Creating a social implementation platform for evidence-based healthcare services

Final report

2024/3/29

Our understanding of the project context and objectives

Context

- For effective healthcare, service providers must **enhance service quality through evidence-building**, while users (companies, governments, insurers, consumers) should opt for and persist with evidence-based services to **promote good health and prevent diseases**
- 2 key challenges exist: firstly, service providers lack **robust metrics and research designs for prevention and wellness** in building evidence for healthcare services; secondly, service users **lack access to organized expert/scientific information** required for informed purchasing decisions

Challenges identified during past projects

- In response, AMED, supported by METI, initiated a **healthcare social implementation platform project** in FY2022. Building on previous efforts by the US and the UK, which included studies on evidence-building support and evaluation frameworks, this project aims to establish the groundwork for delivering evidence-based **healthcare services to the public**. Challenges identified in the studies include:
  - **Establishing guidelines on prevention to encourage service development:** The US Preventative Services Task Force (USPSTF) guidelines are helpful for developing intervention-type services for preventative exercise or nutrition, but Japan only started looking at formulating guidelines for prevention last fiscal year
  - **Creating environments conducive to utilizing RWD from healthcare services:** Foreign early evidence-building initiatives, like Sleepio in the UK, actively leverage RWD. In some countries, regulations have been established to promote the secondary use of RWD. While Japan has advanced in building databases, there is room for improvement in establishing regulatory frameworks and supportive environments
  - **Setting up quality and usability evaluation frameworks to generate incentive for service evidence-building:** In certain non-SaMD projects, UX/UI takes precedence over evidence in influencing user decisions, allowing for the promotion of healthcare services without substantiated evidence. To shift this paradigm, there's a need to establish mechanisms (evaluation frameworks and organizations) for evidence evaluation and create systems incentivizing awareness of evidence presence or absence
- Effective and efficient systems for service providers, users, and stakeholders should not be confined to academia or government offices. They must be promoted through **dialogue and co-creation with diverse stakeholders** involved in raising awareness

Objectives

- This project aims to **comprehensively support the evidence-building to social implementation process in Japan** by reviewing global trends in research and systems. This fiscal year's support includes **studying R&D in healthcare services**, examining **the social implementation of healthcare services**, and contributing to the **healthcare social implementation platform project**

How last year's findings will support this year's studies

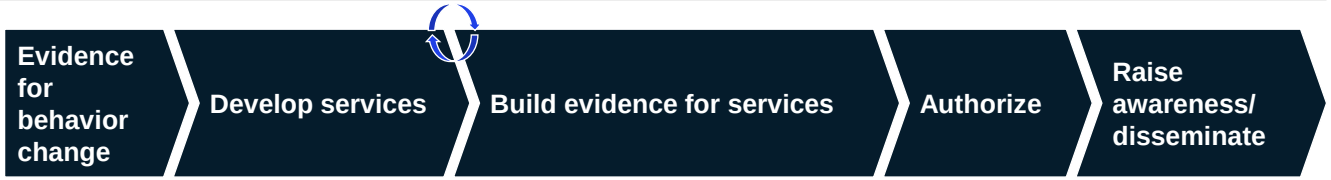
Our aim this fiscal year is to use last year's research findings to deliver comprehensive support for the entire evidence-building and social implementation process in Japan

FY2023 findings

Identified challenges in establishing a cohesive healthcare services social implementation system

- **Creating environments conducive to utilizing RWD from healthcare services**: Foreign evidence-building efforts actively use RWD; Japan needs to enhance regulatory frameworks and supportive environments
- **Setting up quality and usability evaluation frameworks to generate incentive for service evidence-building**: The current environment allows healthcare services to gain popularity without evidence. Changing this requires creating mechanisms to evaluate evidence and incentivizing awareness of its presence or absence
- **Establishing guidelines on prevention to encourage service development:** US Preventive Services Task Force includes guidelines for preventive exercise and nutrition, but Japan lacks established prevention guidelines

Positioning of FY2024 studies



Create prevention guidelines to promote service development
→(3) Support the healthcare social implementation platform project

Establishing evaluation frameworks to incentivize service evidence-building quality and usability
→(2) Study the social implementation of healthcare services

Creating environments conducive to utilizing RWD from healthcare services
(1) Study R&D in healthcare services

Contents

Study R&D in healthcare services

Approach overview

Purpose

- Validate methods for evidence-building and service development **using RWD and lifelog data in the Japanese healthcare industry**
- Suggest **essential support measures** based on overseas cases

Approach

Select disease area targeted for study

Choose an area, preferably lifestyle diseases or mental health, with existing scientific evidence and a service penetration gap between Japan and other countries. Prioritize women's health if significant RWD usage is observed overseas

Select target countries

Prefer the US and UK due to market size and numerous studies; for a comprehensive study, explore other nations like Canada or Germany with advanced RWD cases

Conduct a best-practice survey of services utilizing RWD and/or lifelog data

Study cases where services were developed for the target disease area using RWD and/or lifelog data, with a focus on the US and UK.Use the case studies to outline the process required for data-driven service development

Study support measures/systems at government and public agencies that encourage the use of RWD and/or lifelog data

Examine government support, accelerator programs, approvals, certifications, data utilization laws, platform establishment, standardization, and API integration mechanisms

(Independent) Study the use of healthcare data in domestic systems

Study Japan’s systems for building evidence using RWD and/or lifelog data

Spot potential issues in applying foreign best practices in Japan

Suggest necessary **support measures for Japan to address these challenges**

RWD Utilization in Healthcare: Summary of Findings

US-UK research

- **Clinical effectiveness:** UK companies (Liva, Sleepio, ieso) swiftly conducted RCTs post-service launch for clinical effectiveness, while US companies (Livongo, Talkspace, Wysa, Lyra, Omada) often relied on RWD for effectiveness assessments without necessarily conducting RCTs
  - In regions with universal healthcare impact (e.g., UK, Germany), RCTs promptly establish clinical effectiveness
  - In the US, without universal healthcare, RCTs aren't essential; companies only need to demonstrate comparable or superior outcomes in digital treatment and prevention programs with established evidence
- **Building evidence in line with stakeholders:** Services studied, like Talkspace (labor productivity and economic impact), Sleepio (treatment economic impact), Omada (counseling services ROI), and Lyra (employee retention), established RWE as outcome metrics
  - Omada's research assessed clinical efficacy and behavioral change outcomes by studying program engagement correlation with lower blood pressure, specifically focusing on hypotensive effects
  - Talkspace's research assessed treatment economic impact and labor productivity by examining questionnaire scores (PHQ-9 and GAD-2) to evaluate text therapy effectiveness. Cost-effectiveness was determined by multiplying the difference between in-person and text therapy costs by the number of treatment sessions
  - In the US and other situations where business activities fell outside of the universal healthcare scheme, evidence on labor productivity and economic effectiveness was built in addition to evidence on clinical efficacy depending on stakeholder interest
- **Evidence trends:** Numerous papers look at behavioral change impacts across 2 services, and these studies not only are in an area that fits RWD analysis, but they also demonstrate the usefulness of AMED service lines
- **Sample size:** Sample sizes in rapidly conducted evidence-building research using RWD were typically in the dozens or hundreds. For retrospective RWD research, sample sizes were usually in the thousands or tens of thousands, examining 5+ years post-service launch
- **Evidence evaluations:** All 8 studied companies published a minimum of 5 research papers in scientific journals. Additional information, including abstracts from academic conference presentations, company-issued white papers, case studies, and other content, was often disseminated through company blogs and various media channels

Potential study cases in lifestyle diseases and mental health

Deep-dive into cases using RWD in healthcare services identified in last year's study

		Name	Key services	Evidence-building method/examples
Lifestyle diseases	UK	**Liva app**	Online platform that matches **diabetes and other patients with health coaches**	• Liva studied the effectiveness of coaching on weight loss in obese DM patients (n = 103)
	US	**Omada**	Digital prevention and treatment program **for diabetes, high blood pressure, and other lifestyle diseases**	• Omada observed the clinical efficacy (hypotensive effects) of 12-month coaching on 1,117 patients with high blood pressure
		Livongo	Digital service providing personalized content via monitoring for diabetes patients	• Livongo conducted a feasibility study modeling interventions for individuals based on data from 16,531 participants
Mental health	UK	**ieso**	Online therapy **using AI and NLP**	• ieso tested the correlation between clinical effects and therapist utterances using data from 14,899 participants
		Sleepio	**Digital sleep improvement program**	• Sleepio tested cost reduction compared to traditional therapies using claims data from 10,705 participants
	US	**Talkspace**	Behavioral health service **providing mental health treatment**	• Talkspace studied the effects of their service on depression and anxiety as well as absence, productivity, and ROI with 51 participants
		Lyra	Cognitive behavioral therapy (CBT) **remote mental health service for employees**	• Lyra used data from 6,738 participants to evaluate clinical efficacy (alleviation of anxiety and depression)
		Wysa	AI-based **counseling service**	• Wysa studied the clinical efficacy and engagement level of their service on self-reported depression (n = 129)

