性を確保する上で、性能が変化するたびに検証する必要がある。特にその内容、プロセス管理に加え、変化した性能の検証及びその管理が重要となる。このことを踏まえ、あらかじめ市販後に生じ得る性能変化水準を臨床上許容される範囲かつ統計学的な妥当性に基づいた範囲で目的に応じて規定しておくこと等とともに、その機器に適用可能な検証方法、変化した性能が規定された範囲を逸脱した際の対策等を定めておくことが求められる(6(3)、6(4)項を参照)。また、検証に使用したデータについては、入手元を明らかにした上でその使用根拠や妥当性を示すことが求められる(6(2)2)項を参照)。

2) 性能変化に伴う品質確保

人工知能の性能向上には、その人工知能の学習方式に適した市販後学習が必須であることから、その学習方式を明確にした上で、使用した市販後学習データの詳細を示すことが求められる。(6(2)2)項を参照)。

支援システム等においては、まず、一定量蓄積した臨床データを市販後学習データとして使用することで性能の変化を段階的に行う場合が考えられるが、その場合には、基本的にその性能変化のための学習及び性能検証は製造販売業者が実施することとなり、市販前に求められる要件と同様に性能変化後の品質を確保することが求められる(6(3)1)項を参照)。

一方、市販後学習により性能が連続的又は相当程度高い頻度で変化する機器の場合、製造販売業者の意図に反してその性能が低下し、臨床上許容される水準を下回る可能性がある。このような問題を防ぐために、原則、使用者はネットワークを介して支援システム等と接続する等、人工知能は製造販売業者が管理した上で、当該支援システム等の目的及びリスクを勘案した検証方法や対策を講じることが求められる(6(3)2)②項を参照)。

性能が変化した支援システム等の品質確保に係る作業は製造販売業者が実施することとするが、品質確保の目的で、使用者による市販後学習機能を有さない、誤使用や適応外使用を防止するため使用者を対象としたトレーニングを実施する、性能が下限以下となった場合の対策を使用者側の端末にも施す等、その支援システム等のリスクに応じて必要な追加対策を施すことが望ましい(6(4)項を参照)。

3) 薬事上の手続きにおける考え方

一般的には支援システム等の性能を向上させた結果、臨床上の有効性が向上する等製品の品質、有効性または安全性に影響を及ぼす場合、薬事上の手続きが必要になる場合があるが、その要求範囲は性能変化の内容及びその変化に伴うリスクを勘案した設定が必

要となる。

性能向上に係るバージョンアップが新規製品となる、又はその適用に費用が発生する可能性があることから、その性能変化は対象となり得る販売後の製品全てに適用されず、同じ製品であってもバージョンの異なるものが市場に存在することになる可能性も考えられる。このような、人工知能の流通体系の特殊性を考慮しながら、バージョンアップに関するフォローアップ体制や薬事上の効率的な手続きについても検討が必要になる。

(3) 責任の所在

人工知能を利用した医療機器においては、その使用に伴う責任の所在が曖昧になることについての懸念が議論されている。特に、支援システム等においては、その目的が「診断支援」であるにもかかわらず、医師が最終診断まで委ねてしまう危険性を含んでいる。よって、支援システムの保守、設計・仕様上の不具合や故障等のトラブル対応においては、通常の医療機器と同様、製造販売業者が責任を負うことは変わらないが、支援システム等の使用目的、使用方法等を使用者に対して明らかにし、使用者に対するトレーニング等を実施する等、適正使用のために必要な方策を検討し、実施の実効性を担保する必要がある。

(6 (4) 項を参照)。なお、平成 30 年 12 月 19 日付け医政医発 1219 第 1 号厚生労働省医政局医事課長通知「人工知能 (AI) を用いた診断、治療等の支援を行うプログラムの利用と医師法第 17 条の規定との関係について」において、当該プログラムを利用して診断、治療を行う主体は医師であり、医師がその最終的な判断の責任を負うことが示されていることも留意されたい。

6. 評価に際して留意すべき事項

(1) 基本的考え方

本評価指標で対象とする支援システム等は、その特性上、通常の医療機器と異なり、原理(実装する検出・診断アルゴリズム等)や設計仕様等のみで性能を確保することが難しいことから、性能に影響する要素、効果を確認できる範囲や限界等を規定する必要がある。当該要求事項を達成するためには、医学的かつ統計的に妥当な方法で、それらを規定し、支援システム等の性能を検証することが求められる。なお、支援システム等自体の品質、安全性及び有効性を確保するために必要となる評価は、原則として前出の「コンピュータ診断支援装置に関する評価指標」を参考に、目的に応じた性能水準を設定の上、実施することになる。

上述した内容を考慮した上で、現時点の科学水準を基に、本評価指標の対象となる支援システム等の評価において特に留意すべき点を示す。

(2) 品目の検出・診断の原理、学習、情報セキュリティ等に関する基本的事項

1) 検出・診断の原理 (アルゴリズム)

具体的にその記述が可能な検出・診断アルゴリズムにおいては、従前通り、そのアルゴリズムとプログラム概要を示すことが求められる。深層学習のように処理工程がブラックボックス化する原理を用い、承認申請時に検出・診断アルゴリズムを示すことが困難な場合には、設計開発時の検出・診断用ネットワーク構造とプログラム概要を示す必要がある。なお、技術の進展等によりブラックボックス化したネットワーク構造等の振る舞いが説明できるようになった場合には、可能な限りその情報を示すことが求められる。

2) 学習

本評価指標が対象とする支援システム等は、適切な学習データを使用した人工知能の学習によりその目的を達成するために必要となる性能を有することが求められる。よって、対象とする支援システム等の機構や規定された性能等に応じて、以下の項目例を参考にして必要な項目について内容を明確に示し、またそれらを使用した根拠及び妥当性を示す必要がある。

- ・ 学習アルゴリズムとプログラム概要
(教師あり、半教師あり、強化学習、自己学習、その他等を明確にすること)
- ・ データ (学習データ、バリデーショndata^{注1)} 及びテストデータ^{注2)} について、以下を参考に必要な項目を明記すること)

注1) バリデーショndata: 機械学習アルゴリズムのハイパーパラメータ (例: サポートベクターマシンの目的関数を決定するパラメータ、深層学習のネットワークの層数、特徴マップ数、畳み込みのフィルタサイズ、学習回数等) を決定するためのデータ

注2) テストデータ: システムの性能を評価・検証するためのデータ

- データソース (取得及び管理方法の明確化を含む。特に、市販後に支援システム等の使用に伴い取得した画像等を市販後学習データとする場合には、対象となる患者から同意を取得する方法も含む。)
- 画像データ取得時の撮影パラメータ
- 紐付けした臨床データの種類^{注3)} (アノテーション (例: 病変等の位置やサイズ (ラベル画像も含む。)、良悪性の鑑別結果) も含む。)
- 臨床画像データ以外のデータ使用の有無^{注4)} (Data Augmentation 等の目的で作成したコンピュータシミュレーションやファントム使用による画像データ等)
- 画像データの枚数、サイズ、濃度レベル数等
- 学習前に加工 (例: 匿名化) している場合にはその処理方法

