Molecular Imaging CRO Network

Micron's ViewPoint

Clinical Trials for Medical Imaging Diagnostic Decision Support Systems

- The Social Implementation of Medical Artificial Intelligence -

Contents

Introduction

Medical Imaging Diagnostic Decision Support Systems (MIDDSS), the systems that use medical images such as CT, MRI and ultrasound imaging to assist doctors in diagnosis, are expected to support doctors' diagnoses by preventing diseases from being missed, reducing the time required for doctors to read images, and minimizing variable diagnoses between doctors.[1] In recent years, the development of artificial intelligence (AI) technology has led to significant changes in the provision of medical services, including the creation of a growing market of MIDDSS to facilitate AI.

The programming behind MIDDSS is classified as medical device.[2] Therefore, manufacturers desiring to sell MIDDSS must obtain certification or approval for the products to sell them as medical devices as per the Act on Securing Quality, Efficacy and Safety of Products Including Pharmaceuticals and Medical Devices, Pharmaceuticals and Medical Devices Act in Japan.

Manufacturers and academia may face some obstacles in bringing AI-based MIDDSS to market (Figure 1). In particular, manufacturers and venture companies that have recently entered this market may face additional difficulties due to inexperience in this space. Clinical trials for MIDDSS are an example of one such difficulty manufactures may face in bringing these systems to market. The efficacy and safety of MIDDSS must be demonstrated through clinical trials before such systems can go to market. Poorly conducted clinical trials can be very expensive and time consuming, thus necessitating the development of appropriately designed efficient clinical trials for MIDDSS.[3]

In the previous issue of Micron's ViewPoint, we discussed the importance of reference standards in MIDDSS clinical trials. In this article, we endeavor to outline some issues that need to be considered in the development of MIDDSS clinical trials. It should be noted that this article does not apply to MIDDSS, whose performance will continue to improve after coming to market.

Figure 1. Concerns of MIDDSS Developers

Trends in the Use of AI-Based Medical Devices

In the United States, AI-based medical devices are classified as Software as a Medical Device (SaMD) and are subject to regulation by the Food and Drug Administration (FDA).[4] The number of AI-based medical devices approved by the FDA including both Premarket Notification 510(k), De Novo submission and PMA-Premarket Approval has been increasing steadily, with over 40 such products now on the market[5] (Figure 2). Examples of AI-based medical devices now in the market include IDx-DR[6], a conditional First Reader Type Computer-Aided Detection/Diagnosis (CAD) system for the diagnosis of diabetic retinopathy, and QVCAD[7], a system which assists in interpreting automated breast ultrasound imaging.

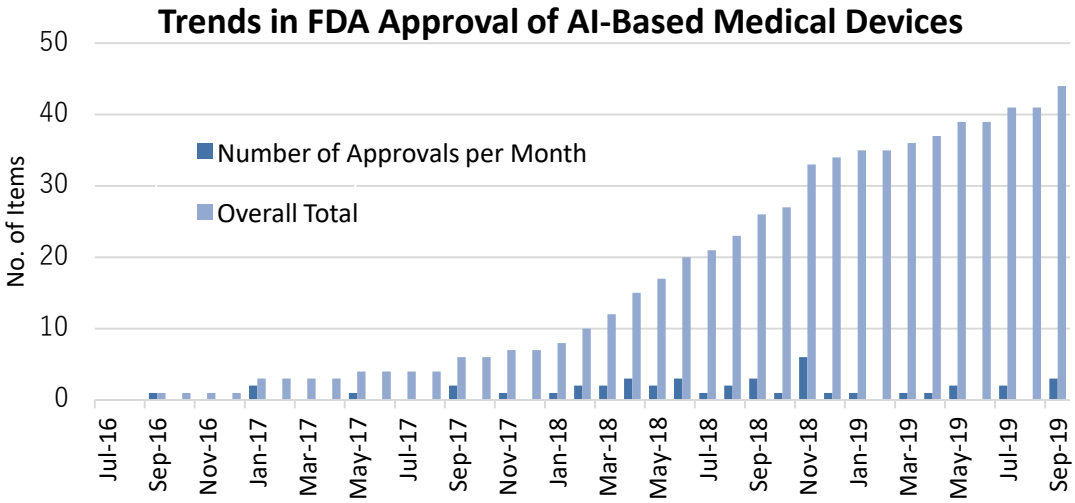

Figure 2. Trends in FDA Approval of AI-Based Medical Devices

In Japan, under the Drugs and Medical Devices Act, two products have been approved as programmed medical devices equipped with AI (as of January 2020): EndoBRAIN®[8] - diagnostic support software for colonoscopy images, and EIRL Aneurysm[9] - a system that uses deep learning to support the diagnosis of cerebral aneurysms from brain MRIs. Further regulatory approvals are expected in numerous medical specialties and diagnostic imaging modalities.