Source: Company websites, FY2023 research and interviews

Lifestyle disease services from the US private sector: Livongo

Digital service providing personalized content via monitoring for diabetes patients

Details of evidence building[1]

Livongo

Research on improved outcomes through personalized recommendations via a monitoring program for diabetes patients

- **Research design:** Retrospective feasibility study (RWD analysis)
- **Proof-of-concept timing:** 5+ years after market release
- **Data selection period:** (not specified)
- **Target:** Type 2 diabetes patients
- **Comparison:** Distribution of intervention (treatment) group vs. control group along five action criteria
- **Sample size:** 16,531
- **Acquired data[1]:** Medical information (self-measured/self-reported), preferences (self-reported), service usage frequency and other engagement data, demographic data
- **Outcome metrics:** Estimated A1c = [mean BG over past 30 days + 46.7] / 28.7
- **Analysis method:** Multivariate analysis for Heterogeneous Treatment Effect (HTE) found that interacting with coaches and self-monitoring blood glucose levels led to the highest average decrease in estimated A1c. These were the most recommended actions for 54% of the population
  - 46% of members were predicted to benefit more from an alternative intervention
  - Engaging members had, on average, a 0.8% larger reduction in estimated A1c within the first 3 months compared to non-engagers

Suggestions for the AMED public offering business

- **Analysis models to estimate** Heterogeneous Treatment Effect (HTE) helps identify optimal interventions for individuals, enhancing clinical effectiveness in service development
- Evidence on effective interventions for **specific patient types** is valuable for shaping future guidelines and diabetes management programs
- **Business ideas for the public offering**
  - Joint corporate-academic research on heterogeneous treatment effects (e.g. high blood pressure, diabetes, fatty liver disease)
  - Systematic review of research analyzing heterogeneous treatment effects on lifestyle diseases

1. Kamath S, Kappaganthu K, Painter S, Madan A. Improving Outcomes Through Personalized Recommendations in a Remote Diabetes Monitoring Program: Observational Study. JMIR Form Res. 2022 Mar 21;6(3):e33329. doi: 10.2196/33329. PMID: 35311691; PMCID: PMC8981007. link

*1: Acquired data includes self-reported medical details (HbA1c, diabetes management, insulin use), preferences (communication channels, interest in health), engagement metrics (website and app use, SMBG checks, coaching sessions, steps, etc.), and demographic information (age, gender, BMI, race)

Private sector services in the American lifestyle disease field, like Omada

Digital management programs for diabetes prevention

Details of evidence building[1]

Omada

Research using invoice data to evaluate the effects of healthcare cost reductions in digital diabetes prevention programs

- **Research design:** Crosscutting observational research, real-world data (RWD) analysis
- **Timing of verification experiments:** Approximately 5 years after release (signoff as digital DPP by CDC in 2015)
- **Data extraction period:** Apr. 2015 – Mar. 2018
- **Target:** Patients at risk of Type II diabetes who are capable of light exercise
- **Comparisons:** Pre-and post-intervention, with a matched comparison group from MarketScan data (individuals not in DPP or similar programs during the observation)
- **Sample size:** Digital DPP intervention group: 2,027, and matched non-intervention group: 2,027
- **Acquired data:** Invoice data linked to unique IDs (e.g.: diagnosis name, insurance information, actual payment amount covered by insurance, status of healthcare service usage*[1]), weight, demographic information and health status
- **Outcome metrics:** Medical costs
- **Analysis method:** Analysis of difference in before/after comparison design (DID), evaluation of significant difference and trust range in regression analysis
- **Results:**
  - In the first year, compared to the control group, per capita healthcare costs for the digital DPP group dropped by USD1,169 (P=0.01), with USD699 attributed to lower hospitalization costs
  - Lower hospitalization costs were due to fewer and shorter hospital stays
  - In other items, there was no significant difference vs. the impact of healthcare cost reductions
  - In the second year, a trend toward reduced healthcare costs was observed, but without significant difference

Suggestions for the AMED public offering business

- Conducted DID analysis on healthcare cost reduction using multi-year invoice data
- **Matched using disguised invoice data as the comparison group.** This can be a reference for research in Japan using NDB-matched comparison groups in insurer or healthcare institution interventions
- **Business ideas with public offering**
  - Collaborate with service providers, insurers, or employers to conduct post-market release evidence-building research
  - Conduct comparative research on in-house services' effectiveness using NDB-matched comparison groups from a healthcare economics perspective
  - Research on healthcare cost reduction impact in prevention programs and services over multiple years using DID analysis
  - Systematic review focusing on post-market release impact on healthcare economics in lifestyle-related diseases

1. Sweet CC, Jasik CB, Diebold A, DuPuis A, Jendretzke B. Cost Savings and Reduced Health Care Utilization Associated with Participation in a Digital Diabetes Prevention Program in an Adult Workforce Population. J Health Econ Outcomes Res. 2020 Aug 18;7(2):139-147. doi: 10.36469/jheor.2020.14529. PMID: 32884964; PMCID: PMC7458495. link

*1 Status of healthcare service usage (1) Number of outpatient checkups (all; annual); (2) Number of emergency checkups (annual); (3) Number of hospitalizations (annual); (4) Number of hospitalization days (annual); (5) Number of medication days (annual)

Private sector mental health services in the UK, like ieso

Online platforms that match patients and therapists, and provide cognitive behavior therapy

Details of evidence building[1]

ieso

Apply deep learning to examine the "black box" relationship between therapist utterances and clinical impact in psychological therapy

- **Research Design:** Retrospective RWD analysis
- **Timing of verification experiments:** After release to market (Established in 2000; adopted for NHS pilot project in 2011)
- **Data extraction period:** Jun. 2012 – Mar. 2018
- **Target:** Text data from dialogues with therapists during app-based therapy sessions
- **Comparisons:** Before and after intervention
- **Sample size:** Text data from 14,899 individuals (initially analyzed 90,934 sessions for 17,572 patients)
- **Acquired data:** Responses to questionnaires, and text data exchanged between patients and therapists (Jun. 2012 ~ Mar. 2018)
- **Outcome metrics:**
  - PHQ-9 (Patient Health Questionnaire) score
  - GAD-7 (Generalized Anxiety Disorder 7-item scale) score
- **Analysis method:** Multivariate analysis (conducted 3 logistic regression analyses using R)
- **Results:**
  - More sessions and method changes (CBT's cognitive and behavioral methods) correlate with increased patient engagement and symptom improvements
  - Therapist utterances outside therapy did not correlate with patient engagement or symptom improvements

Suggestions for the AMED public offering business

- Multivariate analyses linking clinical outcomes and online therapy text data can standardize the untransparent process of psychological therapy in mental health
- Developing Japanese evidence and guidelines can contribute to mental health services' advancement
- **Business ideas with public offering**
  - Multivariate analysis research on text-based mental health services (eiso research Japanese ver.)
  - Systematic review related to the use of text data in the mental health field

1. Ewbank MP, et al. Quantifying the Association Between Psychotherapy Content and Clinical Outcomes Using Deep Learning. JAMA Psychiatry. 2020 Jan 1;77(1):35-43. doi: 10.1001/jamapsychiatry.2019.2664. PMID: 31436785; PMCID: PMC6707006. link

Private sector mental health services in the US, such as Lyra Health

Remote mental health services, offering patient-therapist matching and personalized treatment programs

Details of evidence building[1]

Lyra Health

Study on employee retention rates: Examining social welfare services for 185,000 employees across 14 companies. Patients chose between remote mental health services and traditional therapy, and attrition rates were compared over 12 months

- **Research Design:** Retrospective cohort RWD analysis
- **Timing of verification experiments:** Post-market release (Mental Health Coaching and Blended Care Therapy launched in 2018: aggressive capital procurement from around 2020, with ongoing business growth)
- **Data extraction period:** Jan. 2017 – Oct. 2019
- **Target:** Employees at 14 companies with Lyra contracts
- **Comparisons:** Lyra users, non-users, and a conventional therapy comparison group
- **Sample size:** 184,715 (participants in employee support programs)
- **Acquired data:** Data on individual employment status, use of Lyra, and healthcare invoice data*[1]
- **Outcome metrics:** Attrition rate
- **Analysis method:** Multivariate analysis (all analyses conducted using Python)
  - Confirm employment status after 12 months for persons who did not use Lyra
  - Analysis using Kaplan-Meier method*[1]
  - Assess Lyra therapy and conventional psychological therapy using Cox proportional hazard model*[1]
- **Results:**
  - 11% of Lyra users retired in 12 months, compared to 22% of non-users
  - In three companies with invoice data, Lyra users had a 7% turnover rate, compared to 15% in the conventional therapy group over 12 months
  - In the anxiety-diagnosed subgroup, Lyra users had a 20% lower risk of leaving compared to the conventional therapy group