- 臨床データの最終判断を行った医師等
- 臨床データ以外のデータを使用した場合は、その最終判断を行った責任者等
- テストデータを学習のプロセスから完全に切り離して管理するための方策

注3) これまでの医用画像診断支援システムは医用画像のみを用い、臨床的信頼度が高いアノテーションを用いて設計することがほとんどであった。しかし、技術の進歩により、近年では、学習データとして自然画像や人工画像、さらには紐づけされた臨床情報の信頼性が低いデータ（Weak label データ）を用いることが可能となってきたが、このようなデータを利用する場合は、その使用根拠及び妥当性を示す必要がある。また、テストデータに関する各種留意点に関しては、その評価段階に応じて6（3）項を参照すること。

注4) 特に転移学習の場合には、自然画像や別の医用画像、さらには人工画像等を学習データとして用いることが想定される。自然画像を用いた場合には画像中の対象部分、種類及び枚数等、別の医用画像を用いた場合には通常の医用画像を使用した場合と同等の事項、人工画像の場合にはその作成方法等について、その詳細を示すこと。

学習において使用する画像等をデータベース等から入手して使用する場合は、その妥当性を確認することが求められる。以下に、データベース等に関して承認申請時に記載が必要と考えられる点を例示する。なお、記載すべき項目は、データベースごとに異なるため、適切に選択、追加すること。

- ・ データベース運営者（学会、認定匿名加工医療情報作成業者等）の概要、組織体制等
- ・ 運営者が所有する事業計画書
- ・ 運営者が外部に委託している業務内容
- ・ データベースに保有するデータの種類
- ・ データベースの概要及び設計書
- ・ データ管理に係る各種手順書とその運用状況 等

上記の学習にかかわる全ての項目については、設計開発時の学習と市販後学習で共通の項目、相違する項目を区別して明記する必要がある。例えば、設計開発時の学習と市販後学習で学習アルゴリズムが異なる場合は、それぞれについて明らかにする必要がある。また、市販後学習により機械学習アルゴリズムのハイパーパラメータも再決定される場合は、変化する範囲をあらかじめ規定したうえで、そのハイパーパラメータの変化が規定範囲内となることを明らかにすることが求められる。さらに、市販後学習後のネットワークの変遷については性能変化ごとに必要な記録を保存し、後日検証可能な状態にする

ことが必要となる。

3) 使用環境及び情報セキュリティ

人工知能技術の利用において、その人工知能が使用者のコンピュータ上にあるのか又はネットワークを介して接続しているのかを明確にした上で、次の項目を示すことが求められる。

- ・ 画像撮影装置からのデータ転送方法
- ・ 使用に必要なソフトウェアのインストール等による画像撮影装置又はその他のソフトウェアへの干渉
- ・ 支援システム等がインターネット等外部機器と接続されている場合、平成 27 年 4 月 28 日付け薬食機参発 0428 第 1 号・薬食安発 0428 第 1 号厚生労働省大臣官房参事官（医療機器・再生医療等製品審査管理担当）、厚生労働省医薬食品局安全対策課長連名通知「医療機器におけるサイバーセキュリティの確保について」を参考にサイバーセキュリティ対策に必要な措置を講じていること
- ・ 支援システム等を使用して得られた臨床情報を市販後学習データとする場合は、以下のガイダンス等を参考に適切な管理を行い個人情報の保護に留意すること
 - 「医療・介護関係事業者における個人情報の適切な取り扱いのためのガイダンス」平成 29 年 4 月 14 日、個人情報保護委員会、厚生労働省
 - 「医療情報システムの安全管理に関するガイドライン第 5 版」平成 29 年 5 月、厚生労働省

注) 必要に応じて、個人情報保護法及び次世代医療基盤法も参考にすること

4) 併用する画像撮影装置等の条件

支援システム等が所定の機能を達成するにあたっては、あらかじめ併用する画像撮影装置で得られた医用画像データのフォーマットや解像度等が当該システム等の要求を満たす必要がある。また、市販されている特定の画像撮影装置全てに適用することが可能なものであっても、支援システム等の性能は、人工知能の学習に使用する医用画像データの撮像条件や教師データの作成方法等に影響されるため、必要に応じてそれらについて規定する必要がある。

(3) 安全性、品質及び性能評価

1) 承認申請時

支援システム等の評価においては、対象となる機器の目的やリスクに応じた評価が必要なため、品質、安全性及び有効性の評価に必要な規格値（検出率等）を明確にするとともに、その妥当性を示した上で適切な評価を行うことが必要である。

各種評価においては、6（2）項に記載した項目以外に支援システム等に特有の問題と

して、

- ・ 接続する画像撮影装置の機能（特に既存の画像撮影装置を利用可能なシステムの場合）
- ・ 責任者を含めた性能検証方法の明確化（治験による性能検証の場合は、「正解」となる最終的な確定診断を得るまでのプロトコール、後向き検証の場合は検証用テストデータに関して、その妥当性を示すために必要な情報を示すことが求められる）
- ・ 性能検証に使用したテストデータの種類、入手元及び妥当性（必要に応じて管理方法等も明らかにすること）
- ・ 市販後に生じ得る性能変化の範囲の規定（検出率の下限、偽陽性率及び偽陰性率の上限設定等）及びその規定根拠となる医学的又は統計学的データ等
- ・ 市販後に性能変化が発生しても、有効性及び安全性が確保されていることを確認するための方策（学習させるデータの入手方法、変化後の性能の検証方法等）
- ・ 性能変化に伴って生じ得る不具合

等、その目的に応じて考慮すべき項目を明確にする必要がある。

特に、性能検証用のテストデータセットは、学習データやバリデーションデータとは重複しないことが求められる一方、対象とする母集団の特徴を考慮して、十分な説明が可能となる質と量のデータが必要となる。しかしながら、機械学習の方法、目的及び対象に応じてその必要数は異なり一律に定義することは困難であるため、支援システム等の性能評価に使用したデータセット及び方法論を記載するとともにその妥当性検証スキーム等を個別に記載することが求められる。

2) 市販後

市販後、支援システム等に発生する性能変化に関しては、承認申請時に規定された性能検証方法に従い、市販後学習データにより変化した性能等が臨床上又は統計学的な観点からあらかじめ規定された範囲内を満たし、品質が確保できることを確認の上、その妥当性検証を行うことが求められる。

支援システム等を既存の画像撮影装置に接続して使用する場合、画像撮影装置の性能に依存して入手できる画像データの質が多様となることが想定される。この場合には、得られた画像データが学習データとして使用できる条件を規定する必要がある、その条件を満たさないデータは学習データとしないための対策が必要となる。また、性能検証及びそのために用いるテストデータセットに関する考え方は、基本的に前項と変わらない。加えて、市販後学習データ及びテストデータの臨床最終判断を行った医師等を明確にする必要がある点も、前項と変わらない。

