

Labeling of Direct-to-Consumer Medical Artificial Intelligence Applications for “Self-Diagnosis”

Sara Gerke

I INTRODUCTION

Artificial intelligence (AI), particularly its subcategory, machine learning, is changing our daily lives and the way we receive health care. The digital health apps market is booming, with over 350,000 health apps available to patients and consumers, ranging from wellness and fitness apps to disease management apps.¹ In particular, many direct-to-consumer medical AI apps for “self-diagnosis” (DTC medical self-diagnosing AI apps) are emerging that help individuals to identify a disease or other condition based on entering, for example, symptoms.² DTC medical self-diagnosing AI apps offer new opportunities, but they also raise issues. While the current legal debate has mainly focused on the poor accuracy of DTC medical self-diagnosing apps,³ this chapter will discuss the labeling challenges associated with these apps that have received little attention in the literature.

This chapter will first explore the current and future landscape of DTC medical self-diagnosing AI apps. It will then focus on their regulation and discuss whether DTC medical self-diagnosing AI apps are medical devices under section 201(h)(1) of the Federal Food, Drug, and Cosmetic Act (FDCA). This will be followed by a discussion of two labeling challenges raised by DTC medical self-diagnosing AI apps:

- ¹ Emily Olsen, Digital Health Apps Balloon to More Than 350,000 Available on the Market, According to IQVIA Report, *Mobi Health News* (August 4, 2021), www.mobihealthnews.com/news/digital-health-apps-balloon-more-350000-available-market-according-ivqvia-report.
- ² The term “consumer” is here understood broadly and includes healthy individuals and patients. Aleksandar Ćirković et al., Evaluation of Four Artificial Intelligence–Assisted Self-Diagnosis Apps on Three Diagnoses: Two-Year Follow-Up Study, 22 *J. Med. Internet Res.* e18097 (2020).
- ³ See, for example, Boris Babic et al., Direct-to-Consumer Medical Machine Learning and Artificial Intelligence Applications, 366 *Nature Mach. Intel.* 283 (2021); Sara Gerke et al., Germany’s Digital Health Reforms in the COVID-19 Era: Lessons and Opportunities for Other Countries, 3 npj *Digit. Med.*, 94 (2020); Stephanie Aboueid et al., The Use of Artificially Intelligent Self-Diagnosing Digital Platforms by the General Public: Scoping Review, 7 *JMIR Med. Info.* e13445 (2019). For privacy aspects of DTC AI/machine learning health apps, see Sara Gerke & Delaram Rezaeikhonakdar, Privacy Aspects of Direct-to-Consumer Artificial Intelligence/Machine Learning Health Apps, 6 *Intelligence-Based Med.* 100061 (2022).

First, the concern of labeling DTC medical self-diagnosing AI apps as what I call “information-only” tools, and second, particular issues associated with the use of AI, ranging from bias to adaptive algorithms.

This chapter concludes that the labeling of DTC medical self-diagnosing AI apps as “information-only” rather than “diagnostic” tools is unknown to most consumers. The Food and Drug Administration (FDA) should create user-friendly labeling standards for AI-based medical devices, including those that are DTC. For example, these standards should ensure that consumers are adequately informed about the indications for use, model characteristics, and the risks and limitations of the respective DTC medical self-diagnosing AI apps. Based on a risk-based approach, some of these apps should also be prescribed by physicians rather than being offered directly to consumers over the counter. Physicians can help direct the use of the app in question and point out material facts, such as the risk of false positives and negatives, in the patient–physician conversation. In the long run, it may also be helpful to set up a new federal entity responsible for (at least for the coordination of) all issues raised by mobile health apps, from regulation to privacy to reimbursement. While this chapter focuses on FDA regulation for DTC medical self-diagnosing AI apps, some of the suggested solutions here may also have implications for other DTC apps.

II THE CURRENT AND FUTURE LANDSCAPE OF DTC MEDICAL SELF-DIAGNOSING AI APPS

The US mobile health market is expected to grow continuously over the next decade, with medical apps (compared to fitness apps) representing the bulk of the market.⁴ Before, or instead of, visiting a doctor’s office, consumers are trying more than ever before to self-diagnose their conditions by putting keywords of their symptoms into search engines like Google or using DTC medical self-diagnosing AI apps.⁵ Approximately 80 percent of patients use the Internet for health-related searches.⁶ According to a 2017 US survey, only 4 percent (ages 61 and older) to 10 percent (ages 18 to 29) of adults used apps for self-diagnosis, but 32 percent (ages 18 to 29) to 62 percent (ages 61 and older) of adults said that they could imagine using them.⁷ Since the COVID-19 pandemic, digital health technologies have gained

⁴ Grand Review Research, *mHealth Apps Market Size, Share & Trends Analysis Report By Type (Fitness, Medical), By Region (North America, Europe, Asia Pacific, Latin America, Middle East & Africa), and Segment Forecasts, 2022–2030*, www.grandviewresearch.com/industry-analysis/mhealth-app-market.

⁵ The Smart Clinics, *Rise in Internet Self-Diagnosis*, www.thesmartclinics.co.uk/rise-in-internet-self-diagnosis.

⁶ Maria Clark, 37 Self Diagnosis Statistics: Don’t Do It Yourself, *Etactics* (December 10, 2020), <https://etactics.com/blog/self-diagnosis-statistics>.