Suggestions for the AMED public offering business

- Remote mental health services were effective from an HR perspective, particularly in health management and corporate data analysis (e.g., attrition rates) post-market release
- Example of research using multivariate analysis with strong links to employees on the exit side
- **Business ideas with public offering**
  - Collaborate with service providers, employers, or insurers to research and build evidence using post-market release data
  - Systematic review focusing on post-market release corporate data in employee mental health

1. FALCON, Maja et al. Impact of evidence-based psychotherapy on employee retention. Medical Research Archives, [S.l.], v. 9, n. 11, nov. 2021. ISSN 2375-1924 https://esmed.org/MRA/mra/article/view/2574

*1 Analysis conducted at 3 companies (total: 24,947 employees) that provided invoice data

Private sector mental health services in the US, like Wysa

Apps providing mental health care services; e.g., AI chat that expresses empathy

Details of evidence building[1]

Wysa

Research verifying the effectiveness of AI chat tools for empathy driven digital mental well-being

- **Research design:** Quasi-experimental mixed research (RWD analysis)
- **Timing of verification experiments:** Soon after market release
- **Data extraction period:** Jul.11 ~ Sept.5, 2017
- **Target:** Persons with depressive symptoms using Wysa
- **Comparisons:** Before and after intervention
- **Sample size:** 129 persons
- **Acquired data:** Status of app usage; responses to questionnaires
- **Outcome metrics:** PHQ-9 (Patient Health Questionnaire) score
- **Analysis method:** Mixed research
  - Quantitative analyses: ① Mann-Whitney test for comparing high and low users before and after intervention; ② Reporting analyses based on disguised information
  - Qualitative analysis: ① Thematic analysis to assess effectiveness in daily tasks' feasibility and ease of response; ② Machine learning algorithm analysis of "denial" and "dissatisfaction"
- **Results:**
  - Improvement in feelings (self-reported scores before and after interventions) showed greater enhancement for high-frequency users (Avg.: 5.84 [SD 6.66]; P=.03, impact volume: 0.63) than low-frequency users (Avg. 3.52 [SD 6.15])
  - 67.7% or more of users provided feedback saying that the app was useful and made them feel better

Suggestions for the AMED public offering business

- **Based on the mixed research design,** it's feasible to rapidly and easily assess the real-world effectiveness and engagement of the chatbot
- Moreover, this design may apply to large-scale disguised data analyses
- **Business ideas with public offering**
  - Mixed research on engagement and effectiveness of text-based mental health services
  - Systematic review related to the use of AI robots in the mental health field

1. Inkster B, Sarda S, Subramanian V. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. JMIR Mhealth Uhealth. 2018 Nov 23;6(11):e12106. doi: 10.2196/12106. PMID: 30470676; PMCID: PMC6286427. link

Summary of a system survey promoting real-world evidence (RWE) construction in healthcare services

- **Outline of survey:** Explore overseas frameworks and support systems for constructing RWE in healthcare services, examining their potential applicability to Japan
- **Survey results:**
  - Support system studies had ample information on pharmaceutical and medical device certification but lacked survey results for promoting RWE in healthcare services, especially in the US
  - Among overseas support case studies, we found networks supporting RWE construction essential for service proliferation. Also, case studies noted RWE frameworks by stakeholders and international industry organizations maintaining product-based libraries
- ❶ **Suggestions from surveys of positive case studies:** For constructing RWE in healthcare services, "accompanied support" in matching with stakeholders and academia partners is effective for service proliferation
  - **Sleepio:** Collaborated with NHS through Innovate UK funding, utilizing Oxford AHSN, for research on cost impact using RWD invoice data and in-house app information
- ❷ **Suggestions from system surveys (maintaining environments and platforms):**
  - Key discussions on maintaining environments and platforms focused on national database construction. In Japan, we align our conceptual activities with similar directions taken by the UK and the US
  - In the UK's NICE real-world evidence framework, policies related to RWD emphasize transparency in databases, including data sources, compatibility, and relationships with data quality
- ❸ **Suggestions from system surveys (capabilities)**: As support for service providers' evidence building capabilities, it is effective to maintain RWE frameworks and product-specific libraries
  - AHSN supports the required RWE construction for NHS service proliferation using British research networks, developing research tools and guidelines.
  - The Digital Therapeutics Alliance, an international NPO with Japanese pharmaceutical company participation, focuses on DTx development. The website holds frameworks for exit-side stakeholders and product-based evidence libraries

Summary of implications for Japan in light of the study content

Details follow

	Example studies	Description	Implications for Japan
❶**Positive example studies**	Sleepio ×Innovate UK ×Oxford ASHN	Oxford ASHN supports the RWD building of the individual company Sleepio via Innovate UK's large-scale public fund	Collaborative support, aligning with exit stakeholders and academic partners for joint research, effectively builds RWE for healthcare services
	Promoting joint research Oxford ASHN	Leverages a real-world evaluation program to support practical RWE construction, providing necessary support at all stages of development	
	Liva ×University of Southern Denmark ×the Region for Southern Denmark	3 parties contribute research costs, and ApHER performs RWD analysis and evaluates service utility	Joint research with a third-party institution with expertise is effective for evaluating health economics impact
❷**Systemic studies (environment, infrastructure development, etc.)**	Guidelines NICE Evidence Standards Framework	Guidelines on RWD, particularly "Assessing data suitability," emphasize source transparency, including data origin, fitness for purpose, quality, and their correlation	While maintaining RWD source legitimacy/transparency is crucial, strict Japanese laws on personal data protection can hinder data utilization. Future business needs may include creating guidelines for RWE construction
	Data infrastructure preparation NHS Secure Data Environments	NHS's platform for all health/nursing data, tailored for researchers at trusted institutions; current use by healthcare service providers is unconfirmed	While utilizing the national DB seems beneficial, it's anonymized, similar to Japan, preventing analysis with your own company's user data. Interviews with domestic businesses are essential to assess potential needs
❸**Systemic studies (capabilities)**	Product library DT Alliance (DTA), Product Library	Provides product-specific evidence, leveraging an understanding of DTx linked with scientific evidence to stand out among numerous health applications	A support system that publishes a RWE library organized by product is also effective

Proposed support measures for evidence building (1/2)

Issues faces by businesses and the current support measures (hypothetical)	Support measures	Details
Global examples: In the UK and Denmark, large-scale joint research is fostered through public-private sector collaboration. The UK's NIA offers selected companies tailored support from fellows across various domains (biostatistics, health economics, business development) **Issues for businesses:** Domestic businesses face challenges, such as a lack of specialized expertise in evidence building and connecting with exit stakeholders for joint evidence construction **Issues for support measures:** While there is some similar support, challenges include identifying support companies, support for evidence building and business development, and aligning with existing stakeholders	**Collaborative support that matches support needs** • Matching with stakeholders • Promoting joint research	• Supports startups' use of RWD for building evidence and actual examples leading to service scale-up — **Narrowing down target support areas:** Matching with the main business and focus areas for guideline creation — **Matching with stakeholders:** Supports evidence building with stakeholders who will become an exit for the business (municipalities, employers, public insurers) — **Determination of selected companies:** When selecting companies, consider not just research but also business viewpoints to identify those with high business potential — **Collaborative support according to the phase:** Leverage appropriate expertise to offer collaborative advice tailored to the company's needs across various phases: development, release, expansion, etc • Instead of launching new businesses, broaden the scope of similar existing businesses (preventative/Non-SaMD) or make adjustments to existing businesses (details follow the next page)
Global examples: In the UK, progress is seen in library preparation, like the DT Alliance's "DTx Products," showcasing evidence-building approaches by representative DTx products **Issues for businesses: Difficulty understanding the evidence expectations of** stakeholders for social implementation. Limited skills to assess the evidence of benchmark services, and few chances for users to view the high-level evidence from their company **Issues with existing support:** Japan lacks a public Non-SaMD library. The PMDA provides a list of approved program medical devices, but it doesn't serve as a product library	**Creating recommendations for building evidence that surpasses standards** • Specifying the evidence demanded by stakeholders • Publicly releasing evidence building guidelines/case library for businesses	• Creating recommendations for aligning viewpoints between stakeholders and service businesses about the evidence that should be built (e.g., dementia recommendations) — Designed in a way that attracts the attention of exit stakeholders to the businesses, so that they can understand what kind of evidence is required — In the recommendations, create/publish a case library specializing in preventative Non-SaMD. (e.g., DTA DTx Products) • Issue recommendations for each service group using the same framework as the dementia recommendations, rather than launching a new business. Considering publishing them on the site investigating the certification mark to enhance recognition

1. https://www.meti.go.jp/shingikai/mono_info_service/ninchisho_wg/pdf/2022_002_05_00.pdf

Proposed support measures for evidence building (2/2)