なお、人工知能にネットワークを介して接続する支援システム等において、使用者に必

要な性能及び動作の詳細を漏れなく把握できるよう、性能変化が生じた際には使用者にその旨を通知するシステムを有することが望ましい。

以下に、性能変化の実現方法ごとに必要となる留意点を示す。

① バージョンアップごとの段階的な性能変化を意図する場合

製造販売業者が臨床現場で入手したデータ等を定期的に収集し、データの偏り等にも配慮しながら適切な形で人工知能に市販後学習させることにより支援システム等の性能を変化させる場合がこれに該当する。なお、臨床現場で入手したデータ以外を利用することも想定されることから、それらの利用においては承認申請時に明記するとともにその妥当性を示すことが求められる。

この場合、変化した性能の評価や、その変化に伴う安全性及び品質の確保は製造販売業者によってバージョンアップごとに行うこととなるが、それらの評価にあたっては承認申請時と同様に前出の「コンピュータ診断支援装置に関する評価指標」を参考に行うこと。

また、支援システム等の性能変化において必要となる薬事上の手続きについては、システムの特性にも依存するため、想定可能な範囲については独立行政法人医薬品医療機器総合機構と承認申請前に相談して決定しておくことが望ましい。

② 臨床使用に伴った市販後学習により連続的又は相当程度高い頻度でのバージョンアップを意図する場合

支援システム等を臨床現場で使用した際に入手したデータを市販後学習データとして使用し、連続的又は相当程度高い頻度で性能を変化させる場合には、使用する人工知能は製造販売業者が一括して管理し、各支援システム等とネットワークを介して接続していることが想定される。この場合、自動的に追加されたデータを用いた市販後学習により変化した支援システム等の性能や品質があらかじめ規定された範囲に収まることが特に求められることから、その学習プロセス及びその管理方法を示すことが求められる。加えて、性能が規定範囲に収まることを検証するための機構をあらかじめ付与するとともに、製造販売業者で定期的にその性能検証を実施することが求められる。なお、当該バージョンアップの場合には、想定外の挙動、誤判定等の問題が生じる可能性が高いため、その解決策（リスクマネジメント）を講じておくことが求められる。

(4) リスクマネジメント

1) 基本的考え方

支援システム等におけるリスクとしては、市販後学習の結果性能が変化し「人工知能が規定された変化範囲を逸脱した結果、誤った情報を提示すること」が挙げられる。バージョンアップごとの段階的な性能変化を意図する支援システム等の場合には、その性能変

化によって生じた問題は製造販売業者が解決してから上市することが前提となるため、そのリスクは低い。一方、市販後学習による性能変化が直ちに実装されることで性能が連続的又は相当程度高い頻度で変化する支援システム等の場合には、上記リスクが生じる可能性があるため、そのリスクを回避するためのシステムが必要となる。

また、上記リスクに加え、適応外で使用した際には、使用者が支援システム等における想定外の挙動に気付かない可能性も増大することが考えられるため、適応外使用を防ぐ方法を講じることが求められる。

2) 対策について

対策が必要となる支援システム等には、上記リスクを軽減するために、

- ・ 市販後学習データとする画像に関する詳細なログ(使用環境、診断結果、責任者等)を保存するための機能(6(2)を参照)
- ・ 支援システム等の内部データを利用した自己テスト等により、定期的に性能評価を行い、その時点での性能を使用者に示すことで最低限の性能を確保・確認できる機能
- ・ 上記機能により上記リスクが発生していることが確認された場合に、人工知能を市販時の状態に初期化又は変化直前の状態へと復元する機能
- ・ 上記リスク発生時に、必要に応じて、支援システム等を停止させる機能
- ・ 臨床での使用状況を逐次フィードバックして製造販売業者がその性能を適宜検証可能となる機能

等、当該支援システム等の承認申請時に確認された性能を確保するための機能、又は具体的な対策手順をその支援目的に応じて設定しておくことが求められる。

製造販売業者は、使用者は支援システム等の目的に応じた専門知識を有し、かつ、その使用等に関する教育(トレーニング)を受けた医師を始めとする医療従事者に限定され、確定診断、最終判断等、診療行為における最終判断は医師でなければならないことを使用者に明確にすることが求められる。加えて、使用者の誤使用や支援システム等の目的から逸脱した適応外使用を防ぐため、ユーザトレーニング等の適切な対策を講じておくことが求められる。なお、このユーザトレーニングは関連学会等と連携したものが望ましい。

7. 機器の特性に応じた追加留意事項(案)

(1) 市販後学習による性能変化が診断支援においてハイリスクとなる支援システム等の場合

例えば、病理診断においては、その診断が臨床における最終診断であることから、どのような形式であってもその性能が市販時から変化することは検出率及び正診率低下の可

能性を含むこととなり臨床上高いリスクを伴うこととなる。このような支援システム等の場合、その品質、安全性及び有効性を十分に確保する必要があることから、現時点では、市販後学習プロセス、性能変化及びその検証には必ず病理専門医の関与を要求することが望ましい。

(2) 支援システム等の対象画像が動画の場合

消化器内視鏡機器や超音波画像診断装置のように観察像が動画となる場合、人工知能に対する学習データはその動画をどのような形で利用したかを明確にする必要がある。よって、動画のフォーマットや取得方法を明示するほか、学習データへの変換方法を明らかにする必要がある。

また、当該支援システム等においては、動画の撮影条件及び質が撮像者の技能に左右されることから、学習データとして利用可能となる条件を規定するとともに、その妥当性を明らかにする必要がある。

以上

使用者等が性能変化に関与可能な人工知能技術を利用した 医用画像診断支援システムに関する基本的考え方

1. 背景

今回、人工知能技術を利用した医用画像診断支援システム（支援システム等）に関する評価指標案を作成するにあたり、近年の計算機技術の著しい発展に伴い、実現可能な様々な支援システム等を全て評価指標案の対象に含めるかについて、WG で議論が行われた。特に、人工知能の中でも、近年着目されている深層学習を利用した場合の可能性を考慮し、コンピュータ技術が発達することにより、現在の大型コンピュータと同レベルのスペックの汎用 PC が使用可能となることで、ネットワークを介さずに PC 中に人工知能を有し、特定の臨床現場で市販後学習を行い性能が変化するケースを対象に含めるかも議論された。

このケースには、市販後学習により生じる性能変化に伴う品質管理を製造販売業者が行うことが難しい支援システム等、例えば、施設又は使用者ごとの市販後学習による性能変化が可能なシステム、教師なしデータによる学習が可能なシステム等が含まれる。これらの支援システム等は、解決すべき事項が数多く存在することから、本評価指標の対象外とすることとした。しかしながら、人工知能技術の著しい発展に伴いこれらの支援システム等が近い将来開発される可能性に鑑み、それらの品質、安全性及び有効性を確保するための考え方を別添として取りまとめた。