⁷ Statista, *Percentage of US Adults That Use Apps for Self-Diagnosis as of 2017, by Age*, www.statista.com/statistics/699505/us-adults-that-use-apps-to-self-diagnose-by-age.

popularity to mitigate the spread of the virus,⁸ and the use of medical self-diagnosing apps, including those based on AI, has become a reality for more adults in the USA.⁹

In 2021, Google announced the planned launch of a pilot study of its “AI-powered dermatology tool” to help consumers find answers to their skin, nail, and hair condition questions.¹⁰ With their phone’s camera, consumers simply need to take three photos of their skin, nail, or hair concerns from different perspectives and answer a few questions, such as their skin type and other symptoms.¹¹ The app will then offer a list of possible conditions.¹² Google’s app, dubbed DermAssist, is currently CE-marked as a low-risk (so-called class I) medical device in the European Union (EU) but is being further tested via a limited market release.¹³ The CE marking indicates that the device conforms with the applicable legal requirements.¹⁴ DermAssist is not yet available in the USA and has not undergone an FDA review for safety and effectiveness.¹⁵

But Google is not the only company that is investing in dermatology apps. Indeed, a quick search in mobile app stores like Apple and Google Play reveals that there are already similar apps available to download for US consumers, such as AI Dermatologist: Skin Scanner by the developer Acina. Once consumers download this AI app, they can check their skin by taking a photo of, for example, their mole with their phone’s camera.¹⁶ Within one minute, consumers will receive a risk assessment from AI Dermatologist, including some advice concerning the next steps.¹⁷ It appears that AI Dermatologist is CE-marked as a medical device in the EU but has not undergone premarket review by the FDA.¹⁸

⁸ See, for example, Sara Gerke et al., Regulatory, Safety, and Privacy Concerns of Home Monitoring Technologies During COVID-19, 26 *Nature Med.* 1176 (2020).

⁹ See, for example, Raquel Correia, How Doctors Can Benefit from Symptom Checkers, *Infermedica* (March 2, 2021) <https://blog.infermedica.com/how-doctors-can-benefit-from-symptom-checkers>.

¹⁰ Peggy Bui & Yuan Liu, Using AI to Help Find Answers to Common Skin Conditions, *Google, The Keyword* (May 18, 2021), <https://blog.google/technology/health/ai-dermatology-preview-io-2021>.

¹¹ *Id.*

¹² *Id.*

¹³ Google Health, *DermAssist*, <https://health.google/consumers/dermassist>.

¹⁴ For more information on CE marking, see the new EU Medical Device Regulation (2017/745 – MDR), Art. 2(43) and, for example, Sara Gerke et al., Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare, in *Artificial Intelligence in Healthcare* (1st edn.) 295, 312 (Adam Bohr & Kaveh Memarzadeh eds., 2020).

¹⁵ Google Health, *supra* note 13.

¹⁶ AI Dermatologist, *Say No To Skin Diseases!*, <https://ai-derm.com>.

¹⁷ *Id.*

¹⁸ *Id.* AI Dermatologist is not listed on the FDA’s website for AI/machine learning (ML)-enabled medical devices marketed in the USA. See US Food and Drug Admin., *Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices*, www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices (last updated October 5, 2022). This app can also not be found in the FDA’s databases Devices@FDA, see US Food and Drug Admin., www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.cfm (last updated October 9, 2023), and DeNovo, see US Food and Drug Admin., www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/denovo.cfm (last updated October 9, 2023).

There are also other DTC medical self-diagnosing AI apps already available on the US market. A classic example is Apple's electrocardiogram (ECG) and irregular rhythms notification feature apps.¹⁹ Both apps are moderate-risk (so-called class II) medical devices that received marketing authorization from the FDA in September 2018.²⁰ They are used with the Apple Watch and are addressed directly to consumers. While Apple's ECG app is intended to store, create, transfer, record, and display a single channel ECG,²¹ Apple's irregular rhythms notification feature app detects irregular heart rhythm episodes suggestive of atrial fibrillation.²² Another example is the AI symptom checker Ada. Consumers can manage their health by answering Ada's health questions about their symptoms, such as headaches and stomach problems.²³ Ada's AI will then use its medical dictionary of medical conditions and disorders to deliver possible causes for the symptoms and offer advice.²⁴ Ada's consumer app is currently CE-marked as a class I medical device in the European Economic Area,²⁵ but, similar to AI Dermatologist, it does not appear that the app has undergone a premarket review by the FDA.²⁶

III DTC MEDICAL SELF-DIAGNOSING AI APPS AS MEDICAL DEVICES

Can the FDA regulate DTC medical self-diagnosing AI apps? The answer is yes, if they are classified as medical devices under FDCA section 201(h)(1). This section will discuss the definition of a medical device, the FDA's enforcement discretion, and a relevant exception to the medical device definition.

A *The Medical Device Definition and the FDA's Enforcement Discretion*

Under FDCA section 201(h)(1), a "device" is

an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or

¹⁹ Letter from the FDA to Apple Inc., *ECG App* (September 11, 2018), www.accessdata.fda.gov/cdrh_docs/pdf/8/DEN180044.pdf; Letter from the FDA to Apple Inc., *Irregular Rhythm Notification Feature* (September 11, 2018), www.accessdata.fda.gov/cdrh_docs/pdf/8/DEN180042.pdf.

²⁰ Letters from the FDA to Apple Inc. (September 11, 2018), *supra* note 19.

²¹ Letter from the FDA to Apple Inc., *ECG App* (September 11, 2018), *supra* note 19, at 1.

²² Letter from the FDA to Apple Inc., *Irregular Rhythm Notification Feature* (September 11, 2018), *supra* note 19, at 1.

²³ Ada, *Take Care of Yourself With ADA*, <https://ada.com/app>. For further examples of DTC medical self-diagnosing AI apps, see Ćirković et al., *supra* note 2; Aboueid et al., *supra* note 3.

²⁴ Ada, *supra* note 23.

²⁵ Class IIa under the EU MDR is currently pending; see Ada, *5.1 Is Ada a Medical Device?*, <https://ada.com/help/is-ada-a-medical-device>. The European Economic Area consists of all 27 EU member states, Liechtenstein, Norway, and Iceland.