Issues faces by businesses and the current support measures (hypothetical)

Global examples: NHS's platform for all health/nursing data in the UK, designed for researchers at specific trusted institutions, etc. Current use by healthcare service providers is unconfirmed

Issues for businesses:

- In lifestyle diseases, the high hurdle for RCTs prompts the use of evidence based on RWD. While building evidence using internal data is straightforward, elevating the evidence level necessitates control group data, challenging for the company to collect independently
- For mental health, there's a lack of data on a control group's stress checks, etc. Ideally, a shareable reference database, possibly led by the private sector, would be valuable, but currently, none exists

Issues for support measures: In Japan, progress is underway for a NDB, but using control group data for service evidence building is challenging due to high approval hurdles and limitations to research purposes

Support measures

Securing reference data

- Support approval of the use of the NDB, etc. by private sector businesses
- Build the control group database required for evidence building

Details

- Support approval of the use of the NDB, etc. by private sector businesses
  - Broaden the scope of research using NDB data beyond academic purposes to include private-sector-led research for service evidence building
  - Establish a liaison to support businesses in the submission of the required application documents, etc
- Build the control group database required for evidence building
  - Encourage the formation of consortiums in the required disease areas (e.g., mental health) and jointly build the required control group database

Contents

Study the social implementation of healthcare services

Study social implementation systems in the UK, US and Germany

Study the mechanisms/support systems used to evaluate healthcare service quality and utility in the UK, US, and Germany

→

Propose a mechanism for evaluating the quality/utility of healthcare services in Japan

Based on the results of the overseas study, propose a unified system in charge of the social implementation strategy for healthcare services in Japan

→

Create an execution plan for Japan's social implementation system



Stakeholder meeting

Engage healthcare service businesses and exit stakeholders (employers, municipalities, etc.) to discuss the evidence needed for healthcare services and suggest mechanisms for evaluating their quality and utility

Understand the positioning of foreign case examples and the current system in Japan

Standards are mandatory rules, while guidelines are recommended for voluntary compliance

Details follow

	Create/announce standards and guidelines		Authorize
	Quality (security, accessibility, safety, minimum level of scientific rigor)	**Utility** (achieving health-related outcomes)	
UK	**DTAC** • SaMD/Non-SaMD[1] **standards in the UK** • Japan's **Non-SaMD are also covered** **NICE Evidence Standard Framework** • SaMD/Non-SaMD **guidelines in the UK** • Japan's Non-SaMD are also covered **(healthcare apps primarily fall under Tier B)**		**ORCHA** • They are commissioned by the government to evaluate applications and make announcements on the website • Primarily evaluate based on DTAC
US	**APA App Advisor** • Guidelines **for mental healthcare apps in the US (primarily Non-SaMD in the US)** • Japan's **Non-SaMD are also covered**		
Germany	**DiGAV** • SaMD **Standards in Germany** • Japan's Non-SaMD are also covered in the cases of secondary/tertiary prevention		**BfArM** • **Performs screening of quality/utility for** insurance reimbursement
Japan	**Current proposal** • Standards **or/and guidelines for Non-SaMD** • Position utility in the items as evidence		**Current proposal** • Self-declaration • Third-party institution (ORCHA) • Industry association certification • Industry association participation standards

1. Items that can be purchased/recommended by the NHS

DTAC sets minimum quality criteria for NHS organizations when purchasing or recommending items

Background information

Formulator

- **NHS**

Details

- Established in 2021, the NHS App Library Replacement sets **minimum criteria** for healthcare apps when NHS organizations purchase or recommend items
- Introduces 5 criteria for **"quality"**
  - **Clinical Safety**: Clinical risk management activities, organizational risk management systems, etc
  - **Data protection**: Compliance in the collection, storage, and use of personally identifiable data
  - **Technical Safety**: Security including cyber-attack countermeasures, incident management, etc
  - **Inter-operability**: Seamless data linkage system such as API use
  - **Usability and accessibility** (score evaluation): Understanding of needs, presence/absence of acceptance testing, accessibility based on WCAG (Web Content Accessibility Guideline), etc
- Positioned as an **essential criterion to fulfill NICE's "Evidence Standards Framework,"** including "usefulness" guidelines (e.g., validation by RWE)
- ORCHA, an authorizing body, evaluates apps mainly based on DTAC

Implications

- "Quality" criteria are **developed by public authorities**
- **The linkage to investor and evaluation guidelines** provides clarity for service providers on compliance requirements, though the abundance of frameworks is acknowledged as an issue

Source: DTAC, NHS website

NICE provides guidelines for "quality" and "usefulness”

Background information

Formulator

- **NICE (National Institute for Health and Care Excellence)**

Details

- NICE, created in 2018, aimed to serve investors and technology developers in the absence of a unified guideline within the digital healthcare technology industry
- Categorizes digital healthcare by risk and presents items required for each category
  - Tier A: System Services
  - Tier B: Lifestyle improvement applications, etc (applicable to healthcare applications)
  - Tier C: Medical devices such as software to aid diagnosis and predict the risk of serious illness
- Examples of required levels
  - Quality: Sufficiency of data utilization, safety assurance in case of P2P, environmental considerations, etc
  - Usefulness: Results (RWE, cost-efficacy analysis, etc.), process (involvement of experts, etc.)
- Formulation process
  - Core members (NICE, NHS England, Public Health England, MedCity, an academia group) developed a draft
  - 13 workshops were conducted to refine the draft, engaging with public bodies, providers, and academia to discuss content. Subsequently, feedback on usability was gathered from business operators. The draft was then publicly released, feedback was obtained, and the final version was completed
  - Regularly updated, taking into account survey results and technology trends (e.g. machine learning)

Implications

- **Unified framework for "quality" and "usefulness" guidelines** simplifies compliance understanding for service providers
- Developed by **public organizations** with an emphasis on consensus among stakeholders
- **Regularly updated** to align with technological trends and other relevant factors

Source: NICE interviews, NICE Evidence Standard Framework

APA publishes evaluation items (App Advisor) for mental health care apps not covered by FDA

Background

- Many services are Non-SaMD exempt from FDA regulation, or SaMD[1] subject to discretionary decision
- HIPAA and HITECH protect personal information in healthcare, leaving many digital health services uncovered
- Academic societies and universities publish evaluation standards for digital healthcare services via apps to address concerns about evidence-building for efficacy and safety amid the surge in services
- The American Psychological Association (APA) has published mental health app evaluation items for patients and healthcare workers

Evaluation items

Overview	Details (abridgment)
Access & Background	• In-app costs • Availability of linkage with mobile data
Privacy & Safety	• Privacy protection • Safety of data management
Clinical Foundation	• Availability of evidence of usefulness • Availability of medical evidence
Usability	• Availability of functions that meet patient needs • Clarity of functions
Data Integration towards Therapeutic Goal	• Availability of data sharing with healthcare professionals • Availability of function that lead to behavioral change, improvement of self-care, etc

1. SaMD services subject to FDA discretion include those for delivering voice messages and simple behavioral methods to manage anxiety in patients with depression or anxiety disorders, enhancing interest in ergotherapy through videos and video games, and providing support tools and guidance for borderline diabetics to adopt healthier eating and exercise habits

Source: APA website

(2) Study related to social implementation of healthcare service

DiGA sets criteria for "quality" and "usefulness" applicable to some non-SaMD in Japan as well

Background information

Formulator

- Evaluated by **the German Federal Ministry of Health** and *BfArM (German Federal Institute for Drugs and Medical Devices)

Details

- DiGA, initiated by the DVG Act (2019), establishes reimbursement prices within 3 months from application
- Class 1 non-medical devices in Japan are reimbursable if intended for secondary/tertiary prevention and include patients in intended uses
- DiGAV provides criteria for "quality" and "usefulness"
  - **Safety and suitability for use**: Status of CE marking, etc
  - **Data protection**: Compliance with the EU General Data Protection Regulation (GDPR), the German Federal Data Protection Act (BDSG), DiGAV's 40 checklists, etc
  - **Information security:** Information security management, compliance with Federal Office for Information Security guidelines, etc
  - **Inter-operability**: Guarantee of access rights to patient data (e.g., data collected from DiGA can be exported in a readable format)
  - **Other quality**: Robustness, etc
  - **Clinical benefits**: Evidence for improvement of health status, disease duration, survival, and quality of life can be built through retrospective studies

Implications

- **Some non-SaMD in Japan are also subject to this criteria as SaMD in Germany**
  - Class 1 software for secondary/tertiary prevention, with patients as intended users, falls under this criteria as SaMD
  - Mainly in the field of mental health
- **Criteria for "usefulness" exist, allowing evidence building** through both RCTs and retrospective studies
- **Unified framework for "quality" and "usefulness“ standards** simplifies compliance understanding for service providers

Source: DiGA Guidelines

Key question to be answered through this study

Question: How should quality and utility evaluation standards/certification mechanisms be stipulated in order to encourage appropriate evidence building without obstructing the creation of healthcare services?