2. 問題点と解決策について

(1) 性能変化

例えば、使用者による市販後学習により性能が変化する支援システム等の場合、開発者以外が主体となって市販後学習を行うことになるため、製造販売業者の開発意図と大きく異なる方向にその性能が変化・低下し、臨床上許容される水準を下回る可能性が否定できない。このような問題を防ぐためには、適切な質と量のテストデータによる検証に加え、当該支援システム等のリスクを勘案した対策を講じることが求められる。例えば、製造販売業者による品質管理を徹底するための方策を講じる、製造販売業者と同等の品質マネジメントシステムを臨床現場に求める、使用者を対象としたトレーニングを実施することに加え、性能が下限以下となった場合にはそのリスクに応じた追加対策を施すこと等が求められる。

支援システム等の目的は「診断支援」であることから、確定診断時における医師の最終的な判断によりそのリスクを回避することができると考えられるが、可能であれば生じたリスクを漏れなく製造販売業者に連絡する機能、又は手順を必要に応じて設けることが望ましい。

(2) 責任の所在

その機器の使用に伴う医療行為、適応外使用、不具合等に対する報告に関しては医師が責任を負うことが原則であるが、その原則を一様に適用できない可能性があることに留意して、製造販売業者及び医師の責任範囲を設定する必要がある。

3. 評価に際して考慮すべき事項（品質、安全性及び有効性の評価）

使用者が施設ごとに学習を行うシステムと、自己学習を行うシステムに関する留意点は以下のとおりである。

（１）使用者による施設ごとの市販後学習が可能な支援システム等

特定の医療機関で得られた臨床データや診断支援の結果を元にした市販後学習を使用者が実施するケースが該当する。この場合、市販後学習データの偏りにより、医療機関ごとに異なる方向に性能が変化し、製造販売業者による品質管理が困難となる。加えて、使用者が市販後学習を実施するため、製造販売業者の開発意図と大きく異なる方向にその性能が変化するリスクがある。そのため、性能評価においては、評価指標に記載した事項に加え、製造販売業者は、

- ・ 市販後学習プロトコル。特に使用者が施設ごとに行う市販後学習データの収集方法や品質管理等も含めたプロトコルの詳細
- ・ 市販後学習後に変化したシステムの性能の検証プロトコルに関する詳細。承認申請時のテストデータに対する性能評価結果のみならず、新しいテストデータを追加する場合には、その収集方法や品質の評価法と結果、及びシステムの性能評価結果に関する詳細

に留意し、製造販売業者が市販後学習を実施した場合と同等の品質マネジメントを臨床現場に求めるための工夫が必要である。

また、市販後学習に加え、変化した性能の評価や、その変化に伴う支援システム等の安全性及び品質の確保を使用者等が行う必要があるため、それらに係る評価は製造販売業者が実施する場合と同等の評価を行う必要がある。そのため、当該支援システム等を有する医療機関内に、以下の要件を充足した責任者の指名と、その責務、性能評価及び品質管理に関する手順書の整備等を求めることになると考えられる。

- ・ 対象となる支援システム等の原理及び学習方法を熟知していること
- ・ 市販後学習データの妥当性を示すことができること

なお、その性能変化の範囲が承認申請時に製造販売業者により規定された範囲内であることが前提となるため、性能評価に関しては製造販売業者が承認申請時に規定した検証方法を用いる必要があり、性能変化が製造販売業者により規定された範囲を超えることが生じないよう、評価指標 6（４）項「リスクマネジメント」に記載した事項に加え、当該支援システム等にはそれを防止するための機構、具体的には、そのリスクを製造販売業者に連絡する機能等を設けることが強く求められる。

（２）「教師なし」データによる自己学習が可能な支援システム等

2017年10月、Googleにより囲碁用人工知能において「教師なし」データを用いた自己学習により連続的な性能変化が可能となったことがNatureにて報告された。将来、同様の人工知能を利用して医師の最終判断や製造販売業者の介入を要することなく、医療現場での使用中、自律的にその性能を変化させていく支援システム等が出現する可能性は否定できない。加えて、その人工知能がネットワークから切り離された状態で搭載され、製造販売業者が関与しない状態で納入先の医療機関ごとに性能が変化する支援システム等の開発も予想される。しかしながら、このような支援システム等がどのように市販後学習を進めるかについては不明な点も多い。

上記のような自律性の高いシステムは、現在の科学技術水準では期待されるレベルの診断支援能力を持つには至っていないが、将来実現した場合には、再学習時に技術者や医師による学習データの作成が不要となり、教師あり学習の場合と比べて支援システム等の自律度は高くなる。また、市販後学習用データの収集や市販後学習のタイミングも全て自ら判断して実行し、技術者や使用者による介入が完全に不要な自律した支援システム等が登場する可能性もある。このような人による介入度が低い、又は不要な支援システム等では特に品質管理が重要となり、承認申請時の性能を担保していることの確認やリスクマネジメントに関する要求は自律度の上昇に伴って高くなる。これらの要求等を充足するためには、前項までに挙げた各種の要求事項をより高度なレベルで満たすことが求められる。充足すべき事項及びそれらのレベルに関しては現時点で評価指標として具体的に記載することは困難であることから、承認申請に先立って独立行政法人医薬品医療機器総合機構と詳細な相談をしておくことが必要になると考えられる。

以上

1 行動変容を伴う医療機器プログラムに関する評価指標（案）

2

3 1. はじめに

4 近年、Internet of Things (IoT) 技術の発展に伴い、スマートフォン等を利用し、個々の患者
5 の行動や思考のパターンに応じて情報を提供することで、従来と異なる習慣づけやアウトカムを
6 もたらす医療機器プログラムの開発が活発に進められている。行動変容の手段としては、心理的
7 アプローチ（認知行動療法ほか）、運動療法や栄養指導等が挙げられる。我が国においては、超高
8 齢化社会に向けた健康寿命の延伸や国内医療機器産業の活性化等が国家的戦略として提唱されて
9 おり、このような行動変容を伴う医療機器プログラムの効果や新規産業としての発展が期待され
10 ている。

11 現在開発中の行動変容を伴う医療機器プログラムとしては、糖尿病、高血圧、非アルコール性
12 脂肪肝炎、アルコール健康障害等の生活習慣病、がん、睡眠障害、認知症等の患者に適用する製
13 品が挙げられる。しかし、心理療法等を用いて行動変容を促す医療機器プログラムは新しい分野
14 の製品であり、その開発においては、表示する情報の内容のみでなく、情報を表示するタイミン
15 グや方法、文化的背景も有効性に影響する可能性が示唆されており¹⁾、有効性・安全性に関わる
16 因子の特定が困難である。また、その評価に際しては、既存治療等の併用療法の有無、比較対照
17 の設定方法、治療効果の発現時期と持続性を見込んだ評価期間、再介入時の有効性、ユーザイン
18 ターフェースにカスタマイズ機能を搭載した場合の振れ幅等を考慮する必要がある。承認申請に
19 においては、有効性・安全性を示す事項を承認申請書に記載するが、上述の理由から記載すべき事
20 項が明確に定まっているわけではない。また、評価結果から導かれる使用目的又は効果の範囲に
21 も注意が必要である。さらに市販後もソフトウェアアップデートや製品改良に伴う変更手続き、
22 患者データ等の取り扱い等にも留意する必要がある。

