

Digital Therapeutics の審査ポイント (Draft)

2026年6月22日
プログラム医療機器審査部

目次

1. はじめに	1
2. 定義	1
3. 対象となる製品	2
4. 臨床試験(治験)の要否について	2
5. 臨床試験(治験)デザインに関する基本的な考え方について	3
6. 対照群について	4
7. 併用療法について	5
8. 評価項目について	5
9. まとめ	6

1. はじめに

医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律(以下「医薬品医療機器等法」という。)第二十三条の二の五 第 1 項では、「医療機器の製造販売をしようとする者は、品目ごとにその製造販売についての厚生労働大臣の承認を受けなければならない。」とされており、同条の第 2 項に該当する場合は承認を与えないとされている。プログラム医療機器を含む医療機器の製造販売承認審査では、医薬品医療機器等法第二十三条の二の五 第 2 項 第三号のうち、主にイ及びロの各要件に該当しないことを確認している。

第二十三条の二の五 第 2 項 第三号

- イ 申請に係る医療機器又は体外診断用医薬品が、その申請に係る効果又は性能を有すると認められないとき。
- ロ 申請に係る医療機器が、その効果又は性能に比して著しく有害な作用を有することにより、医療機器として使用価値がないと認められるとき。
- ハ イ又はロに掲げる場合のほか、医療機器又は体外診断用医薬品として不適当なものとして厚生労働省令で定める場合に該当するとき。

製造販売業者は、開発品の品質、有効性及び安全性(以下「有効性等」という。)を、開発過程で実施される非臨床試験(プログラム医療機器においては、主に性能評価試験)や臨床試験等において評価することになる。また、疾病治療用プログラム医療機器に関しては、その有効性及び安全性を非臨床試験のみで評価することが難しいことから、想定される使用者における有効性及び安全性が確保されていることを臨床試験(治験)により確認することが必要となる。

プログラム医療機器は心理療法等を開発コンセプトにしていることが多く、プログラム医療機器の有効性及び安全性を評価するために実施する治験を設計する場合は、プラセボ効果、ホーンソン効果等の心理的な影響による効果に配慮することが必要になる。医薬品の治験では、心理的な影響による効果を踏まえて、プラセボ薬を用いた二重盲検比較試験を実施することが考慮されるが、後述するように、プログラム医療機器においてプラセボ薬と同様のシャムアプリの作成方法は確立されていない。

本審査ポイントでは、PMDA で実施している対面助言や製造販売承認審査における経験を踏まえ、プログラム医療機器における対照群を設定する際の考え方や留意点等について示している。なお、本審査ポイントに示されている留意事項等が唯一のものではなく、製品の特性を踏まえて個別に検討することが必要となるため、必要に応じて PMDA の対面助言を利用することを推奨する。

2. 定義

- プログラム医療機器: 医療機器のうちプログラムであるもの(医療機器プログラム)、又はプログラムを記録した記録媒体も含むもの。
- Digital Therapeutics (DTx): 科学的な根拠に基づいた治療効果を有する医学的介入であり、

疾患、障害、状態、又は傷害を治療または軽減することを意図したヘルス・ソフトウェア (Digital Therapeutics Alliance のウェブサイト¹より引用し、PMDA により翻訳を実施)。日本においては、疾病治療用プログラム医療機器の一部が該当する。

- 開発品: 製造販売承認を目指す疾病治療用プログラム医療機器。
- 臨床評価: 製造業者の意図に従い使用したとき、医療機器の臨床的安全性及び性能を検証するための、医療機器に関する臨床データの評価及び分析 (ISO13485 より引用)。
- シャムアプリ: 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ェア (「次世代医療機器評価指標の公表について」(令和 4 年 6 月 9 日付け薬生機審発 0609 第 1 号)の別添 2「行動変容を伴う医療機器プログラムに関する評価指標」(以下「行動変容評価指標」という。)より引用)。
- プラセボ効果: 製品を使用していると考えることによって被験者に改善が見られること (「臨床試験における対照群の選択とそれに関連する諸問題」について) (平成 13 年 2 月 27 日付医薬審発第 136 号より一部引用)。
- ホーソン効果: 人は一般に注目されることを好み、特別な扱いを受けると、さらに効果を上げようとする傾向があること (デジタル大辞泉より引用)。

3. 対象となる製品

本審査ポイントでは、DTx 製品とされる疾病治療用プログラム医療機器を対象とする。本審査ポイントで示す考え方は、一律にすべての製品や領域に適用できるものではなく、個別製品ごとに議論が必要になることに注意が必要である。また、プログラム医療機器に該当しないソフトウェア、製品構成の一部に有体物を含むもの、ソフトウェアの使用に際し、医療機器たる用途に限定された専用のハードウェアが必要なもの、医薬品医療機器等法における医療機器の定義に該当しないものは、本審査ポイントの対象ではないことにも留意する必要がある。