²⁶ Ada is not listed on the FDA's website for AI/ML-enabled medical devices marketed in the USA, see FDA, *supra* note 18. This app can also not be found in the FDA's databases Devices@FDA and DeNovo, *supra* note 18. For more information, see also *infra* Section III.A.

accessory, which is ... *intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease*, in man ..., and which does not achieve its primary intended purposes through chemical action within or on the body of man ... and which is not dependent upon being metabolized for the achievement of its primary intended purposes....²⁷

From the outset, the FDA can only regulate software functions that are classified as medical devices under FDCA section 201(h)(1) (so-called “device software functions”).²⁸ In other words, the FDA has no statutory authority to regulate software functions that are *not* considered medical devices under FDCA section 201(h)(1).²⁹ There are different types of software classifications. A relevant one is “Software as a Medical Device” (SaMD), which is standalone software and, as such, counts as a medical device.³⁰ The International Medical Device Regulators Forum defines SaMD as “software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device.”³¹ For example, Apple’s ECG and irregular rhythms notification feature apps are both SaMD because they are software-only apps intended for a medical purpose.³²

Only recently, in September 2022, the FDA updated its *Guidance for Device Software Functions and Mobile Medical Applications (Mobile Medical App Guidance)* to reflect recent changes, such as the issuance of the FDA’s final *Guidance on Clinical Decision Support Software*.³³ Although the *Mobile Medical App Guidance* contains nonbinding recommendations, it represents the FDA’s current thinking on its regulatory approach to device software functions, including mobile medical apps.³⁴ The FDA defines “mobile medical apps” as mobile apps

²⁷ 21 USC § 321(h)(1) (emphasis added).

²⁸ Sara Gerke, *Health AI For Good Rather Than Evil? The Need For a New Regulatory Framework For AI-Based Medical Devices*, 20 *Yale J. Health Pol’y L. & Ethics* 433, 446 (2021).

²⁹ See US Food and Drug Admin., *FDA’s Legal Authority* (April 24, 2019), www.fda.gov/about-fda/changes-science-law-and-regulatory-authorities/fdas-legal-authority.

³⁰ Gerke, *supra* note 28, at 446. For more information on the different types of software, see, for example, US Food and Drug Admin., *Software as a Medical Device (SaMD)* (December 4, 2018), www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd.

³¹ International Medical Device Regulators Forum, *Software as a Medical Device (SaMD): Key Definitions* 6 (2013), www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf.

³² Letters from the FDA to Apple Inc., *supra* note 19; Gerke, *supra* note 28, at 447.

³³ US Food and Drug Admin., *Policy for Device Software Functions and Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff* (2022), www.fda.gov/media/80958/download; US Food and Drug Admin., *Device Software Functions Including Mobile Medical Applications* (September 29, 2022), www.fda.gov/medical-devices/digital-health-center-excellence/device-software-functions-including-mobile-medical-applications. For the new Clinical Decision Support Software Guidance, see US Food and Drug Admin., *Clinical Decision Support Software: Guidance for Industry and Food and Drug Administration Staff* (2022), www.fda.gov/media/109618/download.

³⁴ US Food and Drug Admin., *Mobile Medical App Guidance*, *supra* note 33, at 1, 3.

that incorporate device software functionalities that meet the medical device definition in the FDCA, and either are “intended ... to be used as an accessory to a regulated medical device; or ... to transform a mobile platform into a regulated medical device.”³⁵

The “intended use” is relevant for determining whether a mobile app is considered a medical device.³⁶ The term means “the objective intent of the persons legally responsible for the labeling of devices.”³⁷ Such persons are usually the manufacturers whose expressions determine the intent.³⁸ The intent can also be shown by the circumstances surrounding the product’s distribution.³⁹ For instance, the objective intent can be derived from advertising materials, labeling claims, and written or oral statements by the product’s manufacturer or its representatives.⁴⁰

In its *Mobile Medical App Guidance*, the FDA clarifies that it intends to focus its regulatory oversight on only those device software functions whose functionality could present a risk to the safety of patients if they were not to function as intended.⁴¹ This means that the FDA intends to exercise enforcement discretion over those software functions that are or may be medical devices under FDCA section 201(h)(1) but present a lower risk to the public.⁴² Enforcement discretion means that the agency does not aim to enforce requirements under the FDCA.⁴³

For example, the FDA intends to apply its regulatory oversight to device software functions that analyze images of skin lesions using mathematical algorithms and provide users with risk assessments of the lesions.⁴⁴ In contrast, for instance, the FDA considers apps exclusively intended for patient education, such as an app that helps guide patients to ask the right questions to their physician concerning their disease, as *not* being medical devices, and, thus, those apps fall outside of the FDA’s statutory authority.⁴⁵ An example of a mobile app that may meet the medical device definition, but for which the FDA intends to exercise enforcement discretion because it poses a lower risk to the public, is an app that provides a “Skill of the Day” behavioral technique that patients with diagnosed psychiatric conditions can access when experiencing increased anxiety.⁴⁶

³⁵ US Food and Drug Admin., *Mobile Medical App Guidance*, supra note 33, at 5. A mobile app is “a software application that can be executed (run) on a mobile platform (i.e., a handheld commercial off-the-shelf computing platform, with or without wireless connectivity), or a web-based software application that is tailored to a mobile platform but is executed on a server.” *Id.* at 5.

³⁶ *Id.* at 6.

³⁷ *Id.* at 6 and n.20. See also 21 CFR § 801.4 (defining the words “intended uses”).

³⁸ US Food and Drug Admin., *Mobile Medical App Guidance*, supra note 33, at 6 and n.20.

³⁹ *Id.*

⁴⁰ *Id.*; 21 CFR § 801.4.

⁴¹ US Food and Drug Admin., *Mobile Medical App Guidance*, supra note 33, at 2, 11.

⁴² *Id.* at 2, 14, 24.

⁴³ *Id.* at 2, 13.

⁴⁴ *Id.* at 27.

⁴⁵ *Id.* at 18.

⁴⁶ *Id.* at 24.