23 米国 Food and Drug Administration (FDA) は「医療機器ソフトウェア機能及びモバイル医療ア
24 プリケーションに関する指針(2019)」等、各種の関連ガイダンスを発行しており、ドイツでは2019
25 年の法改正により低リスクの医療機器プログラムを対象としたファストトラック制度を制定し、
26 既に20*品目を仮保険償還の対象にする等、海外では医療機器プログラムの早期実用化を支援する
27 行政的施策が講じられつつある。(*2021年9月現在)

28 そこで本評価指標では、当該分野の適切な発展に寄与することを目的として、我が国固有の医
29 療環境を考慮した上で、行動変容を伴う医療機器プログラムに求められる安全性、有効性に関す
30 る評価の留意点を、ユーザビリティ等の因子を考慮しつつ取りまとめた。

31

32 2. 用語の定義

33 本評価指標における用語の定義は、以下のとおりとする。

34

35 (1) 患者等

36 当該製品を治療目的で使用する患者や、自身の疾病予防のために使用する健常者

37

38 (2) 医療者

39 医師、看護師、保健師等

40

41 (3) 一次予防

42 生活習慣を改善して健康を増進し、生活習慣病等の発症を予防すること。発症前の健常者に対
43 して、疾病の原因と考えられる要素の除去や忌避に努め、健康を増進を図って疾病の発生を防ぐ
44 等の予防措置を講じること。

45

46 (4) 二次予防

47 疾病の早期発見、早期治療を行うことにより、疾病の重篤化を防ぐこと。発症した患者を可能
48 な限り早期に発見し、早期治療を行い、疾病の進行を抑え、重篤化しないように努めること。

49

50 (5) 三次予防

51 発症後、必要な治療を受け、機能の維持・回復を図ると共に、再発や合併症の発症を予防する
52 こと。疾病が進行した後、後遺症治療、再発防止、残存機能の回復・維持、リハビリテーショ
53 ン、社会復帰等の対策を講じ、実行すること。

54

55 (6) 心理療法

56 対話等の言語的又は非言語的な交流と介入により、認知や情緒、行動等の心理的側面に変容を
57 もたらし、症状や苦痛を改善・緩和する治療法を指す。精神療法とほぼ同義で用いている。

58

59 (7) 認知行動療法、認知療法、行動療法

60 認知行動療法とは、認知的偏りや非適応的行動を検証し、認知再構成法等の認知的アプローチ、
61 行動活性化や曝露療法等の行動的アプローチによって認知や行動の修正を行い、気分の改善や具
62 体的な問題の解決、対処能力の向上や再獲得を図る治療法を指す。

63 認知行動療法は、歴史的な変遷の下、異なった理論体系を背景にした認知療法と行動療法の要
64 素が融合した治療法である。認知療法とは、認知理論に基づく認知的側面への介入により変容を
65 促す治療法である。行動療法とは、学習理論に基づいた行動の分析や制御の促進によって精神状
66 態を変化させる治療法である。

67

68 (8) 行動変容

69 疾病の予防・治療等を支援するために、個々の患者等の情報を処理し、得られた結果をその患
70 者等に応じた適切な情報として提示する、又はそれを活用した心理療法等により介入し、日常の
71 生活習慣を含めた行動を変化させることで、臨床的に意義のある効果をもたらすこと。

72

73 (9) 主観的評価

74 本評価指標において「主観的評価」とは、被験者の状態に関して被験者自身又は代理者による

75 回答を記すことによる評価、又は被験者の回答に基づく医療者の評価、さらには医療者が被験者
76 を観察した評価である。同じ主観的評価の範疇であるが、医学的な専門性を持たない被験者や代
77 理者が記した評価と、医療者の評価とはその特性が異なることに留意する必要がある。一方、「客
78 観的評価」とは、被験者や評価者の判断の影響を受けない、定量的な検査結果等を用いた評価で
79 ある。

80

81 (10) シャムアプリ

82 一般に盲検性を確保することを目的に用いられ、医薬品でいうプラセボ薬に相当するソフトウ
83 エア。

84

85 3. 本評価指標の対象

86 本評価指標は、「医師の指導の下で使用され、個々の患者等に応じて情報提供することで患者
87 等の行動変容を促す医療機器プログラム」を対象とする。非医療機器であるヘルスケア製品に相
88 当する製品は本評価指標の対象外であるが、一次予防を目的とした製品のうち、医療機器に該当
89 する製品は対象となり得る。一方、二次予防及び三次予防については、疾病の徴候の検出や疾病
90 の進行の抑制、機能の維持等、診断、治療の側面があるため医療機器に該当し、本評価指標の対
91 象となり得るものと考えられる。なお、患者等の行動変容を伴わずに治療効果を発揮することを
92 意図する製品、並びに診断のみに基づき行動変容を伴わず受診勧奨を行うプログラムは本評価指
93 標の対象外とする。

94 その他の機能を有する医療機器プログラムや、行動変容を伴う医療機器プログラムの機能を含
95 む有体物の場合は、プログラムの該当する部分について本評価指標を参照することを推奨する。

96 なお、医療機器プログラムの該当性の判断については、「プログラムの医療機器該当性に関する
97 ガイドラインについて」²⁾を参照の上、厚生労働省医薬・生活衛生局監視指導・麻薬対策課又は
98 医薬品医療機器総合機構審査マネジメント部 SaMD 一元的相談窓口にご相談することが望ましい。

99 その他にも、厚生労働省が公開している「医療機器プログラム事例データベース」も参考となる
100 ³⁾。

101

102 4. 評価指標の位置づけ

103 本評価指標は、技術開発が著しいデジタル技術を対象とすると共に、行動変容を伴う医療機器
104 プログラムを取巻く状況も日々変化していることを勘案し、現時点で重要と考えられる事項を示
105 したものである。今後の技術革新や医療現場での知見の集積等を踏まえて改訂されていくべきも
106 のであり、承認申請内容に対して、拘束力を持つものではない。本評価指標が対象とする製品の
107 評価にあたっては、個別の製品の特性を十分に理解した上で、科学的な合理性を背景にして、柔
108 軟に対応する必要がある。本評価指標の他、国内外のその他の関連ガイドライン等も参考にす
109 ること。