- What kind of services should be promoted by this study?
- What are the topics and current situation related to the evidence-based social implementation of the target services?
  - Who are the stakeholders who will be involved in the dissemination of the target services?
  - What evidence is required from the viewpoint of service users?
  - What are topics related to the social implementation of evidence-based healthcare services?
  - What are the current countermeasures for these topics?
- With regard to quality and utility evaluation standards, what kinds of standards should be stipulated and what kind of certification system should be built in order to encourage appropriate evidence building without discouraging innovation?

(1) What kind of services should be promoted by this study?

Lifestyle diseases

Disease: Assuming primarily hypertension and diabetes

Target:

- These are services that involve wearable devices or smartphones keeping track of the PHRs of individuals, including basic information (gender, age, etc.), dietary and sleep habits, and health condition (blood sugar and blood pressure levels, etc.), and **that provide multifactorial intervention programs or feedback to encourage appropriate exercise, diet, and sleep habits**
- The primary target is non-pharmacological intervention, and does not include those that perform only monitoring

Prevention area: Assumes that primary prevention will be the main focus, but depending on the service secondary or tertiary prevention could also be included (including specific health guidance services)

Example services: Pep Up.compass, kencom, Linkx, Kenko Try, Mystar, Omada (US), Livongo (US)

These are the targets for prioritized investigation, but are not all of the healthcare services expected to be included in this project

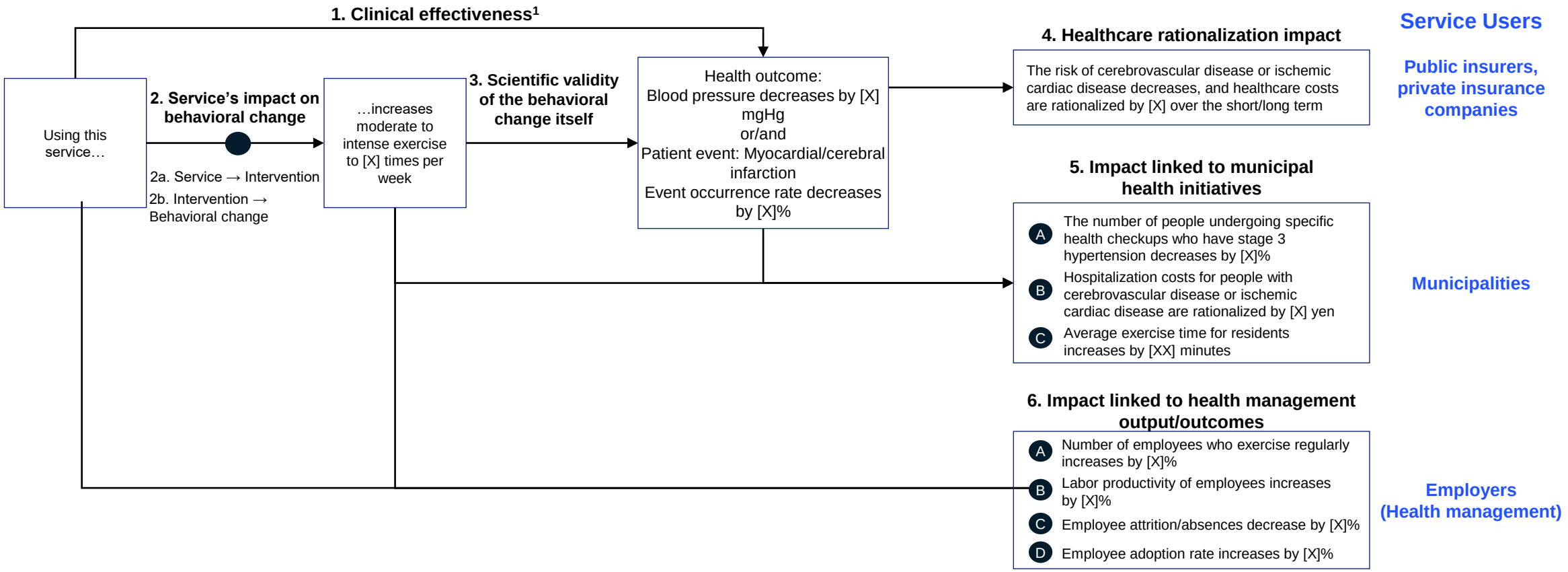
(2) Who are the stakeholders who will be involved in the disseminations of the target services?

Key stakeholders involved in the dissemination of healthcare services in Japan

Stakeholders		Dissemination methods and types of users reachable through service dissemination	Incentives offered to recommend/bear the cost of service	Activities required to promote recommendation/bearing the cost
Employers		Recommending the service as part of employee benefits, and (partially) bearing the cost of it	• Improving productivity or reducing attrition/absences by promoting better employee health (e.g., mental health) • Cultivating the corporate brand/improving adoption rate	• Designing incentives for employers to recommend/bear the cost of service (e.g., Health Management) • Presentation of impact linking the service's health management output with outcomes
Public insurers	**Employee health insurance**	Primarily recommends the service to employees (or public workers in the case of Kyosai) and their dependents, and (partially) bears the cost	• Using effective public health programs to get high ratings via insurer incentives • Rationalizing healthcare costs by promoting prevention by insured people (e.g., preventing the occurrence/worsening of lifestyle diseases, insurance for the very elderly including frailty prevention)	• Designing incentives for insurers to recommend/bear the cost of service (e.g., linking with a database portal site, linking with insurer incentives) • Demonstration of the service's clinical effectiveness and health economic impact
	National Health Insurance	Primarily recommends the service to the self-employed, pensioners, the informally employed, etc. and (partially) bears the cost		
	Insurance for the very elderly	Primarily recommends the service to people age 75 or older, and (partially) bears the cost		
Private insurers		Recommends the service to private insurance policyholders and (partially) bears the cost	• Rationalization of healthcare costs by promoting prevention by policyholders	• Demonstration of the service's health economic impact
Municipalities		Recommends the service to the region's residents and (partially) bears the cost	• Encourages regional revitalization and the creation of a regional brand by promoting the health of the region's residents	• Demonstration of impact linking the service to municipal health initiatives
Doctors (industrial doctors, etc.)		Recommends the service to patients	• Promotes health and prevention by patients	• Demonstration of the service's clinical effectiveness
Patients		The patient themselves selects the service and uses it	• Promotion of health and prevention for the patient themselves	• Demonstration of the benefits of the service to the individual

What kind of evidence is required from the viewpoint of service users?

Types of evidence of service utility (using hypertension as an example)

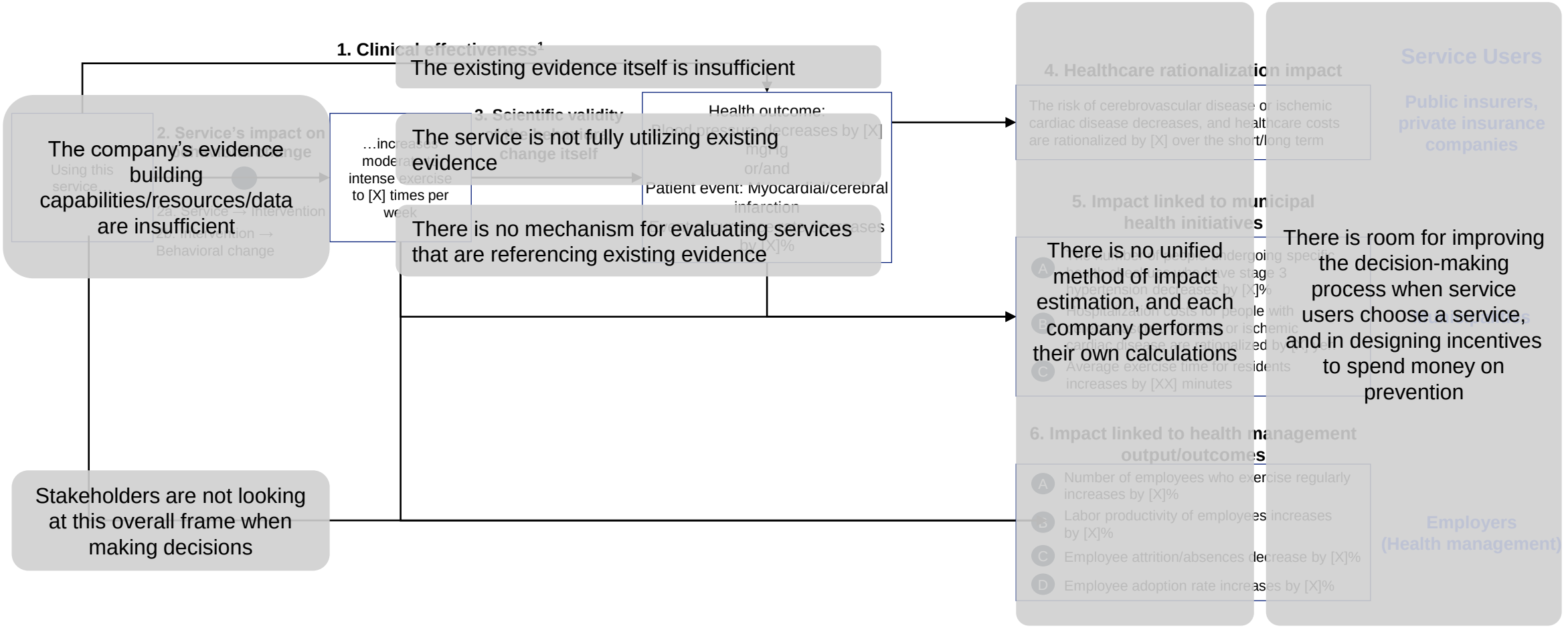


1. Even if a service demonstrates both "2. Service's impact on behavioral change" and "3. Scientific validity of the behavioral change itself," it is not possible to make claims about "1. Clinical effectiveness"

Source: https://www.meti.go.jp/shingikai/mono_info_service/jisedai_health/kenko_toshi/pdf/021_06_00.pdf

What kind of evidence is required from the viewpoint of service users?