When applying the FDA's current thinking in the *Mobile Medical App Guidance* to DTC medical self-diagnosing AI apps, some of these apps are considered device software functions that are the focus of the agency's regulatory oversight. Take as an example Apple's ECG and irregular rhythms notification feature apps. Both apps are considered class II (moderate-risk) medical devices and had to undergo a pre-market review by the FDA via the so-called De Novo process before being placed on the US market.⁴⁷

However, even if DTC medical self-diagnosing AI apps are considered medical devices because they help individuals identify a disease or other condition and are considered to be "intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease,"⁴⁸ the FDA may exercise enforcement discretion over some of them if they are considered to pose a low risk to the public. For example, as mentioned previously, the consumer app Ada is currently CE-marked as a class I (low-risk) medical device in the European Economic Area.⁴⁹ However, it seems that Ada has not undergone a premarket review by the FDA.⁵⁰ One option why this is likely the case is that Ada (may) meet(s) the medical device definition in FDCA section 201(h)(1),⁵¹ but falls within the FDA's enforcement discretion because it is considered to pose a lower risk to the public. This analysis also seems to be consistent with the *Mobile Medical App Guidance*. In Appendix B of its *Guidance*, the FDA lists examples of software functions that may meet the medical device definition but for which the agency exercises enforcement discretion, including:

- "Software functions that use a checklist of common signs and symptoms to provide a list of possible medical conditions and advice on when to consult a health care provider" and
- "Software functions that guide a user through a questionnaire of signs and symptoms to provide a recommendation for the type of health care facility most appropriate to their needs."⁵²

In addition, most class I medical devices under the FDCA are also a priori exempt from premarket notification (510(k)) requirements.⁵³

⁴⁷ See supra Section II and Letters from the FDA to Apple Inc., supra notes 19.

⁴⁸ FDCA section 201(h)(1). See also infra Section III.B. (discussing whether DTC medical self-diagnosing AI apps fall under the medical device exception in FDCA section 520(o)(1)(B)).

⁴⁹ See supra Section II; Ada, supra note 25.

⁵⁰ See supra Section II.

⁵¹ See infra Section III.B. (discussing whether DTC medical self-diagnosing AI apps fall under the medical device exception in FDCA section 520(o)(1)(B)).

⁵² US Food and Drug Admin., *Mobile Medical App Guidance*, supra note 33, at 24–25.

⁵³ US Food and Drug Admin., *Class I and Class II Device Exemptions* (February 23, 2022), www.fda.gov/medical-devices/classify-your-medical-device/class-i-and-class-ii-device-exemptions. For more information on health-related products that straddle the line between devices and general wellness products, see also David Simon et al., At-Home Diagnostics and Diagnostic Excellence, 327 JAMA 523 (2022).

B *The Medical Device Exception, FDCA Section 520(o)(1)(B)*

Section 3060(a) of the 21st Century Cures Act introduced five exceptions to the medical device definition for certain software functions. One of these exceptions is particularly relevant for DTC AI apps – namely FDCA section 520(o)(1)(B), which states that “the term device, as defined in section 201(h), shall not include a software function that is intended ... for maintaining or encouraging a healthy lifestyle and is unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition;”

In 2019, the FDA issued nonbinding *Guidance on Changes to Existing Medical Software Policies Resulting from Section 3060 of the 21st Century Cures Act (Cures Act Guidance)*, in which the agency, among other things, expresses its current interpretation of FDCA section 520(o)(1)(B).⁵⁴ In particular, the FDA clarifies that FDCA section 520(o)(1)(B) means software functions that belong to the first category of general wellness intended uses, as defined in the FDA’s nonbinding *Guidance on General Wellness: Policy for Low Risk Devices (General Wellness Guidance)*,⁵⁵ and that are “unrelated to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition.”⁵⁶ Software functions that fall within the first category of general wellness intended uses are intended for “maintaining or encouraging a general state of health or a healthy activity.”⁵⁷ For example, an app that assists users with weight loss goals and does not make any reference to diseases or conditions falls under FDCA section 520(o)(1)(B), and, thus, is *not* considered a medical device under FDCA section 201(h)(1).⁵⁸

In its *Cures Act Guidance*, the FDA also clarifies that software functions that fall within the second category of general wellness intended uses, as defined in the *General Wellness Guidance*, are not covered by FDCA section 520(o)(1)(B).⁵⁹ Software functions that fall within the second category of general wellness intended uses have “an intended use that relates the role of healthy lifestyle with helping to reduce the risk or impact of certain chronic diseases or conditions and where it is well understood and accepted that healthy lifestyle choices may play an important role in health outcomes for the disease or condition.”⁶⁰

In contrast to the first category of general wellness intended uses, this second category relates to the prevention or mitigation of a disease or condition, and, thus, software functions that fall within this second category are *not* excluded from the

⁵⁴ US Food and Drug Admin., *Changes to Existing Medical Software Policies Resulting from Section 3060 of the 21st Century Cures Act: Guidance for Industry and Food and Drug Administration Staff* 4–7 (2019), www.fda.gov/media/109622/download.

⁵⁵ US Food and Drug Admin., *General Wellness: Policy for Low Risk Devices: Guidance for Industry and Food and Drug Administration Staff* (2019), www.fda.gov/media/90652/download.

⁵⁶ US Food and Drug Admin., *supra* note 54, at 5.

⁵⁷ US Food and Drug Admin., *supra* note 55, at 3.

⁵⁸ See *id.*; US Food and Drug Admin., *supra* note 54, at 5–6.

⁵⁹ US Food and Drug Admin., *supra* note 54, at 5–6.

⁶⁰ US Food and Drug Admin., *supra* note 55, at 3.

medical device definition.⁶¹ For example, if the app in the previous example makes reference to diseases or conditions – for instance, if it claims that maintaining a healthy weight will aid living well with type 2 diabetes – this app falls outside of the scope of FDCA section 520(o)(1)(B).⁶²

As understood here, DTC medical self-diagnosing AI apps help users to identify a disease or other condition based on entering, for example, symptoms. They are related “to the diagnosis, cure, mitigation, prevention, or treatment of a disease or condition” and, thus, do *not* fall under the medical device exception in FDCA section 520(o)(1)(B).⁶³ To sum up, DTC medical self-diagnosing AI apps are medical devices under FDCA Section 201(h)(1) that are either the focus of the FDA’s regulatory oversight or for which the agency exercises its enforcement discretion. Figure 10.1 summarizes the regulation of mobile health apps, including DTC medical self-diagnosing AI apps.