The Classification of MIDDSS

MIDDSS for detecting and diagnosing the presence or absence of disease and conditions are also known as CAD (Computer Aided Detection/Diagnosis).[1] There are multiple different possible functions of CAD and numerous types of applications in clinical workflow[10] (Figure 3). Consequently, the evidence required for regulatory approval varies between different CAD systems. It is thus important to clearly define them early in the development phase in order to efficiently bring CADs to market.

Functions of CAD

- CADe (Computer Aided Detection)

  The computer automatically detects suspicious lesions in the image, and displays marked candidate lesions to the physician.

- CADx (Computer Aided Diagnosis)

  In addition to detecting suspicious lesions, CADx also provides physicians with information about the nature of diagnosis, such as whether a target lesion is benign or malignant, or the stage of a disease.

Clinical Application

- Second Reader Type

  Physicians interpret the images without using CAD (i.e. normal image interpretation). They then interpret the images again with added perspective of CAD.

- Concurrent Reader Type

  Physicians interpret all images while simultaneously referring to CAD results.

- First Reader Type

  CAD interprets the images in isolation and identifies potential lesions. Physicians read only the target lesions proposed by CAD.

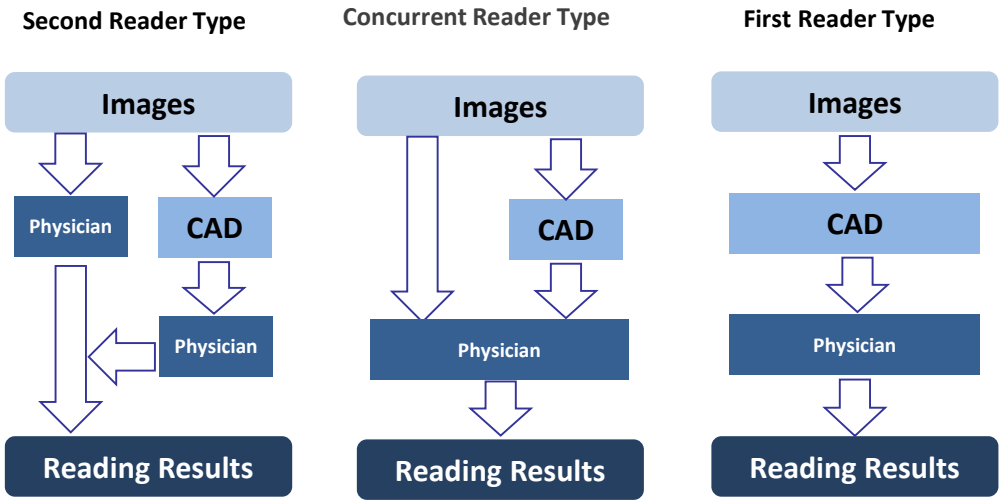

Figure 3. Clinical Application of MIDDSS

A Flow Chart Example of MIDDSS Clinical Trial

Figure 4 illustrates an example of the process of a clinical trial for a MIDDSS in the form of a flow chart. Please note that this process may change according to the classification of a MIDDSS and in consultation with Pharmaceuticals and Medical Devices Agency (PMDA).

PMDA Consultation	There are different types of interviews at each stage of the development process, such as pre-development consultation or protocol consultation. It is suggested to consult closely with PMDA to ensure an efficient development process.
Preparation of Study Protocol	Develop a standalone performance assessment and interpretation test plans. In order to develop a robust testing plan, it is important to involve a multidisciplinary team and set up efficient study design.
Collection of Image Data	There are two types of collection according to the product's application: retrospective collection - the collection of image data previously acquired; and prospective collection - that image data acquired after subject registration, with images are taken according to the imaging protocol.
Data Cleaning	Involves reviewing data to detect and correct errors. Where medical judgement is required to determine the validity of the data, an expert committee may need to be assembled.
Create a Gold Standard	The gold standard refers to the true state of patients as per the most accurate diagnostic results. This provides the reference data (accurate data) to compare with the output results of the MIDDSS.
Standalone Performance Assessment	One of the goals is to calculate system performance, such as sensitivity and specificity, by comparing individual output results of the MIDDSS with gold standards.
Preliminary Test	The preliminary test is a small-scale test before the interpretation test. One of the purposes of the preliminary test is to statistically calculate the number of cases and readers required for the interpretation test.
Interpretation Test	This is a test to evaluate the benefits of interpretation results made by physicians using MIDDSS compared with those without MIDDSS.
Statistical Analysis	By calculating sensitivity, specificity and area under ROC curve, the usefulness of MIDDSS is proved statistically.
Approval Application	According to the statistical analysis results - a summary report of the experiment is created, and the required materials and clinical data package are prepared for approval submission.