110

111 5. 評価にあたって留意すべき事項

112 (1) 基本的事項

113 1) 設計開発の経緯、品目の仕様、開発機器の原理（アルゴリズムを含む）、対象疾患、使用目的
114 又は効果、類似品の国内外での使用状況、使用場所、使用方法等を明らかにすること。具体例
115 を以下に列挙する。

116 ① 対象疾患、想定される使用者の範囲とその要件

- 117 ・ 対象とする疾患やその重篤度等
- 118 ・ 使用者については、患者等若しくは医療者又はその両者の別
- 119 ・ 使用者の要件及び使用に適さない患者又は医療者

120 ② 使用目的又は効果

- 121 ・ 予防（二次予防、三次予防）、治療等の別

122 ③ 臨床的位置づけ

- 123 ・ 開発コンセプト（当該疾患の治療の中で、本品が介入し解決する課題及びその達成手段）
- 124 ・ 介入内容、介入する治療段階及び介入頻度
- 125 ・ 既存治療法等（薬物療法、心理療法、ガイドラインに基づいた個々の患者に応じた医療的
126 指導、患者教育等のいずれか若しくはこれらの組み合わせ）との併用（上乘せ）、補完・
127 代替（置き換え）、又は新規治療アプローチの別
- 128 ・ 上記項目も考慮した上での現状の診療・治療法と製品との差分

129 ④ 使用方法

- 130 ・ 医療機器プログラムの提供方法
- 131 ・ 患者等に提供されるセンサ等プログラム以外の機器、又は医療者用に提供されるプログラ
132 ム及び機器の有無
- 133 ・ 患者等に提示される情報と、医療者に提示される情報の内容
- 134 ・ 医療者の関与方法
- 135 ・ 操作マニュアル等の文書化とその内容（故障時の対応を含む）

136 ⑤ 原理

137 行動変容を促すシナリオ、指示の決定フローチャートや判断基準等、有効性・安全性に影響
138 する要素や根拠の詳細、根拠とする診療ガイドライン等

139 ⑥ プラットフォームの OS、製造業者により開発されたソフトウェア部品以外の使用するアプリ
140 ケーションソフトウェア及びソフトウェア部品

141 ⑦ 共存するその他のソフトウェアとの干渉の有無

142 ⑧ モバイル機器やセンサ等、併用する機器の要件

143

144 2) 行動変容を伴う医療機器プログラムについては、以下の事項を参考として、有効性・安全性に
145 影響する項目を明示する必要がある。それぞれの項目について、具体的なデータ又はその他の科
146 学的根拠等をもって明らかにすること。

147 ① 提示される心理療法等の根拠

- 148 ・ これまでに有効性について検証されたことのない新規手法が含まれる場合

149 その手法を用いることの適切性について説明すること。何らかの文献報告がある場合には
150 それも提示すること。

151 ・ 標準治療や治療ガイドラインに基づく場合
152 標準治療や治療ガイドラインに基づく場合であってもサロゲートエンドポイントを使用し
153 て評価する際は、その適切性を示すこと。臨床的な位置づけが確立していないパラメータを使用
154 する場合は、その適切性の科学的根拠を明示すること。

155 ・ 検証的臨床試験で有効性が確認されている場合
156 疾患ガイドラインには記載されていないものの、既に当該疾患領域で認知されている方法
157 や、有効性についての検証試験が行われている場合は、その文献情報を提示すること。

158 ② 患者等の行動変容を促す機能の仕様
159 ・ ユーザーインターフェース
160 ・ 出力するメッセージの表現
161 ・ 提示する情報の選択や表示のアルゴリズム

162 ③ 開発時に当該製品の機能設計・性能評価・検証等に用いられたデータ
163 開発時に当該製品の機能設計・性能評価・検証等に用いられたデータについて説明すること。
164 製品の目的に合致しない、偏った対象者、網羅性に欠ける方法で収集されたデータ等を用いた場
165 合は、汎化性に欠けるプログラムとなる可能性があることに留意すること。
166

167 (2) 非臨床試験に関する事項
168 「行動変容を伴う医療機器プログラム」の安全性等について、以下の事項を参考として評価す
169 ること。なお、本評価指標では、製品に含まれるセンサ等の有体物については言及しないが、認
170 証基準等も参考に、別途適切に評価する必要がある。

171 1) プログラム部分の評価
172 意図したとおりにプログラムが動作することを評価すること。

173 2) 安全性に関する評価
174 開発機器の特性を踏まえて、適切に評価すること。

175 ① 品質マネジメント、リスクマネジメント⁴⁾
176 ソフトウェアライフサイクルプロセス^{5, 6)}について評価すること。

177 ② サイバーセキュリティ⁷⁾⁸⁾
178 開発機器の特性に応じて、想定されるサイバーリスクを明確化し、最新のセキュリティ基準・
179 規格に応じた適切な対策が講じられていることが望ましい。

180 ③ 個人情報保護⁹⁾
181 臨床情報を取得する機器にあつては、その取得する範囲と目的を特定し、必要に応じ患者等の
182 同意を得ること。また、必要な個人情報保護の対策が講じられていること。臨床情報を保存する
183 機能を有する機器にあつては、情報の保管や廃棄の際の取り扱いについても考慮されていること
184 が望ましい。

185 ④ 情報セキュリティ¹⁰⁾

186 患者等のデータを保管するためにサーバーやクラウド等を利用する場合は、最新の技術に照ら
187 し合わせ、適切な対策がなされていることを確認すること。また、情報取扱についての社内ルー
188 ルを規定し徹底すること。また、市販後の脅威の監視と対応体制を明記されていることが望まし
189 い。

190

191 (3) 臨床評価に関する事項

192 1) 臨床試験（治験）の必要性の考え方

193 臨床評価に関する資料の必要性については、個々の医療機器プログラムの特性、非臨床で評価
194 できる範囲等を基に総合的に判断されることから、医薬品医療機器総合機構の対面助言を活用す
195 ることが望ましい。¹¹⁻¹³⁾

196 現時点において、行動変容を伴う医療機器プログラムを対象とした承認基準は存在しない。ま
197 た、行動変容を伴う医療機器プログラムについては、承認前例が限られると共に、その有効性・
198 安全性を非臨床試験のみで評価することが難しいため、臨床試験成績に関する資料の必要性が高
199 いと考えられる。

200

201 2) 臨床試験（治験）のデザイン

202 臨床試験の実施にあたっては、5/(1)/2)/③項「臨床的位置づけ」に基づき、まずは二重盲検ラ
203 ンダム化比較試験の実施要否について検討すること。また、優越性又は非劣性の検証を選択した
204 理由の適切性について確認すること。

205 対照群の設定にあたっては標準治療、既存の行動変容を伴う医療機器プログラム、シyamアプ
206 リ等の要否を適切に検討し、臨床的有効性のエビデンスを構築すること。また、効果の持続性
207 について評価する必要がある場合は適切に観察期間を設けること。