IV LABELING CHALLENGES FOR DTC MEDICAL SELF-DIAGNOSING AI APPS

As established above, DTC medical self-diagnosing AI apps, as understood here, are medical devices that are either the focus of the FDA’s regulatory oversight or for which the agency exercises its enforcement discretion. This section will focus on the labeling challenges for DTC medical self-diagnosing AI apps. It will first give an overview of medical device labeling and the relevant terminology. It will then focus on labeling challenges for DTC medical self-diagnosing AI apps and make suggestions on how to overcome them.

A Labeling

Device software functions are organized into one of three classes based on their risk level, ranging from class I (lowest risk) to class III (highest risk).⁶⁴ Depending on the device classification, manufacturers must follow the associated controls – that is, General Controls, Special Controls, and/or Premarket Approval.⁶⁵ In principle, General Controls apply to all device software functions.⁶⁶ For instance, the General Device Labeling Requirements in Part 801 of Title 21 of the Code of Federal Regulations (CFR) are General Controls.⁶⁷ 21 CFR Part 801 includes, among other

⁶¹ US Food and Drug Admin., *supra* note 54, at 6.

⁶² See *id.*; US Food and Drug Admin., *supra* note 55, at 4–5.

⁶³ See US Food and Drug Admin., *supra* note 54, at 5–6; US Food and Drug Admin., *supra* note 55, at 4–5.

⁶⁴ US Food and Drug Admin., *Mobile Medical App Guidance*, *supra* note 33, at 11.

⁶⁵ *Id.* For more information on the regulatory controls, see also US Food and Drug Admin., *Regulatory Controls* (March 27, 2018), www.fda.gov/medical-devices/overview-device-regulation/regulatory-controls.

⁶⁶ For exemptions by regulations, see US Food and Drug Admin., *supra* note 65.

⁶⁷ For more information on device labeling, see, for example, US Food and Drug Admin., *Device Labeling* (October 23, 2020), www.fda.gov/medical-devices/overview-device-regulation/device-labeling.

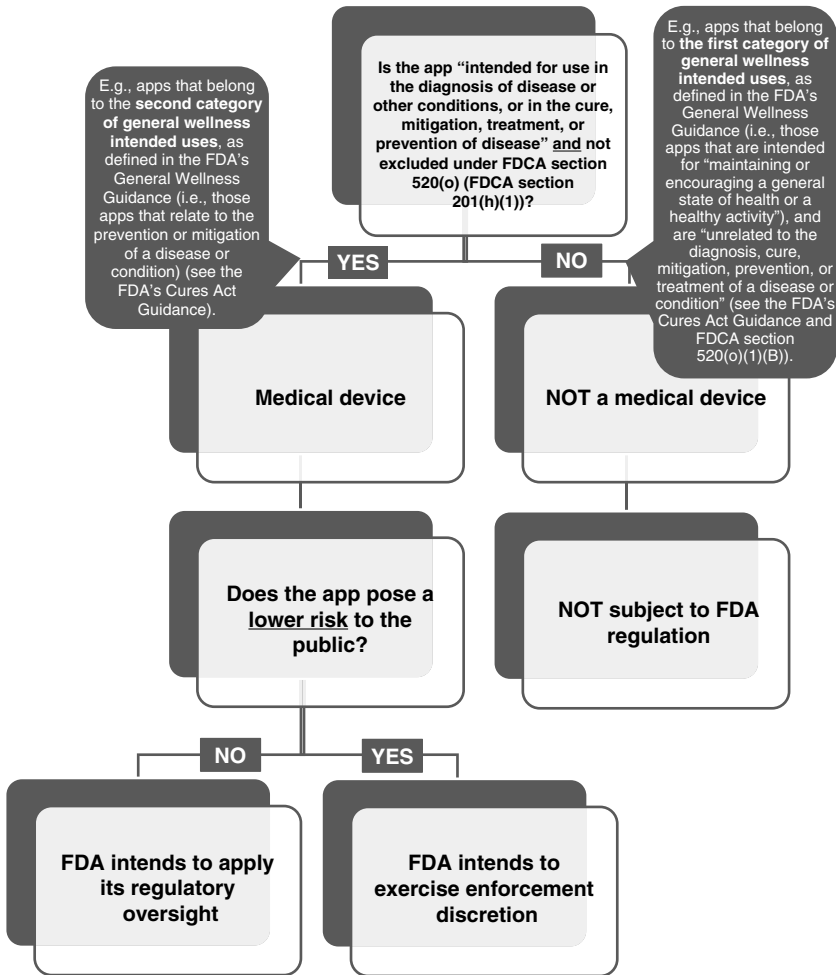


FIGURE 10.1 Regulation of mobile health apps, including DTC medical self-diagnosing AI apps^a

^a Figure inspired by the FDA’s *Mobile Medical App Guidance*, supra note 33; the FDA’s *Cures Act Guidance*, supra note 54; the FDA’s *General Wellness Guidance*, supra note 55.

things, general labeling provisions, such as the name and place of business, adequate directions for use, and the use of symbols, as well as special requirements for specific devices, such as hearing aids, and labeling requirements for unique device identification and over-the-counter devices.⁶⁸

⁶⁸ For more information, see Sara Gerke, “Nutrition Facts Labels” for Artificial Intelligence/Machine Learning-Based Medical Devices – The Urgent Need for Labeling Standards, 91 *Geo. Wash. L. Rev.* 79, Section III.A.3 and Box 1.

Labeling is defined in FDCA section 201(m) as “all labels and other written, printed, or graphic matter (1) upon any article or any of its containers or wrappers, or (2) accompanying such article.” It is a generic term that also includes all labels.⁶⁹ Under FDCA section 201(k), the term “label” means

a display of written, printed, or graphic matter upon the immediate container of any article; and a requirement made by or under authority of this Act that any word, statement, or other information appear on the label shall not be considered to be complied with unless such word, statement, or other information also appears on the outside container or wrapper, if any there be, of the retail package of such article, or is easily legible through the outside container or wrapper.