4. 臨床試験(治験)の要否

疾病治療用プログラム医療機器に属する DTx 製品の製造販売承認申請に際して評価すべき、ハイレベルな内容は、以下のとおり整理される。

<評価すべき内容>

- ① 意図した使用者において、臨床的に意義がある程度に患者転帰等を改善できること。
- ② 意図した期間、使用が継続できること。
- ③ 本品の機能が意図とおりに動作すること。
- ④ ソフトウェア開発ライフサイクルが適切に管理されていること。
- ⑤ サイバーセキュリティへの対応が適切になされていること。

¹ <https://dtxalliance.org/understanding-dtx/what-is-a-dtx/>

⑥ ユーザビリティエンジニアリングプロセスが適切に管理されていること。

製造販売承認申請では、プログラム医療機器の開発コンセプト、臨床的位置づけ、既承認の類似医療機器との差分、概念的な要求事項を検討し、性能試験、非臨床試験、既存臨床成績、開発品に関する学術論文等を踏まえ、総合的に臨床評価を実施する。製造販売承認申請を検討する際、治験でなければ評価できない論点が未評価のまま残る場合には、新たな治験の実施が必要となる。一般的に、疾病治療用プログラム医療機器のうち DTx 製品の有効性(開発品を使用したことによる効果)及び安全性を非臨床試験のみで評価することは困難であるため、原則、製造販売承認申請に際しては、検証的治験の試験成績の提出が必要になる。

製造販売承認申請に際して提出する臨床試験成績に関する資料は、原則として治験によって得られた試験成績であることが必要である。一方で、特定臨床研究で得られた情報の薬事利活用の是非については、得られた試験成績によるプログラム医療機器の評価の可否、根拠資料の保存状況、治験の実施・再現の困難度等も踏まえて個別に判断する必要がある。特定臨床研究で得られた試験成績の承認申請への利用を検討する場合は、早期に PMDA に相談すること。

5. 臨床試験(治験)デザインに関する基本的な考え方

すべての DTx 製品に適用可能な治験デザインに関する基本的な考え方が整理されたガイドライン等は作成されていないが、「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等の行動変容を促す医療機器プログラム」を対象とした行動変容評価指標が作成され、厚生労働省より通知されている。一方で、行動変容評価指標における行動変容とは、「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者等に応じた適切な情報として提示するなど、患者由来の情報を活用した心理療法等により介入し、日常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。」とされており、行動変容評価指標は、認知行動療法等を含む心理療法を開発コンセプトとするプログラム医療機器を想定して作成されている。開発品によっては、行動変容評価指標を参考にすることが適切ではない場合もあるが、多くの DTx 製品の開発においては、行動変容評価指標は参考にすべき指標である。

行動変容評価指標では、治験デザインについて以下のような考え方が示されており、シャムアプリ等を用いた二重盲検ランダム化比較試験を検討する必要性が言及されている。

- ・ 二重盲検ランダム化比較試験の実施要否について検討すること。
- ・ 二重盲検ランダム化比較試験における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、シャムアプリ等の要否を適切に検討し、臨床的有効性の二重盲検ランダム化比較試験エビデンスを構築すること。

ICH E9(臨床試験のための統計的原則)や中枢神経系用薬及び末梢神経系用薬等の医薬品の臨床評価に関するガイドラインにおいても、参考表 1 のように同様の考え方が示されており、行

動変容評価指標で示されている考え方は、プログラム医療機器の治験デザインのみで考慮すべき事項ではなく、治験を設計する際に考慮すべき事項である。一般的に、測定値等の客観的な評価指標ではない、患者本人による自記式の評価指標等の主観的な評価指標を主要評価項目とする場合、プラセボ効果、ホーソン効果等の心理的な影響による効果が含まれるため、製品としての有効性を評価するためには、原則として対照群にシャムアプリ等を設定した二重盲検ランダム化比較試験を行う必要がある。

6. 対照群

行動変容評価指標では、シャムアプリについて以下のような考え方が示されている。

- ・ シャム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ャムアプリの盲検性が適切に評価されていることが理想である。その結果等から、二重盲検ランダム化比較試験の実施可能性やシャムアプリの効果量の影響等についても考慮すること。
- ・ 盲検性を保ちつつ効果のないシャムアプリを作成することには困難を伴うことが考えられ、二重盲検ランダム化比較試験の実施自体が難しいことも想定される。この場合は可能な限り、ホーソン効果等といった患者の心理的な影響を受け難い評価方法及び評価項目を選択すること。主観的指標を使用する場合は、臨床的位置づけに立ち戻り、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。