208 評価項目の設定については、可能な限り広く認知された標準的な客観的指標を用いること。対
209 象とする疾患等、行動変容を伴う医療機器プログラムの臨床的位置づけによっては、主観的指標
210 ^{14, 15)}を用いざるを得ない場合があるが、その際は評価すべき内容に応じて可能な限り信頼性・妥
211 当性が検証されている、又は当該診療領域において標準的に広く受け入れられている適切な指標
212 を選択すること。選択した指標を用いて評価する際は、治験全体のデータの質を向上させるため
213 に適切なトレーニングや回答方法の説明等を実施し意図したデータを収集すると共に、得られた
214 成績の正確性、再現性、妥当性等について注意深く検討する必要がある。可能であれば、副次評
215 価項目等として客観的な指標を用いて評価することが望ましい。

216 試験デザインについては、個々の医療機器プログラムの特性を十分に検討した上で、症例数、
217 実施期間、対照群、及び行動変容のアウトカムの特性を考慮した有効性の指標となる臨床的意義
218 がある変化量等を設定すること。

219 個々の医療機器プログラムの臨床試験デザインについては、以下の項目について考慮すること。

220 ① シyamアプリ

221 シyamアプリを使用する場合には、パイロットスタディー等でシyamアプリの盲検性が適切に
222 確保されていることを確認すること。その際、プラセボ効果の影響や群間差についても考慮する

223 こと。シムアップの作成には困難を伴うことが多いと考えられるため、PMDA との相談を活用し
224 て試験プロトコルについて検討することを推奨する。

225 ② プラセボ効果

226 二重盲検ランダム化比較試験の実施が困難な場合、可能な限り、プラセボ効果の影響を受け難
227 い評価項目を選択すること。プラセボ効果が排除できない評価項目を使用する場合は、臨床的位
228 置づけに立ち戻り、副次評価項目も含めた試験デザイン全体として評価することも考えられるが、
229 具体的内容については PMDA の相談を活用することを推奨する。

230

231 3) 臨床成績に影響する事項

232 ① 人種や文化的背景の影響

233 海外における使用実績や、臨床試験成績がある製品であっても、行動変容を伴う医療機器プロ
234 グラムにおいては、人種差のみならず、宗教、道徳観、生活環境等の文化的背景が有効性に影響
235 し得ることを考慮し、必要に応じて国内において臨床試験を実施すること。また、世代の違いに
236 よる影響や、地域性の影響についても評価することが望ましい。

237 ② 開発時期や臨床試験が実施された年代の影響

238 評価された時代背景が有効性・安全性へ与える影響について評価すること。例えば、10 年前に
239 開発された製品を承認申請する場合や、疾患ガイドライン等が改訂された場合は、性能に及ぼす
240 影響の有無、当該機器の臨床的位置づけについて再評価する必要がある。

241 ③ 患者アドヒアランスへの影響

242 行動変容を伴う医療機器プログラムは、継続的に使用することで効果を発揮するものもあると
243 考えられることから、使用継続率に影響し得る要素も、当該機器の臨床成績に影響する可能性が
244 あることに留意すること。例えば、以下のような項目が挙げられる。

- 245 ・ フォントや背景色等、グラフィカルなユーザインターフェース
- 246 ・ 出力するメッセージの表現
- 247 ・ 方言やキャラクター等、嗜好に依存する事項

248 ④ ユーザインターフェースのカスタマイゼーション

249 患者等がカスタマイズ等できる仕様を含む場合はその範囲を明らかにすると共に、その影響に
250 ついて評価すること。

251 ⑤ 第三者との比較や交信の取扱い

252 医療機器プログラム上で患者等が他の患者等と交信したり、達成状況等を比較したりする機能
253 を含む場合は、その有効性だけでなく、安全性や適切性についても検討すること。ピアサポート
254 は、属する集団によって有効性・安全性に差が生じ得ることも留意すること。

255 ⑥ 既存治療終了後の延長効果と開発機器の効果の区別

256 当該医療機器プログラムの臨床的位置づけによっては、既存療法を完了、中止等した際の延長
257 効果を考慮した上で、標榜する有効性が適切に評価されていることを示すこと。

258

259 4) 安全性に関わる留意事項

260 臨床上許容できないハザードが存在し得る製品に関しては、確実なアウトカムとリスク評価が
261 必要となる。適応対象となる患者等に対し、不適切な介入（例えば、高齢者に対する過度の運
262 動、食事制限がある患者等に対する不適切な食事指導等）によるリスクについて検討し、製品の
263 仕様や注意喚起の方法等も踏まえたリスクアセスメントを実施することが重要である。例えば、
264 うつ病や糖尿病では不適切な情報介入が行われると、自殺企図の悪化や低血糖発作の惹起といっ
265 た生命に関わる事態を招く可能性がある。このように介入によってもたらされ得る効果が重大な
266 生命リスクに直結する疾患もあるため、対象とする疾病だけでなく関連する合併症等を含めて検
267 討すること。

268

269 6. 市販後に留意すべき事項

270 本評価指標は、承認審査に係る留意事項を取りまとめるものであるが、承認後に予見される医
271 療機器プログラム特有の課題について、付言する。

272 1) 一部変更承認申請・軽微変更届出の取扱い

273 医療機器プログラムでは、プラットフォームのOSやセンサを含めた併用機器の更新又は改良が
274 必要となる場合がある。変更手続きについては、平成29年7月31日付け薬生機審発0731第5号
275 厚生労働省医薬・生活衛生局医療機器審査管理課長通知「医療機器の一部変更に伴う軽微変更手
276 続き等の取扱いについて」¹⁶⁾を参照の上、適切に対応すること（別添参照）。

277

278 2) 一部変更承認申請の必要性の判断

279 有効性・安全性に影響する変更がある場合は一部変更承認申請を行う必要がある。有効性・安
280 全性に影響する事項は製品毎に異なることから、開発段階から何がその医療機器プログラムの有
281 効性を規定する因子かについて検討、把握しておく必要がある。判断に悩む場合は、適宜PMDAに
282 相談することを推奨する。

283 なお、5/(1)/3)/①項「提示される心理療法等の根拠」として用いた診療ガイドライン等に変更
284 があった場合には、当該医療機器プログラムの有効性・安全性への影響を検討して判断すること。

285

286 3) 変更計画確認手続き制度（Improvement Design within Approval for Timely Evaluation 287 and Notice, IDATEN）について

288 市販後に当該機器の性能向上が可能となる場合も考え得る。連続的又は高頻度の改良を計画す
289 る場合は、当該制度¹⁷⁻¹⁸⁾の利用についても検討する選択肢がある。

290

291 4) 臨床評価が必要な変更の範囲

292 診療ガイドライン等の更新に伴い新たな臨床評価を必要とする変更を行う場合は、5/(3)/1)項
293 「臨床試験（治験）の必要性の考え方」に準じて臨床試験を実施すること。変更目的や変更内容
294 に応じて、必要なデータパッケージは初回申請時と異なることもある。