Figure 4. Flow Chart Example of MIDDSS Clinical Trial

Factors to Consider in MIDDSS Clinical Trials

Figure 5 demonstrates the key factors that should be considered when designing MIDDSS clinical trials. Like any project, diligence in the initial phase is of utmost importance in preventing unwanted delays later on and unnecessary backtracking. Accordingly, PMDA consultation and research program preparation should be the most time-consuming and resource-intensive part of the project, with this being the key to ensuring a successful clinical trial.

MIDDSS can begin to be commercialized based on the following considerations.

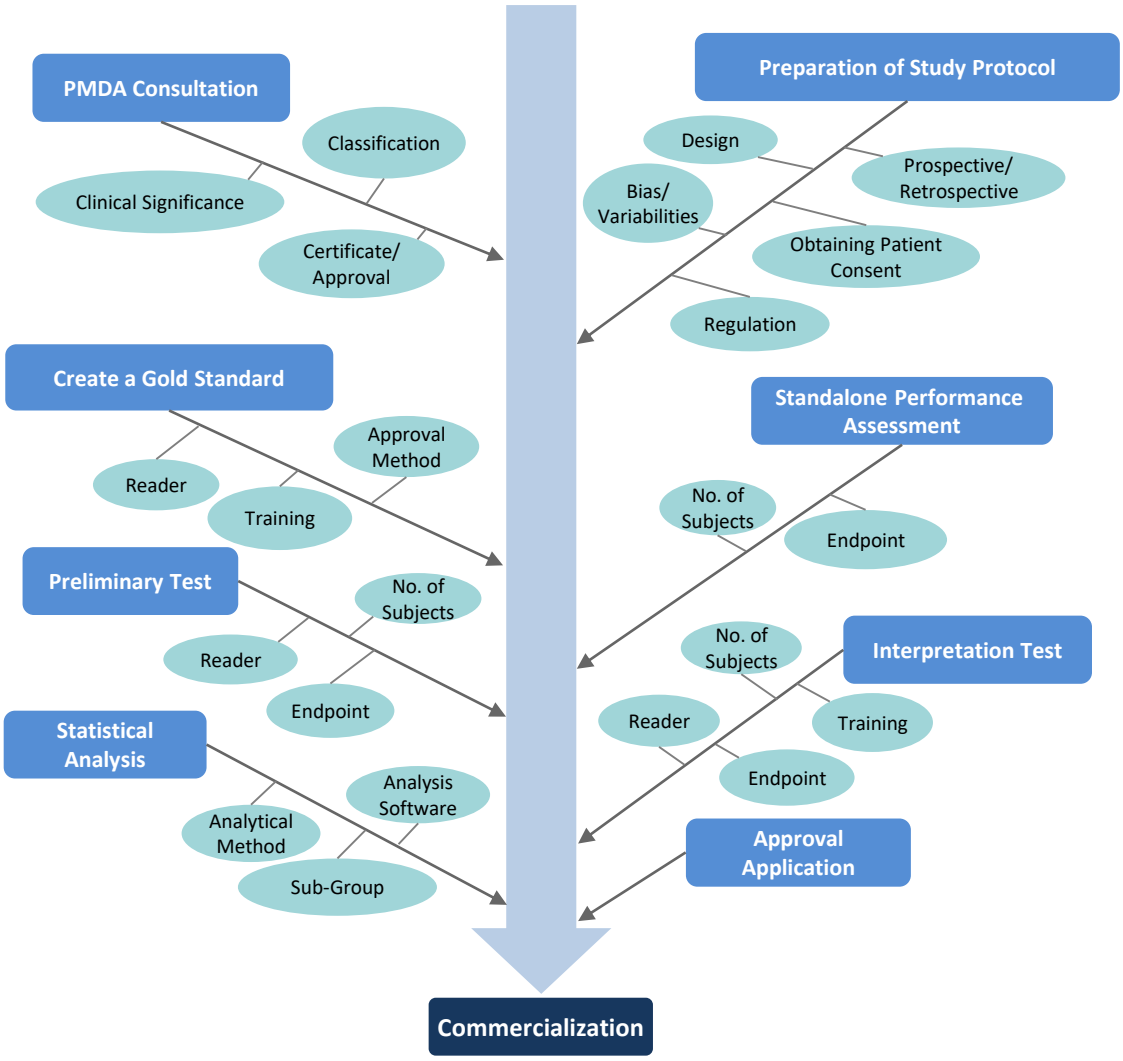

Figure 5. Factors to Consider in MIDDSS Clinical Trials

The Clinical Significance of MIDDSS

Just as the consideration of the potential clinical significance of MIDDSS often lies at the foundation of developing medical device, it is also of paramount importance in the design of clinical trials.

Product developers are often motivated by a desire to design new and innovative products that perform well. Clinical sites however are only interested in products that assist in solving an existing problem. Thus it is ideal if developers and clinical sites collaborate early in product development to ensure medical devices are developed according to the needs of clinical sites (Figure 6). The clinical