Types of evidence of service utility (using hypertension as an example)

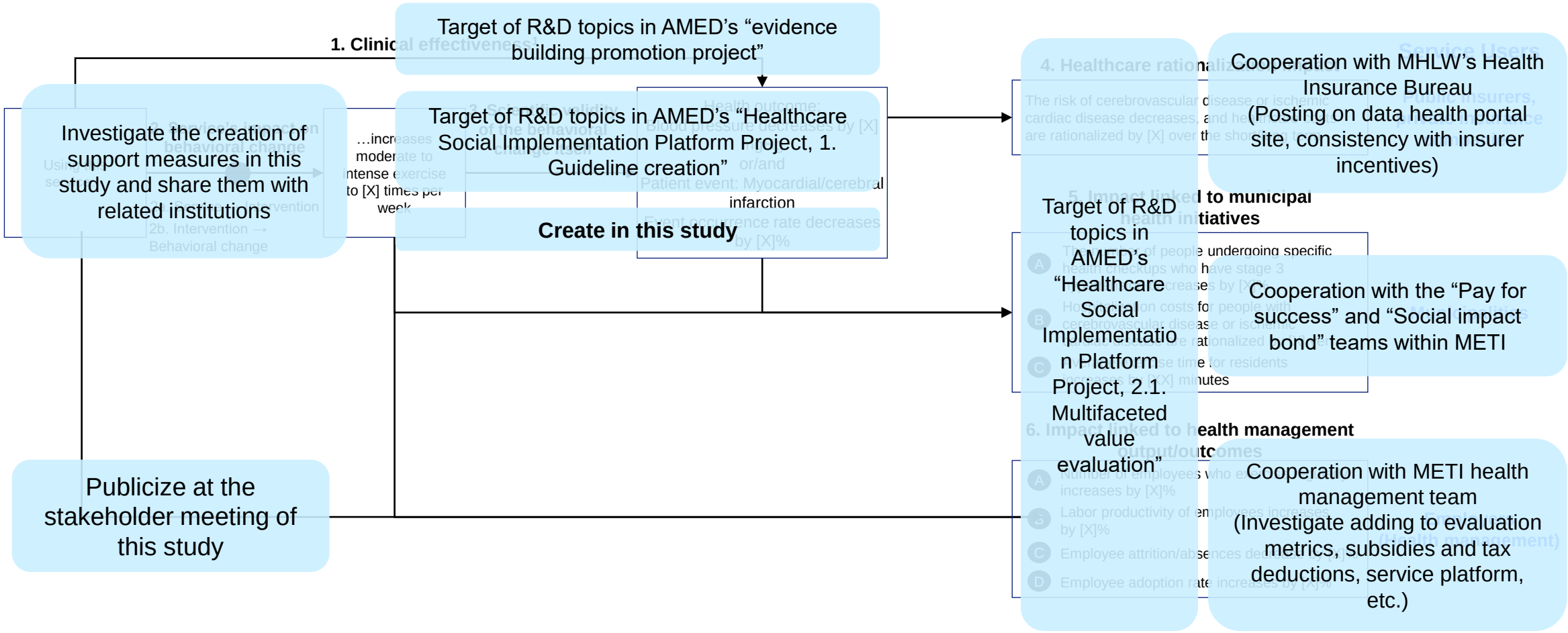


1. Even if a service demonstrates both "2. Service's impact on behavioral change" and "3. Scientific validity of the behavioral change itself," it is not possible to make claims about "1. Clinical effectiveness"

Source: https://www.meti.go.jp/shingikai/mono_info_service/jisedai_health/kenko_toshi/pdf/021_06_00.pdf

What kind of evidence is required from the viewpoint of service users?

Types of evidence of service utility (using hypertension as an example)



1. Even if a service demonstrates both "2. Service's impact on behavioral change" and "3. Scientific validity of the behavioral change itself," it is not possible to make claims about "1. Clinical effectiveness"

Source: https://www.meti.go.jp/shingikai/mono_info_service/jisedai_health/kenko_toshi/pdf/021_06_00.pdf

Details: Current proposal regarding utility standards

Discussion points	Current proposal
How to add conditions to "2. Service's impact on behavioral change"	It is acceptable if any of the following conditions are met • Condition 1: Expert supervision is used when determining the impact validation targets, evaluation metrics, and analysis methodology. The experts are affiliated with a domestic university or domestic research institution in a related field (not limited to the medical disease research area), or they have a certification demonstrating specialization (e.g., they are a certified nutritionist) • Condition 2: The behavioral change impact of the individual service has been published in a peer-reviewed paper The data/research design used in evidence building must be publicly released on the site Notes: NICE ESF requires expert supervision • They are required to be involved in design, development or research, and in determining the appropriateness of the target users • The definition of an expert does not go beyond someone who is a domestic expert in the healthcare field
How to add conditions to "3. Scientific validity of the behavioral change itself"	Indicates that service development is being performed in accordance with the contents of AMED's "prevention/fitness guidelines" In the case of areas not included in AMED's "prevention/fitness guidelines," one of the following conditions must be met • Condition 1: Evidence derived using a certain research design level, such as systematic review/meta-analysis or at least 1 RCT • Condition 2: Evidence included in a peer-reviewed paper • Condition 3: Evidence in a paper published in a top journal

(3) What kind of standards should be stipulated for quality/utility evaluations, and what kind of certification system should be built?

	Service creation	Ensuring service quality/utility	Certification	Dissemination
Support structure encouraging social implementation of healthcare services	**Creation/publication of "prevention/fitness guidelines" by AMED** • Demonstrating "3. Scientific validity of the behavioral change itself" • The service provider can refer to existing evidence when performing service development that focuses on "2. Service's impact on behavioral change"	**Creation/publication of certification standards by the standards creator** • Demonstrating standards related to part of quality and utility (proposed) — Quality: Clinical safety, technical safety, data protection, interoperability, accessibility, usability — Utility: "2. Service's impact on behavioral change" (with conditions) and "3. Scientific validity of the behavioral change itself" (compliance with guidelines) • Because these are not rules, but rather voluntary certification standards, they will not discourage innovation • For businesses that will be undergoing domestic SaMD approval, are planning overseas expansion, or are planning to enter Japan, the domestic SaMD approval standards should be kept as consistent as possible with overseas standards	**Certification mark given by the certifying party** • Proposed: Give certification mark — Certification mark for only meeting quality standards — Mark for meeting quality + "3" — Mark for meeting quality + "2" and "3"	**Publication of the list of certified services by the certifying party** • Publicly release a list of certified services on the website • Users (employers, public insurers, private insurance carriers, municipalities, individuals, etc.) can search for services with guaranteed quality/utility, and can give priority to the ones that have been certified • Each stakeholder is embedded in existing related mechanisms (e.g., data health, health management)
Process of social implementation of healthcare services	**Refer to the "prevention/fitness guidelines" to start service development**	Because the certification mark will encourage service dissemination, they will follow the certification standards to build service evidence that meets part of the quality/utility standards	**They voluntarily apply to the certifying party,** and based on their compliance status with the certification standards, they will be **given the certification mark by the certifying party**, thus obtaining the certification that will encourage dissemination	Uses a dedicated site to make it easier for service users to notice and **encourage dissemination**

Contents

(1) Study related to healthcare service R&D

(2) Study the social implementation of healthcare services

(3) Support the healthcare social implementation platform project

The difference between the USPSTF and the CPSTF is that the USPSTF is geared more towards medical experts, while the CPSTF handles a broader range of general social topics