In the context of DTC medical self-diagnosing AI apps, the label will usually be available in non-physical form through the app itself.

It is also worth noting that if the “labeling is false or misleading in any particular,” the device is considered misbranded under FDCA section 502(a)(1). The term “misleading” means that the labeling proves deceptive to device users and creates or leads to a false impression in their minds.⁷⁰ For example, this can be the case if the label contains exaggerated claims or if it fails to inform users about relevant facts.⁷¹

B Challenges

DTC medical self-diagnosing AI apps raise labeling challenges. This section will discuss two: First, the concern of labeling DTC medical self-diagnosing AI apps as what I call “information-only” tools, and second, particular issues associated with the use of AI, ranging from bias to adaptive algorithms. It will also make suggestions on how to address these challenges. While the following remarks focus on medical devices, they may also have implications for those DTC apps that fall outside the FDA’s statutory authority.

i Labeling as “Information-Only” Tools

Apple’s ECG and irregular rhythms notification feature apps used with the Apple Watch are both over-the-counter class II medical devices that received marketing authorization from the FDA in September 2018.⁷² As previously mentioned, Apple’s ECG app is intended to store, create, transfer, record, and display a single channel ECG.⁷³ The indications for use, however, also include, among other things, the following sentences:

⁶⁹ Id. at 123.

⁷⁰ US Food and Drug Admin., *Labeling: Regulatory Requirements for Medical Devices* (1989) 4, www.fda.gov/media/74034/download.

⁷¹ Id. For more information on misbranding, see also Gerke, *supra* note 68, at Section III.A.2.

⁷² See *supra* Section II and letters from the FDA to Apple Inc., *supra* notes 19.

⁷³ See *supra* Section II and the letter from the FDA to Apple Inc., *ECG App*, *supra* note 19, at 1.

The user is *not intended to interpret or take clinical action based on the device output without consultation of a qualified healthcare professional*. The ECG waveform is meant to supplement rhythm classification for the purposes of discriminating AFib [atrial fibrillation] from normal sinus rhythm and *not intended to replace traditional methods of diagnosis or treatment*.⁷⁴

The FDA created a new device type, namely “electrocardiograph software for over-the-counter use,” regulated in 21 CFR 870.2345, for Apple’s ECG app and substantially equivalent devices.⁷⁵ Interestingly, 21 CFR 870.2345(a) also states that “this device is not intended to provide a diagnosis.”

Moreover, as mentioned, Apple’s irregular rhythms notification feature app detects irregular heart rhythm episodes suggestive of atrial fibrillation.⁷⁶ But, much like Apple’s ECG app, this app’s indications for use include, *inter alia*, the following phrases:

*It is not intended to provide a notification on every episode of irregular rhythm suggestive of AFib and the absence of a notification is not intended to indicate no disease process is present; rather the feature is intended to opportunistically surface a notification of possible AFib when sufficient data are available for analysis. These data are only captured when the user is still. Along with the user’s risk factors, the feature can be used to supplement the decision for AFib screening. The feature is not intended to replace traditional methods of diagnosis or treatment.*⁷⁷

The FDA also created a new device type, namely “photoplethysmograph analysis software for over-the-counter use,” laid down in 21 CFR 870.2790, for Apple’s irregular rhythms notification feature app and substantially equivalent devices.⁷⁸ Similar to 21 CFR 870.2345, this regulation also clarifies that “this device is not intended to provide a diagnosis.”⁷⁹

But Apple’s apps are not the only DTC medical self-diagnosing AI apps that articulate that their device “is not intended to provide a diagnosis.” For example, Google’s 2021 announcement of its AI-powered dermatology tool says:⁸⁰ “The tool is *not intended to provide a diagnosis* nor be a substitute for medical advice as many conditions require clinician review, in-person examination, or additional testing like a biopsy. Rather we hope *it gives you access to authoritative information so you can make a more informed decision* about your next step.”⁸¹

⁷⁴ Letter from the FDA to Apple Inc., *ECG App*, supra note 19, at 1 (emphasis added).

⁷⁵ *Id.* at 1–2.

⁷⁶ See supra Section II and the letter from the FDA to Apple Inc., *Irregular Rhythm Notification Feature*, supra note 19, at 1.

⁷⁷ Letter from the FDA to Apple Inc., *Irregular Rhythm Notification Feature*, supra note 19, at 1 (emphasis added).

⁷⁸ *Id.* at 1–2.

⁷⁹ 21 CFR § 870.2790(a).

⁸⁰ For more information on this tool, see supra Section II.

⁸¹ Bui & Liu, supra note 10 (emphasis added).

In addition, Google's website states: "DermAssist is intended for informational purposes only and does not provide a medical diagnosis."⁸² The same is also true for the AI Dermatologist: Skin Scanner app.⁸³ When looking up information about the app in an app store, the preview states: "It is essential to understand that an AI-Dermatologist is not a diagnostic tool and cannot replace or substitute a visit to your doctor."⁸⁴ App store previews of Ada say something similar: "CAUTION: The Ada app cannot give you a medical diagnosis.... The Ada app does not replace your healthcare professional's advice or an appointment with your doctor."⁸⁵

Consequently, DTC medical self-diagnosing AI apps are labeled as "information-only" rather than "diagnostic" tools.⁸⁶ Irrespective of whether DTC medical self-diagnosing AI apps are medical devices that are the focus of the FDA's regulatory oversight or for which the agency exercises its enforcement discretion, these apps seem to have in common that their manufacturers claim they are "not intended to provide a diagnosis." This is likely due to their over-the-counter nature, although Apple's clinical study of the ECG app, for example, showed that the app correctly diagnosed atrial fibrillation with 98.3 percent sensitivity and 99.6 percent specificity.⁸⁷ As a comparison, a prescription device is a "device which, because of any potentiality for harmful effect, or the method of its use, or the collateral measures necessary to its use is not safe except under the supervision of a practitioner licensed by law to direct the use of such device."⁸⁸ But do patients and consumers really understand that Apple's ECG app and similar apps are not intended to replace traditional diagnosis and treatment methods, let alone that some have been FDA reviewed and others have not?