relevance of the devices will become clearer as the development progresses. Once clinical benefits are identified, development concepts, accuracy of target detection/diagnosis and risks involved can be derived.

However, in the upsurge of AI, there has been a development based by AI technology, and the demand for clinical practice has often been placed in a secondary position. Therefore, if the clinical significance is defined at least at the time of the initial PMDA consultation, the subsequent clinical trials plan and application approval will likely proceed smoother.

Communication and discussion in clinical sites are particularly important for medical devices, because they may be modified or improved according to the feedback of clinical sites even during clinical trials.

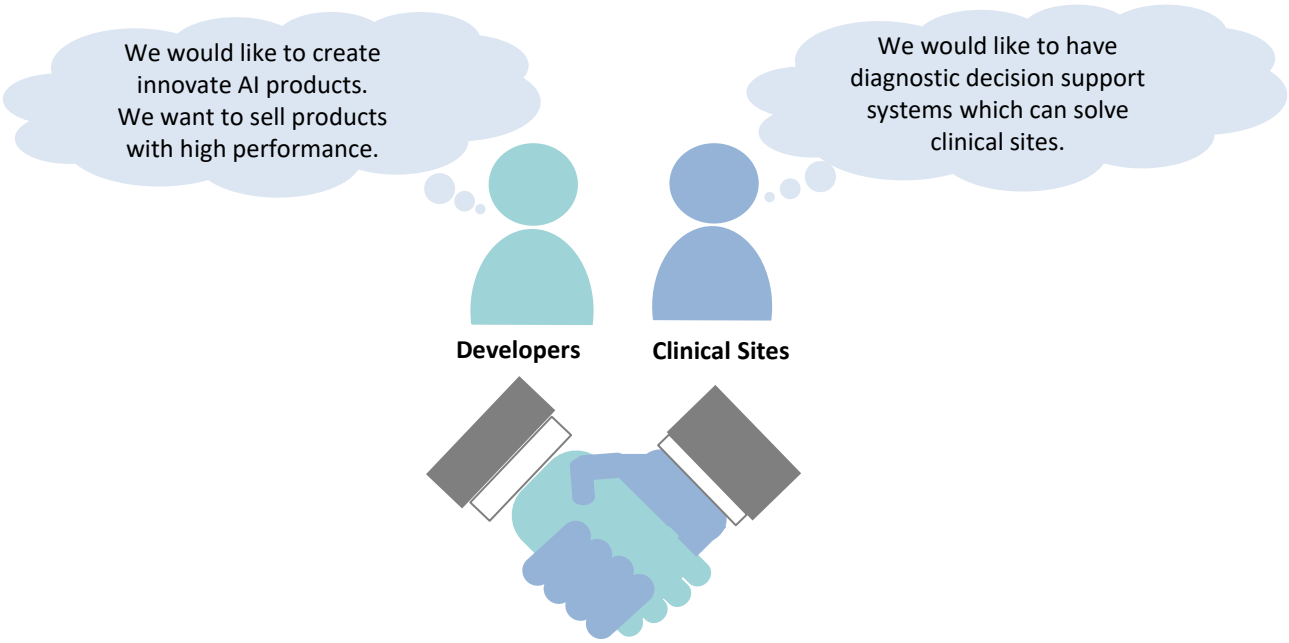

Figure 6. Development of Medical Devices Based on Needs

The Six Levels of Evaluation of the Clinical Efficacy of MIDDSS

When developing and commercializing MIDDSS, the impact of their results on patients, medical professionals, and society must be considered. The impact of diagnostic results using MIDDSS on efficacy can be divided into six levels[11] (Figure 7). If the six levels of the evaluation of the clinical efficacy of MIDDSS is kept in mind in the development process, the clinical significance of the system will become clearer.