Differences between USPSTF and CPSTF

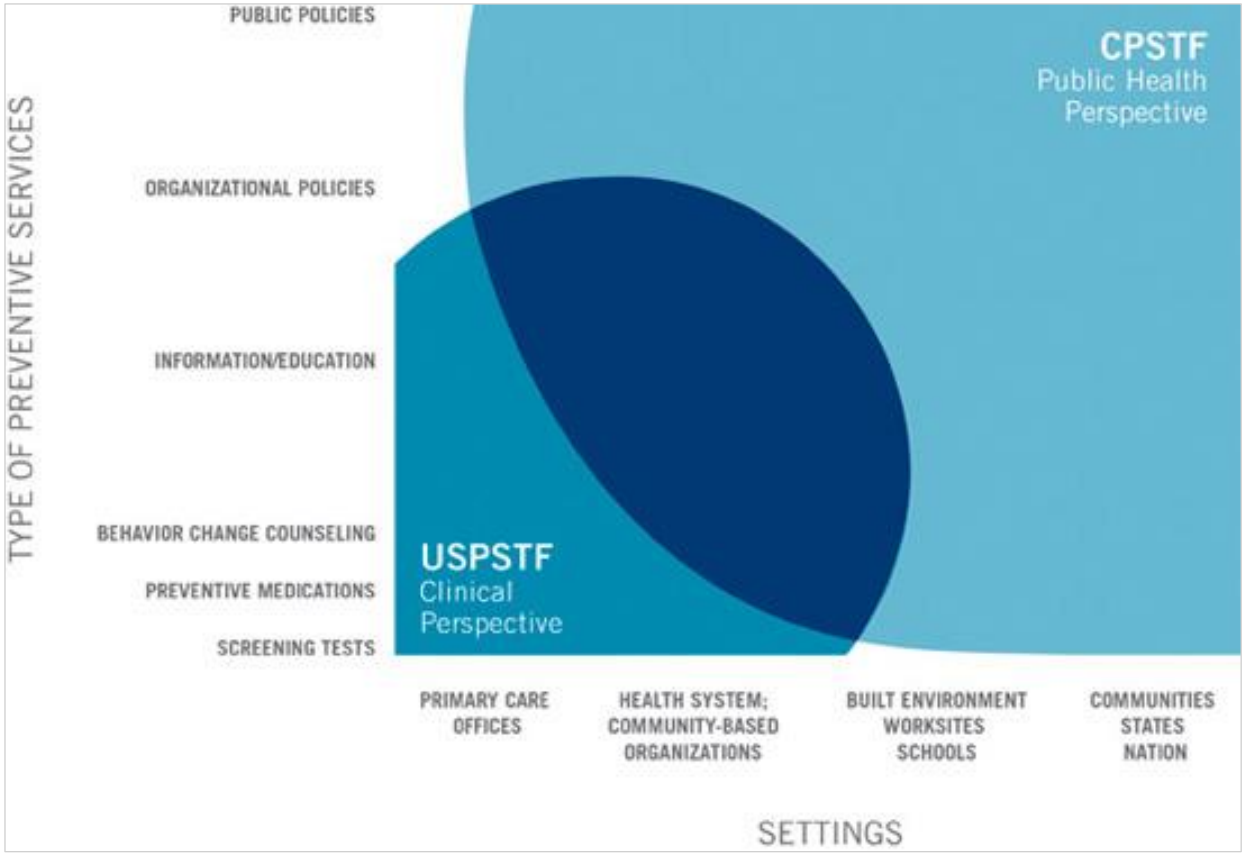
USPSTF and CPSTF

USPSTF provides evidence-based recommendations for healthcare professionals and decision makers on clinical preventive services based on systematic reviews of available evidence. The AHRQ provides administrative, scientific, and technical support

The CPSTF, on the other hand, was established by the DHHS in 1996 to **cover a broader range of public health topics** and complement the work of the USPSTF

Together, CPSTF and USPSTF identify effective, evidence-based preventive interventions. The task forces work together to ensure there is no duplication of effort

The chart on the right shows their fields of activity



Source: CPSTF

The USPSTF is in charge of developing guideline and related activities

Details on the USPSTF and hints for Japan

Organizational setup

What they are

Background and role

Established in 1984 for evidence-based medicine. Independent panel of experts specializing in scientifically-backed medical care recommending preventative services to be used in primary care settings

Current organization

- Independent, volunteer group of 16 national experts in the field of primary care. The panel is led by a chair and 2 vice chairs
- The FY2022 budget was USD 11.5 million (around JPY 15 oku-yen)

Governance

- The AHRQ director appoints members for 4 year terms
- Members are screened for conflicts of interest

Cooperative partnerships

- Various academic societies, medical institutions, medical authorities, and others
- Drafting of guidelines is led by nine Evidence-based Practice Centers (EPCs), universities sponsored by the AHRQ

What they do

Developing guidelines

- The USPSTF does not do its own primary research, instead conducting systematic reviews to develop guidelines
- Task Force guidelines focus exclusively on primaly care services or services that are referable by a primary care clinician, and apply to patients who have no signs or symptoms of the disease or condition
- Guidelines cover preventive service topics for people across the lifespan
- Guidelines are kept as current by routinely updating existing ones and developing new editions

How guidelines are developed

- Prepares processes for developing guidelines
- Policies for alignment with various national standards
- Prepares methodologies for developing guidelines, including ways of assigning grades and identifying research needs and evidence gaps

Key activities

Defining recommendation grades

Grades are assigned based on the strength of evidence for the preventative service and the balance of benefits and harms

Conflicts of interest are disclosed for all members

- All information on financial and non-financial conflicts of interest is disclosed and regularly updated during members' appointment terms
- Conflicts of interest are classified as level 1, 2, or 3. Actions are taken towards those with significant impact (e.g. the member may not participate in the relevant recommendation)

Source: USPSTF

The USPSTF has issued medical professionals over a hundred preventative and other guidelines for a variety of patients, with clear guidelines towards preventative interventions

Overview of USPSTF guidelines

Overview of USPSTF guidelines

What it looks like

All 129 guidelines posted

Can be filtered by multiple parameters

- Status: Published, In progress
- Grade: A, B, C, D, I
- Age: Adolescent, Adult, Pediatric, Senior
- Category: Cancer: (see left)
- Sex/Gender/Pregnancy Status: Female/Woman, Male/Man, Pregnant Persons
- Type of Preventive Service: Counseling, Preventive medication, Screening

Category and numbers

Focus areas

Many guidelines target development and behavior, infectious diseases, and metabolic, nutritional, and endocrine conditions

	Number	%
Cancer	19	12.9
Cardiovascular Disorders (Heart and Vascular Diseases)	15	10.2
Development and Behavior	**18**	**12.2**
Infectious Diseases	**18**	**12.2**
Injury Prevention	3	2.0
Mental Health Conditions and Substance Abuse	14	9.5
Metabolic, Nutritional, and Endocrine Conditions	**20**	**13.6**
Miscellaneous	7	4.8
Musculoskeletal Disorders	4	2.7
Obstetric and Gynecologic Conditions	16	10.9
Perinatal Care	7	4.8
Vision and Hearing Disorders	6	4.1
Total	147	100.0

Note: Total exceeds 147 as some guidelines fall into more than one category

Highlight: grading system

Grades are assigned based on the strength of evidence for the preventative service and the balance of benefits and harms

Grade	Definition	Recommendation
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial	Offer or provide this service
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial	Offer or provide this service
C	The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences. There is at least moderate certainty that the net benefit is small	Offer or provide this service for selected patients depending on individual circumstances
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits	Discourage the use of this service
I	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined	Read the clinical considerations section of USPSTF Guideline Statement. If the service is offered, patients should understand the uncertainty about the balance of benefits and harms

Source: USPSTF, Health Promotion Research Center (Kaori Yoshiba and Masakazu Nakamura. Efficacy of Behavioral Change Counseling in Primary Care--Latest Insights from the US Preventative Services Task Force)

The USPSTF works to systematically develop multiple guidelines, effectively utilizing consistent processes and notation methods as target illnesses and target patients increase

Examples of USPSTF guidelines

Examples

Selection criteria

Filters

1. External standards: 13 hits
   - Published
   - Grade A or B
   - Adult
   - Counseling, preventive intervention
   - Not Pregnant
2. Target patients: Lifestyle diseases and mental health (similar to themes selected by R4' AMED) 12 hits
3. The 3 guidelines on the right were selected after eliminating "out of date" guidelines issued in 2015 or earlier