盲検性を保ちつつプラセボ薬のように効果をもたないシャムアプリの作成方法は確立されておらず、多くの場合は、開発品において有効性を発揮すると考える一部の機能を排除した製品をシャムアプリとすることが多いと推測する。一方で、一部の機能やコンテンツが共通することで、一定の有効性が発揮する可能性もあると考える。有効性を発揮する可能性のあるアプリは真のシャムアプリとはみなせないため、以降はこういったアプリは、対照群に提供されるアプリとして、「コントロールアプリ」と定義する。

多くの開発者は、有効性が発揮する可能性のあるコントロールアプリを治験の対照群に用いることに抵抗感がある場合も考えられる。しかしながら、被験者には、①疾病治療用プログラム医療機器の治験に参加することで経験できるアプリ操作を伴う新しい介入に対する期待及び②デジタルツールで管理されることに対するホーソン効果が存在すること等を考慮すると、対照群に対して標準治療又は通常治療のみを提供するだけでは、①及び②のような心理的な影響を評価することができず、治験における対照群として妥当ではない可能性がある。

以上を踏まえると、可能な限り盲検性を維持しつつ、より適切な有効性評価を実施するため、治験における対照群には、コントロールアプリを提供することを検討する必要があると考える。結果として、盲検性の確保に疑義が生じたとしても、一定の有効性が発揮する可能性のあるコントロールアプリが提供される対照群の成績は、標準治療又は通常治療のみが提供される対照群よりも高い成績を示すことになると想定され、コントロールアプリが提供される対照群はより保守的な対照群と言える。この対照群に対する優越性を示すことができるのであれば、開発品の臨床的有用性を説

明できる可能性があると考ええる。

なお、コントロールアプリに、どのような機能を搭載することが適切であるか判断に迷う場合、前述のとおり、開発者が開発品において有効性を発揮すると考える機能を排除することでおおむね問題ない。ただし、機能を排除しすぎること、被験者に提供された段階からコントロールアプリであることが明らかになってしまう問題が生じる可能性が否定できないため、保守的な対照群とするためにも、例えば多くの心理療法で取り入れられているセルフモニタリングに関する機能（例えば、日記機能等）は残すことが考えられる。

7. 併用療法

人体に対して直接的な侵襲性がない DTx 製品は、治療アルゴリズムの中で最初に処方される選択肢となる可能性があり、開発者によっては、患者に対して診療の最初に処方される位置づけ（ファーストライン）を想定して開発を進める場合もある。一方で、DTx 製品の作用機序や効果が発現するまでに必要な期間等が医薬品等とは異なるため、多くの場合は、DTx 製品を単独使用したり、DTx 製品のみで治療を完結させたりするのではなく、DTx 製品以外の医薬品等と併用等しながら、患者にとって最適となるように調整し治療を進める可能性が想定される。こういった状況を踏まえると、DTx 製品の治験における開発品群及び対照群は、標準治療又は通常治療に加えて開発品又はコントロールアプリを上乗せするような群構成とすることが、市販後の使用状況を想定した評価につながると考える。

また、開発者によっては、コントロールアプリの必要性に疑問を持ち、開発品群を開発品＋標準治療又は通常治療とし、対照群にはコントロールアプリを用いず、標準治療又は通常治療のみとする群構成を検討する場合も想定される。この場合、プラセボ効果、ホーソン効果等の心理的な影響が、開発品群のみならず、対照群に対してもどの程度影響しているかを定量的に説明することが必要になる。一般的に心理的な影響を定量的に説明することは難しく、開発品そのものの有効性を正確に評価することは困難であり、開発品を使用することに伴う臨床的な有用性を評価又は説明することが必要になる。つまり、開発品そのものの有効性を正確に評価できなかったとしても、開発品が臨床現場で使用される際の状況において、開発品がどのような臨床的有用性をもたらすかを評価又は説明する必要がある。しかし、開発品と医薬品等との併用等が想定されているにも関わらず、併用等に関する有効性及び安全性に関する試験成績が存在しない場合、開発品の臨床的有用性を評価又は説明することが難しくなる。

評価項目行動変容評価指標では、評価項目について以下のような考え方が示されている。

- ・ 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を選択すること。選択した主観的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ともに、得られた成績の正確性、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。

- ・ 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、実施期間、対照群及び行動変容のアウトカムを考慮した有効性の指標となる臨床的意義がある変化量等を設定すること。