The six levels of clinical efficacy evaluation are also useful in considering the endpoints of MIDDSS clinical trials. The second Level 2 “Assessment of the Accuracy of Diagnosis” refers to the sensitivity and specificity of the system itself, and Level 3 “Effectiveness in Diagnosis” refers to the influence of the systems on the diagnostic ability of readers (such as ROC analysis for readers with or without systems). It is also important to consider the endpoints of clinical trials in the six levels of clinical efficacy evaluation, because the basic purpose of efficacy-based clinical trials is to prove that there is a significant difference in detection and diagnostic performance between readers using MIDDSS and those not using MIDDSS.

Level 6 Significance of Social Influence	Cost efficiency which is highly dependent on cost-effectiveness analysis
Level 5 Effectiveness of QOL	From the diagnostic systems, be able to tell if there is an improvement in patient health or well-being
Level 4 Effectiveness in Treatment	Diagnostic system results impact patient treatment
Level 3 Effectiveness in Diagnosis	Does the diagnostic test impact a physician’s opinion on the presence or absence of disease?
Level 2 Assessments of the Accuracy of Diagnosis	Abilities to be tested, compared with the gold standard, the presence or absence of disease (sensitivity, specificity and ROC analysis, etc.)
Level 1 Assessments of Technical Quality	Are the anatomical and pathological details properly reflected in the image (physical measurement such as contrast, spatial resolution and noises etc.)?

Figure 7. 6 Levels of Evaluation of the Clinical Efficacy of MIDDSS

The Endpoints of MIDDSS

An important part of detailing the clinical significance is to determine who will be in the clinical environment, under what circumstances, and what the expected outcomes are. From these elements, the endpoints of clinical trials can be obtained and an appropriate test implementation plan can be constructed.

Under What Circumstances?

What is the expected outcome of the MIDDSS in a particular clinical situation?

For example, the population of patients and the characteristics to be detected (sensitivity and specificity adjustment) will change between screening and examination. The concept of endpoint also depends on whether the system is used as a second reader, concurrent reader or first reader in clinical application.

Who?

Who is the expected target to benefit from the use of the MIDDSS?

For example, will the systems only be used by radiologists, or also by internal medicine physicians and surgeons? Is it for large general hospitals or small and medium-sized clinics? Does it allow junior doctors to diagnose like experienced doctors, improve the diagnostic results of experienced doctors, or is it only suitable for specially trained doctors?

Expected Outcomes?

How do patients or doctors benefit from use of the MIDDSS?

For example, does the MIDDSS assist them in avoiding missing lesions? Does it reduce reading time? Does it only identify suspicious cases? Or does it speed up interpretation time, an important factor for acute situations? Does the MIDDSS prioritize the urgency of interpretation for physicians?