There appears to be a significant discrepancy between the user's perception of the intended use of DTC medical self-diagnosing AI apps and their actual intended use (i.e., not to diagnose). Indeed, a recent study on AI-assisted symptom checkers revealed that 84.1 percent of respondents perceive them as diagnostic tools.⁸⁹ In addition, 33.2 percent of respondents use symptom checkers for deciding whether to seek care, and 15.8 percent of respondents said they use them to receive medical advice without seeing a doctor.⁹⁰ However, as seen above, apps like the ones from Apple

⁸² Google Health, *supra* note 13.

⁸³ For more information on this app, see *supra* Section II.

⁸⁴ AI Dermatologist, *App Store Preview*, <https://apps.apple.com/tnt/app/ai-dermatologist-skin-scanner/id1511472597>.

⁸⁵ Ada, *App Store Preview*, <https://apps.apple.com/app/id1099986434?mt=8>. For more information on Ada, see also *supra* Section II.

⁸⁶ The indications for use are usually included in the directions for use and part of the labeling requirements of over-the-counter devices; see 21 CFR § 801.61(b).

⁸⁷ US Food and Drug Admin., *De Novo Classification Request for ECG App*, 11, www.accessdata.fda.gov/cdrh_docs/reviews/DEN180044.pdf.

⁸⁸ 21 CFR § 801.109.

⁸⁹ Ashley ND Meyer et al., Patient Perspectives on the Usefulness of an Artificial Intelligence-Assisted Symptom Checker: Cross-Sectional Survey Study, 22 *J. Med. Internet Res.* e14679 (2020).

⁹⁰ *Id.*

and other companies have clear indications for use, and, thus, are likely *not* considered deceptive to device users and, thus, not “misleading” under FDCA section 502(a)(1).⁹¹ Nevertheless, even if one cannot establish misleading labeling under the FDCA, there is this misperception among users that these apps are diagnostic tools.

This misperception can also be due, among other things, to the fact that many users may not read the labels. Labeling has many benefits, including helping patients and consumers to make more informed decisions, such as by informing them about the potential limitations of an app. But if users do not read the labels and accompanying statements and language like “this device is not intended to provide a diagnosis” is buried somewhere within them, using DTC medical self-diagnosing AI apps can become risky and jeopardize patient health. For example, imagine a patient uses an app like AI Dermatologist and screens herself for skin cancer. What if the AI misses a melanoma, and the patient does not see a doctor because she perceives the app as a diagnostic tool and believes everything is alright?

Regulators and stakeholders, such as app developers, need to better educate users of DTC medical self-diagnosing AI apps, for example, about the indications for use, whether the app has undergone FDA review, and its risks. With the right design, labels could help to achieve these goals. Several groups have already shown the benefits of “eye-popping” label designs, such as with the help of “nutrition” or “model facts” labels.⁹² In particular for apps, there is a multitude of possible design options (e.g., pop-up notifications in plain language) to make users more effectively aware of important information.⁹³ Thus, regulators like the FDA could – with the help of stakeholders and label designers – develop user-friendly label design options for DTC medical self-diagnosing AI apps.⁹⁴ Once created, additional educational campaigns could be used to promote the proper reading of the label.⁹⁵ Human factors testing would also be helpful, particularly to see whether users understand when to seek medical help.⁹⁶

In addition, as part of its regulatory review, the FDA should consider whether some of these apps should be prescribed by doctors rather than being offered directly to consumers over the counter.⁹⁷ The advantage could be that physicians could

⁹¹ For more information on misleading labeling, see *supra* Section IV.A.

⁹² See, for example, Mark P. Sendak et al., Presenting Machine Learning Model Information to Clinical End Users With Model Facts Labels, 3 *npj Digit. Med.*, 41, 3 (2020); Andrea Coravos et al., Modernizing and Designing Evaluation Frameworks for Connected Sensor Technologies in Medicine, 3 *npj Digit. Med.*, 37, 8 (2020). For more information, see Sara Gerke, *supra* note 68, at Section IV.B.

⁹³ See Sara Gerke, *Digital Home Health During the COVID-19 Pandemic* (1st edn.) 141, 160 (I. Glenn Cohen et al. eds., 2022).

⁹⁴ See also Gerke, *supra* note 68, at Section IV.B (suggesting “nutrition facts labels” as a promising label design for AI/ML-based medical devices).

⁹⁵ See *id.*

⁹⁶ See *id.*; Gerke et al., *supra* note 8, at 1178.

⁹⁷ See Babic et al., *supra* note 3, at 286; Gerke et al., *supra* note 3, at 1–2.

assist patients with the use of the app in question and point out material facts in the patient–physician conversation. A risk-based approach may likely be useful here to determine such “prescription apps.”

Moreover, there is a general question of whether the FDA’s current approach to practice enforcement discretion over many DTC medical self-diagnosing AI apps is convincing. Other countries have come up with different regulatory designs to better protect consumers. For example, Germany incentivizes manufacturers of even low-risk apps (i.e., class I devices) to produce high-quality apps that comply with specific standards (e.g., safety, privacy, etc.) by offering insurance coverage for their apps in return.⁹⁸ While the FDA does not currently seem to have the resources to execute a similar approach and review all DTC medical self-diagnosing AI apps, the flood of mobile health apps and all the associated issues,⁹⁹ ranging from poorly designed products to inadequate data protection, to labeling issues and misperceptions concerning their use, requires a new regulatory approach in the long run. A better option might be to create a new federal entity in the future that would be responsible for (at least the coordination of) all issues raised by mobile health apps, including DTC medical self-diagnosing AI apps, from regulation over privacy to enforcement actions and reimbursement.

ii Particular Issues of AI: From Bias to Adaptive Algorithms

Another labeling challenge that DTC medical self-diagnosing AI apps raise is that they are not only directly addressed to consumers without a licensed practitioner’s supervision, but that they also operate using AI. Indeed, AI-based medical devices, including DTC medical self-diagnosing AI apps, are very different from traditional medical devices, such as simple tongue depressors.¹⁰⁰

First, DTC medical self-diagnosing AI apps may use methods like deep learning that make them opaque (often dubbed “black boxes”).¹⁰¹ This means that the end users of the DTC medical self-diagnosing AI app (and likely even the app developers) cannot understand how the AI reached its recommendations and/or decisions. Second, DTC medical self-diagnosing AI apps may be biased. AI tools are prone to different types of bias, ranging from biased data fed to them (e.g., a skin cancer screening app that is largely trained on white skin images) to label choice biases (e.g., the algorithm uses an ineffective proxy for ground truth).¹⁰² Third, DTC

⁹⁸ Gerke et al., *supra* note 3, at 1–2.