得られた治験成績を踏まえて、標準治療又は通常治療と開発品との臨床的位置づけ等を整理するためにも、行動変容評価指標にて示されているように、開発領域の治験において用いられている標準的な評価項目・評価指標を用いることが推奨される。検証的治験において、標準的な評価項目・評価指標以外の評価項目・評価指標を主要評価項目としたい場合には、探索的治験において、少なくとも検証的治験で主要評価項目に設定しようとしている評価項目・評価指標に関する以下のような情報を収集・整理し、その結果を踏まえて検証的治験における主要評価項目を決定することが必要と考える。

- ・ 開発品の有効性を感度高く検出できる評価項目・評価指標であるか。
- ・ 既存の標準的な評価項目との基準関連妥当性があるか。
- ・ どの程度変化すれば、またどの程度の群間差を有すれば、臨床的に有用であるといえるか。

8. まとめ

本審査ポイントでは、PMDA で実施している対面助言や製造販売承認審査における経験を踏まえ、DTx 製品に関する臨床試験(治験)デザインに関する基本的な考え方を示した。一方で、DTx 製品の臨床的位置づけについては、開発する領域等により、様々な考え方がある。本審査ポイントは、PMDA における経験を踏まえて整理された情報であり、必ずしもすべての開発事例に当てはまる考え方ではない。本審査ポイントの考え方を十分に理解し、多くの DTx 製品が早期に臨床現場で活用されることを期待する。

以上

参考表 1 医薬品の臨床評価に関するガイドラインにおける臨床試験デザインに関する基本的考え方

No.	ガイドライン名	臨床試験デザインに関する基本的考え方	通知名
1	臨床試験のための統計的原則	<u>臨床試験で偏りを回避するための最も重要な計画上の技法は、盲検化及びランダム化(無作為化)であり、これらは承認申請に利用することを目的とするほとんどの比較臨床試験で標準的に採用すべきである。(略)</u>	「臨床試験のための統計的原則」について(平成10年11月30日付医薬審第1047号)
2	神経障害性疼痛治療薬の臨床評価に関するガイドライン	薬剤開発では、よく計画され適切に実施された臨床試験により、有効性を検証し安全性が示され、有効用量とその用量範囲を明確にする必要がある。神経障害性疼痛の治療薬の開発においては、 <u>プラセボ群を設定しない実薬対照非劣性試験で有効性を評価することには限界があり、プラセボ対照無作為化二重盲検並行群間比較試験を実施し、有効性及び安全性を検討することが必要</u> である。これは、疼痛の治療薬の効果を検討する臨床試験では、 <u>プラセボに対する反応性が高く、また一定しないことから、実薬対照非劣性試験においては、無効同等の可能性が排除できないため</u> である。(略)	「神経障害性疼痛治療薬の臨床評価に関するガイドラインについて」(令和2年12月28日付薬生薬審発1228第1号)
3	睡眠薬の臨床評価方法に関するガイドライン	薬剤開発においては、臨床試験により治験薬の有効性及び安全性を検討し、治験薬の有効用量とその用量範囲を明確にする必要がある。睡眠薬の開発においては、 <u>実薬対照非劣性試験(又は同等性試験)では無効同等の可能性が排除できないことから、本試験デザインを用いて治験薬の有効性を検証することには限界があり、プラセボ対照無作為化二重盲検並行群間比較試験を実施し、有効性及び安全性を検討することが必要</u> である。(略)	「睡眠薬の臨床評価方法に関するガイドライン」について(平成23年12月13日付薬食審査発1213第1号)
4	抗うつ薬の臨床評価方法に関するガイドライン	薬剤開発においては、臨床試験により薬剤の有効用量とその用量範囲を明確にする必要がある。抗うつ薬の開発においては、 <u>プラセボを含まない又は治験薬の複数用量が用いられない実薬対照非劣性試験(又は同等性試験)により治験薬の有効用量を検討し有効性を検証することには限界があり、抗うつ薬の有効性の検討には、プラセボ対照二重盲検比較試験が必要</u> である。これは、うつ病患者を対象とした臨床試験では、 <u>プラセボに対する反応性が高</u>	「抗うつ薬の臨床評価方法に関するガイドライン」について(平成22年11月16日付薬食審査発1116第1号)

No.	ガイドライン名	臨床試験デザインに関する基本的考え方	通知名
		<u>く、また一定しないことから、内部妥当性を示す指標が存在しない実薬対照非劣性試験(又は同等性試験)においては、無効同等の可能性が排除できないため</u> である。(略)	