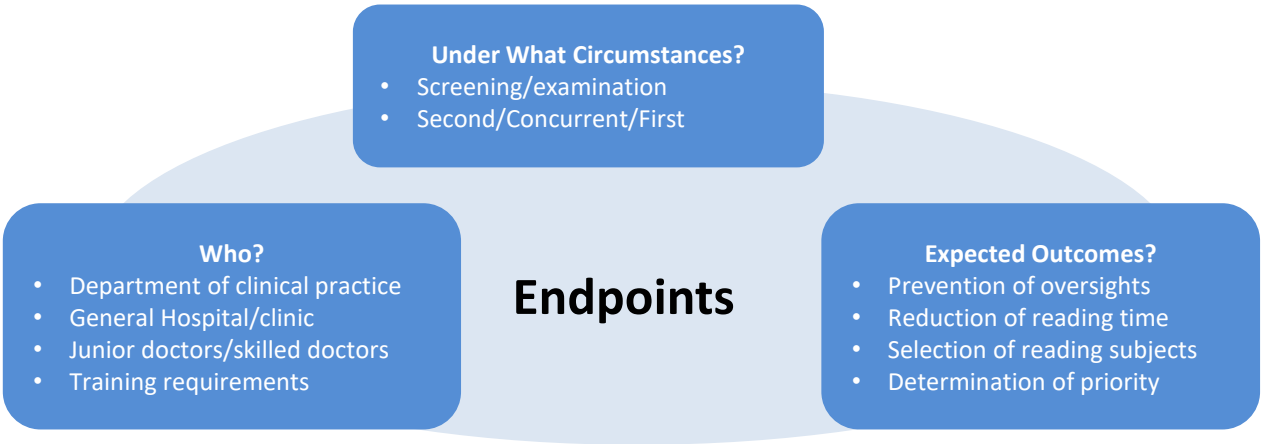

Figure 8. Endpoints of Evaluation of the Clinical Efficacy of MIDDSS

Conclusion

AI-based imaging diagnostic decision support is one of the six priority areas of the development of AI promoted by the Ministry of Health, Labor and Welfare (MHLW) in Japan. The field has strong momentum, with rapid advances in both research and development, and social implementation. The expected role of MIDDSS is mainly to reduce the diagnostic burden on doctors and improve the accuracy of diagnosis, so the best possible medical care can be provided for individual patients.

There are still many uncertainties surrounding AI-based MIDDSS. These include the role and responsibility of physicians in AI diagnosis, the interpretability that arises with the black box of output of deep learning, and plasticity, the idea that the performance of a system will change continuously after being put on the market. Regardless, it is important that all parties involved in AI work together to promote its social implementation. One of the key milestones in social implementation is the approval of the PMDA. If necessary, clinical trials are needed to prove efficacy and safety. However, if speed and quality are required, utilization of a Contract Research Organization (CRO) may be considered, which has expertise in this field. [12]

In this paper, we summarized the clinical trials of MIDDSS. In future issues of Micron's ViewPoint, we will describe in more detail how to reduce bias in clinical trials and ROC analysis, which is regularly used for endpoints in clinical trials.

Micron provides a full range of services for MIDDSS authentication and approval, from consulting, research design and operation, to PMDA approval. We also provide a range of training services depending on the type of diagnostic imaging systems you wish to authenticate/approve.

If you have any questions regarding MIDDSS, please contact Micron.

Company Overview

Headquarters (Tokyo)	3-13-16 Mita, Minato-ku, Tokyo, 108-0073, Japan Phone: +81-3-6631-3691
Osaka Branch	4-5-36 Miyahara, Yodogawa, Osaka Phone: +81-6-6399-0007
Nagoya Office	7-430 Morioka-cho, Obu, Aichi Phone: +81-562-46-2105
Business Details	1. Development support for drugs, diagnostic pharmaceuticals, and biomarkers with medical imaging techniques and know-how 2. Clinical development support (monitoring, quality control, imaging core-lab, image analysis, support for central review) 3. Support for and PET manufacturing in accordance with GMP 4. Consulting services for clinical development
Website	https://micron-kobe.com
Linkedin	https://www.linkedin.com/company/micron-imaging/
Email	imagingbiomarker@micron-kobe.com

Reference

1. 「医用画像診断支援システム（人工知能技術を利用するものを含む）開発ガイドライン２０１９（手引き）」（令和元年12月　経済産業省／国立研究開発法人日本医療研究開発機構）
https://www.meti.go.jp/policy/mono_info_service/healthcare/iryou/downloadfiles/pdf/47_guideline.pdf
2. 医療機器開発ガイド. 菊池　眞監.じほう
3. 「医療機器の迅速かつ的確な承認及び開発のための治験ガイダンス」（平成29年11年17日付け厚生労働省医薬・生活衛生局 医療機器審査管理課）
4. https://www.fda.gov/medical-devices/digital-health/software-medical-device-samd
5. 「平成30年度次世代医療機器・再生医療等製品評価指標作成事業人工知能分野審WG 報告書」（平成 31 年３月国立医 薬品食品衛生 研究所）
http://dmd.nihs.go.jp/jisedai/Imaging_AI_for_public/H30_AI_report.pdf
6. https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180001.pdf
7. https://www.accessdata.fda.gov/cdrh_docs/pdf15/P150043B.pdf
8. www.olympus.co.jp/news/2019/nr01157.html
9. https://lpixel.net/news/press-release/2019/9757/
10. 医用画像ディープラーニング入門. 藤田　広志監. オーム社
11. Fryback DG et al. The efficacy of diagnostic imaging. Medical Decision Making 1991; 11: 88–94.
12. 三澤　将史. INNERVISION. vol.34, no.12, pp54-57, 2019