⁹⁹ See, for example, Babic et al., *supra* note 3; Gerke, *supra* note 68; Gerke & Rezaeikhonakdar, *supra* note 3; Simon et al., *supra* note 53.

¹⁰⁰ See Gerke, *supra* note 68, at Section III.B.

¹⁰¹ See *id.* at Sections I.A.2 and III.B.2. For more information on deep learning, see, for example, Kun-Hsing Yu et al., *Artificial Intelligence in Healthcare*, 2 *Nature Biomed. Eng’g* 719, 720 (2018).

¹⁰² See Gerke, *supra* note 68, at Section III.B.1; Ziad Obermeyer et al., *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 *Science* 447 (2019).

medical self-diagnosing AI apps may continuously learn from new data (e.g., health information, images, etc.) supplied by consumers using such apps (so-called “adaptive algorithms”).¹⁰³ These apps are, thus, much more unpredictable in terms of their reliability and would preferably need constant monitoring to avoid introducing new biases, for example.¹⁰⁴ Lastly, the human–AI interaction is complex. In particular, DTC medical self-diagnosing AI apps that have unique characteristics as their outputs are often probabilistic and, thus, require consumers to incorporate the information received into their own beliefs.¹⁰⁵ In addition, DTC medical self-diagnosing AI apps are usually available for little money or even for free.¹⁰⁶ They can easily be used as often as consumers wish.¹⁰⁷ For example, consumers of a skin scanner app may decide to scan their moles many times (rather than just once), which increases the chance of false-positive results – that is, the app detects a potential disease that is not actually present.¹⁰⁸ Because consumers are typically risk-averse about their health outcomes, they may seek medical help when it is not needed, further overburdening the health care system and taking away limited resources from patients who are more likely to need them.¹⁰⁹

Despite the differences between AI-based medical devices, such as DTC medical self-diagnosing AI apps, and traditional medical devices, such as simple tongue depressors, there are currently no labeling requirements for medical devices specifically aimed at AI (see Title 21 of the CFR).¹¹⁰ The FDA has not yet developed any labeling standards for AI-based medical devices, let alone those directly addressed to consumers.¹¹¹ Thus, when creating the optimal design labels for DTC medical self-diagnosing AI apps,¹¹² the FDA should also focus on the content and use this opportunity to develop labeling standards for AI-based medical devices, including those that are DTC.¹¹³

It is crucial that consumers know and understand, among other things, the indications for use, model characteristics, and the risks and limitations of AI-based medical devices.¹¹⁴ For example, users of DTC medical self-diagnosing AI apps should be made aware of the type of AI used (e.g., a black box, an adaptive algorithm, etc.) and the risks associated with using the app in question. They should also be informed about the various risks of bias and warned against blindly relying on the app’s

¹⁰³ See Gerke, *supra* note 68, at Sections III.B.3.

¹⁰⁴ See Boris Babic et al., *Algorithms on Regulatory Lockdown in Medicine*, 366 *Science* 1202, 1204 (2019).

¹⁰⁵ Babic et al., *supra* note 3, at 284.

¹⁰⁶ *Id.* at 283.

¹⁰⁷ *Id.*

¹⁰⁸ *Id.* at 284–85.

¹⁰⁹ *Id.* at 283.

¹¹⁰ See Gerke, *supra* note 68, at Section III.A.3.

¹¹¹ *Id.*

¹¹² See *supra* Section IV.B.i.

¹¹³ See Gerke, *supra* note 68, at Section IV.A.

¹¹⁴ *Id.*

outputs. Moreover, consumers should be alerted to the fact that increased testing can lead to an increased chance of false positives and generally be educated about the risks of false-positive and false-negative results, including when to see a doctor. A discussion with stakeholders needs to occur as soon as possible on the content of the labels of AI-based medical devices, including DTC medical self-diagnosing AI apps.¹¹⁵ In particular, the language used for the labeling of these devices will need to be plain when they are DTC.¹¹⁶

V CONCLUSION

The digital health apps market is booming, and DTC medical self-diagnosing AI apps are emerging that help users to identify a disease or other condition based on entering, for instance, symptoms. Examples of such apps include Apple's ECG and irregular rhythm notification feature apps, Google's AI-powered dermatology tool, the AI Dermatologist: Skin Scanner app, and the symptom checker Ada. DTC medical self-diagnosing AI apps raise a multitude of challenges, including questions of labeling. What should labels directly addressed to consumers look like? What information should be included in such a label?

This chapter has argued that the FDA should develop user-friendly labeling standards for AI-based medical devices, including DTC medical self-diagnosing AI apps. For example, consumers need to be effectively informed about the type of AI used (e.g., a black box, an adaptive algorithm, etc.), the various risks of bias, the risks of false-positive and negative results, and when to seek medical help. In particular, the design of such labels needs to promote their reading so that users are made aware that the DTC medical self-diagnosing AI app in question is an "information-only" tool and is "not intended to provide a diagnosis." Additionally, some of these apps should be prescribed by a doctor, not offered over the counter, based on a risk-based approach so that the doctor can point out key facts. In the long run, it may also be helpful to create a new federal entity responsible for (at least the coordination of) all issues raised by mobile health apps, ranging from regulation to privacy to reimbursement.

¹¹⁵ *Id.*

¹¹⁶ *Id.* at 145, 160.