295

296 参考資料

- 297 1) 令和2年度 次世代医療機器・再生医療等製品評価指標作成事業 行動変容を伴う医療機器プ
298 ログラムに関する調査研究班報告書 令和3年3月
- 299 2) 令和3年3月31日付け薬生機審発0331第1号、薬生監麻発0331第15号「プログラムの医
300 療機器該当性に関するガイドラインについて」
- 301 3) 厚生労働省ホームページ「医療機器プログラムについて」
302 https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000179749_00004.html
- 303 4) ISO 14971:2019 Medical devices — Application of risk management to medical
304 devices / JIS T 14971:2020 医療機器-リスクマネジメントの医療機器への適用
- 305 5) IEC 62304:2006/Amd 1:2015 Medical device software — Software life cycle processes
- 306 6) JIS T 2304:2017 医療機器ソフトウェアソフトウェアライフサイクルプロセス
- 307 7) 令和3年12月24日付け薬生機審発1224第1号 厚生労働省医薬・生活衛生局医療機器審査
308 管理課長通知、薬生安発1224第1号 厚生労働省医薬・生活衛生局医薬安全対策課長通知
309 「医療機器のサイバーセキュリティの確保及び徹底に係る手引書について」
- 310 8) 令和2年5月13日付け薬生機審発0513第1号 厚生労働省医薬・生活衛生局医療機器審査
311 管理課長通知、薬生安発0513第1号 厚生労働省医薬・生活衛生局医薬安全対策課長通知
312 「国際医療機器規制当局フォーラム(IMDRF)による医療機器サイバーセキュリティの原則及
313 び実践に関するガイダンスの公表について(周知依頼)」
- 314 9) 「個人情報保護に関する法律についてのガイドライン(通則編)」第一版、平成28年11
315 月(令和3年10月一部改正) 個人情報保護委員会
- 316 10) 「医療情報を取り扱う情報システム・サービスの提供事業者における安全管理ガイドライ
317 ン」令和2年8月 総務省・経済産業省
- 318 11) 平成17年2月16日付け薬食機発第0216001号 厚生労働省医薬食品局審査管理課医療機器
319 審査管理室長通知「医療機器の製造販売承認申請に際し留意すべき事項について」
- 320 12) 平成20年8月4日付け薬食機発第0804001号 厚生労働省医薬食品局審査管理課医療機器審
321 査管理室長通知「医療機器に関する臨床試験データの必要な範囲等について」
- 322 13) 平成29年11月17日付け薬生機審発1117第1号 厚生労働省医薬・生活衛生局医療機器審
323 査管理課長通知、薬生安発1117第1号 厚生労働省医薬・生活衛生局医薬安全対策課長通知
324 「医療機器の「臨床試験の試験成績に関する資料」の提出が必要な範囲等に係る取扱い(市
325 販前・市販後を通じた取組みを踏まえた対応)について」
- 326 14) 治験における Patient Reported Outcomes ～臨床開発担当者のためのPRO利用の手引き～
327 2016年6月 日本製薬工業協会 データサイエンス部会 タスクフォース7
- 328 15) Guidance for Industry Patient-reported outcome measures: use in medical product
329 development to support labeling claims, 2009年12月 FDA
- 330 16) 平成29年7月31日付け薬生機審発0731第5号 厚生労働省医薬・生活衛生局医療機器審査
331 管理課長通知「医療機器の一部変更に伴う軽微変更手続き等の取扱いについて」

- 332 17) 令和2年8月31日付け薬生機審発0831第14号 厚生労働省医薬・生活衛生局医療機器審査
333 管理課長通知「医療機器の変更計画の確認申請の取扱いについて」
- 334 18) 令和2年10月30日付け事務連絡「医療機器の変更計画の確認申請に関する質疑応答集
335 (Q&A) について」
- 336

337 市販後変更手続き及び変更計画確認手続き制度（IDATEN）の留意点

338

339 変更計画確認手続き制度（IDATEN）¹⁴⁻¹⁵⁾を利用する場合や市販後に一部変更申請・軽微変更を行
340 う際は、以下の事項を参考として、必要な情報を記載すると共に、適切な評価を行うこと。

341 ① 市販後の収集データに基づいた改良

342 市販後に収集したデータを解析し、当該機器の性能変更を用いる場合には、改良モデルの開発
343 時に使用するデータの品質と同様に、収集されたデータの適切性、過学習の可能性を考慮するこ
344 と。

345 なお、改良時に使用するデータについては、利用目的に応じて患者等の同意が必要となる場合
346 がある。特に市販後のデータ収集においては、当該機器に患者自身が入力した情報や自動取得さ
347 れたバイタルデータ等を収集して利用するため、患者個人の情報を含むデータの利用となること
348 が想定される。したがって、同意の要否については、個人情報保護法の取扱いを十分に確認の上、
349 対応すること。

350 ② OS のアップデート及び併用機器の更新に伴う変更

351 プラットフォームの OS やセンサを含めた併用機器の更新又は改良が開発機器の有効性・安全性
352 に及ぼす影響について評価すること。なお、それらの更新に伴って当該機器が正常に作動するため
353 に行った変更は一部変更承認申請を要しない。

354 ③ 疾患ガイドラインの更新に伴う変更

355 標準治療や併用治療等が更新される可能性を踏まえて、当該プログラムの位置づけ、標準療法
356 との関係性について明確化しておき、ガイドラインの変更に伴う当該プログラムの有効性・安全
357 性への影響を検討し、必要に応じてプログラムの改修を行うこと。この改修が臨床アウトカムに
358 影響を及ぼす場合は一部変更承認申請をすること。

359 ④ 有効性に影響するアルゴリズム自体の変更

360 新旧アルゴリズムの相違点を明示した上で、アルゴリズムの改良が当該機器の有効性・安全性
361 に及ぼす影響を明らかにすること。医療機器プログラムにおいては、プログラムの全容を明示す
362 ることは難しいこともあるが、有効性・安全性に影響する要素が明確化されていることが望まし
363 い。

令和3年度次世代医療機器・再生医療等製品評価指標検討会（厚生労働省）／
医療機器開発ガイドライン評価検討委員会（経済産業省）合同検討会

行動変容を伴う医療機器プログラム審査ワーキンググループ 委員名簿

座 長：佐久間一郎 東京大学大学院工学系研究科附属医療福祉工学開発評価研究センター
バイオエンジニアリング専攻 教授

委 員（五十音順）：

菊地俊暁	慶應義塾大学医学部 精神・神経科学教室 専任講師
後藤 励	慶應義塾大学大学院 経営管理研究科／健康マネジメント研究科 教授
鈴木孝司	公益財団法人医療機器センター 認証事業部 審査役
田村雄一	国際医療福祉大学三田病院 心臓血管センター（肺高血圧症センター） 国際医療福祉大学医学部 循環器内科学 教授