Selection results and details

Title	Population	Recommendations	Risk Assessment /Screening test	Interventions
Tobacco Smoking Cessation in Adults, Including Pregnant Persons: Interventions (2021)	Adults 18 years or older, including pregnant persons	Provide behavioral interventions and pharmacotherapy for nonpregnant adult smokers Grade A	Ask all adult smokers, including pregnant persons, about their tobacco use "5 A's": Ask, Advise, Assess, Assist, Arrange follow-up "Ask, Advise, Refer" "Vital Sign": Treat smoking status as a vital sign	Provide cessation interventions for smokers Provide behavioral counseling to pregnant smokers
Unhealthy Alcohol Use in Adolescents and Adults: Screening and Behavioral Counseling Interventions (2018)	Adults, including pregnant persons	Screen for unhealthy alcohol use and provide persons engaged in risky or hazardous drinking with brief behavioral counseling interventions Grade B	Numerous brief screening instruments can detect unhealthy alcohol use with acceptable sensitivity and specificity in primary care settings 1-3-item screening instruments have the best accuracy for assessing unhealthy alcohol use in adults 18 years or older These instruments include the Alcohol Use Disorders Identification Test (AUDIT-C) and the SASQ	Brief behavioral counseling interventions were found to reduce unhealthy alcohol use in adults 18 years or older, including pregnant women Effective behavioral counseling interventions vary in their specific components, administration, length, and number of interactions The USPSTF was unable to identify specific intervention characteristics or components that were clearly associated with improved outcomes
Weight Loss to Prevent Obesity-Related Morbidity and Mortality in Adults: Behavioral Interventions (2018)	Adults with a BMI ≥30	Offer or refer to intensive, multicomponent behavioral interventions Grade B	More than 35% of men and 40% of women in the United States have obesity Obesity is associated with health problems such as increased risk for coronary heart disease, type 2 diabetes, various types of cancer, gallstones, and disability Obesity is also associated with an increased risk for death, particularly among adults younger than 65 years	Effective intensive behavioral interventions were designed to help participants achieve or maintain a ≥5% weight loss through a combination of dietary changes and increased physical activity Most interventions lasted for 1 to 2 years, and the majority had ≥12 sessions in the first year Most behavioral interventions focused on problem solving to identify barriers, self-monitoring of weight, peer support, and relapse prevention Interventions also provided tools to support weight loss or weight loss maintenance (e.g. pedometers, food scales, or exercise videos)

Suggestions for Japan

The following may be of use for improving Japan guidelines

How guidelines are created

- All guidelines are formulated according to a consistent process based on established editing rules
- Over 100 guidelines have been issued; version updates are managed with older ones flagged as "out of date"

How details are described

- Risk screening methods indicated
- Intervention methods specified:
  - Group or individual target
  - Timeframe and frequency
  - Target behavioral changes (exercise, diet, etc.) indicated

Source: USPSTF

USPSTF recommendation grades are determined based on a combination of evidence certainty and risk-benefit analysis

How the USPSTF determines grades

Primarily based on levels of certainty

Level of certainty	Definition
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Middle	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: • The number, size, or quality of individual studies • Inconsistency of findings across individual studies • Limited generalizability of findings to routine primary care practice • Lack of coherence in the chain of evidence As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: • The limited number or size of studies • Important flaws in study design or methods • Inconsistency of findings across individual studies • Gaps in the chain of evidence • Findings not generalizable to routine primary care practice • Lack of information on important health outcomes More information may allow estimation of effects on health outcomes

Grades are then assigned in consideration of net benefit

Grades are assigned by the 16 Task Force members

Level of certainty	Magnitude of Net benefit			
	Substantial	**Moderate**	**Small**	**Zero/Negative**
High	Grade A	Grade B	Grade C	Grade D
Middle	Grade A	Grade B	Grade C	Grade D
Low	Insufficient			

Magnitude of Net Benefit

Substantial: Substantial net benefit

Moderate: Moderate net benefit

Small: Small net benefit

Zero/Negative: No net benefit or harms outweigh benefits

Net benefit: Overall efficacy of the preventative service in consideration of potential harms

In estimating “Magnitude of net benefit”, the Task Force gives equal attention to benefits and harms because preventive interventions may result in harms as a direct consequence of the service or for other downstream reasons. Magnitude is calculated by estimating the likelihood of benefits and harms occurring during large RCTs

Note: Details available at the link below

Update on Methods: Estimating Certainty and Magnitude of Net Benefit | United States Preventive Services Taskforce (uspreventiveservicestaskforce.org)

Source: USPSTF

The USPSTF uses a 5-stage process develop guidelines, involving various stakeholders and integrating various review processes

USPSTF Guidelines Development Process

	Step 1 Review topic nominations	Step 2 Develop draft research plan	Step 3 Review public comments and finalize research plan	Step 4 Review evidence and develop draft recommendation	Step 5 Review public comments and finalize recommendation
Development process	1. Anyone can nominate a new topic 2. The USPSTF reviews nominated topics based on the potential impact on prevention and primary care, and importance for public health 3. The USPSTF selects and prioritizes topics for review	1. The USPSTF and EPCs develop a draft research plan 2. The USPSTF posts draft research plan for public comment	1. The USPSTF and the EPCs review all comments 2. The USPSTF posts the final research plan on its website	1. EPCs review evidence on the topic from peer-reviewed studies and draft an evidence review 2. The USPSTF weighs the potential benefits and harms and drafts a recommendation 3. The USPSTF posts the draft recommendation and draft evidence review to its website for public comment	1. The USPSTF and EPCs review all comments and the EPCs finalize the evidence review 2. The USPSTF reviews all comments and finalizes the recommendation 3. The USPSTF posts the final recommendation and final evidence review on its website and publishes them in a peer-reviewed journal
Stakeholders	**16 member USPSTF and support staff from AHRQ**				
	EPCs 9 Evidence-based Practice Centers (EPCs), universities sponsored by the AHRQ. EPCs can be US universities, medical centers, or research institutions. Researchers are medical experts such as medical doctors, pharmacologists, psychotherapists, and physical therapists. External reviewers are asked to participate as needed.				
	Anyone can nominate a topic on the website	4 week public comment period From partner organizations, physicians, researchers, the general public, etc.		4 week public comment period Emails sent out to 60,000 registered mailing list members	Active dissemination • Emails sent out to the list of parties involved in developing the recommendation • Media posts, press conferences • Academic conference presentations, website posts • Recommendation viewing apps and podcasts for physicians
Considerations	• Is the topic of the right scope for the USPSTF? (target group, primary or secondary preventive service, relevant to primary care) • Is the topic appropriate? (relevancy, importance, impact, existence of new evidence)	• Draft research questions, analytical framework, and evidence selection criteria	• Add a section summarizing public comments and note the resulting corrections	• Look at the strength of the evidence and consider net benefits to assign a grade • See if the benefits exceed harms, whether the risks outweigh benefits, and whether the evidence is clear • Assign a grade via 16-member vote	

Source: USPSTF

Task Force members must disclose conflict of interest information and update it during their terms; conflicts are assigned different levels and may prevent participation in developing guidelines

How the Conflict of Interest Disclosures work

Overview

- USPSTF requires members to disclose all information regarding possible financial or non-financial conflicts of interests for topics discussed by the Task Force
- Member candidates must also disclose potential conflicts prior to being selected
- Disclosure information is updated regularly during appointment terms
- All conflict of interest information is reviewed as a criterion for serving as a Task Force chair

Levels and how they are handled

Degree of participation in developing guidelines is adjusted according to COI level. Adjustment results are posted on the USPSTF website to ensure transparency

Level	Definition	Response
1	Non-financial disclosures that are not anticipated to affect the Task Force member's judgment on a topic and smaller financial disclosures	No special action. These disclosures do not limit the Task Force member's participation in the topic process
2	Non-financial disclosures that are not anticipated to affect the Task Force member's judgment on a topic and smaller financial disclosures under USD 1,000	These disclosures do not limit the Task Force member's participation in the topic process
3	Relevant financial disclosures over $1,000 and significant nonfinancial disclosures that may affect the Task Force member's view on the topic	Vary according to the nature of the conflict and may include preventing the member from serving as lead of a topic or on the workgroup of a topic, or preventing the member from taking part in all topic activities

Examples of financial conflicts of interest (regardless of the amount of money)

- Stocks, consulting fees, research grants, license fees, expert committee income, advisory fees, honoraria or travel fees, etc.

Examples of non-financial conflicts of interest

- Public comments or testimony relevant to a specific topic, leadership role on a panel, substantial career efforts in a single topic area, previously published opinions, advocacy, etc.

Suggestions for Japan

- Members are checked for conflicts of interest upon selection, and COIs are regularly updated even after they are appointed
- Every effort is made to avoid conflicts of interests, placing a heavy responsibility on both the USPSTF and the members themselves
- Taking such careful measures to prevent conflicts of interest is likely a reflection of how seriously COIs impact the authenticity of guidelines

Source: USPSTF

CPSTF conducts many activities related to the review of proposed interventions, to identify the social and economic benefits of improving people's health

Details on the CPSTF and hints for Japan

Organizational details

Appearance of the organization

Community Preventive Services Task Force

Founding history

The Department of Health and Human Services (DHHS) was established in 1996. It creates guidelines related to scientific evidence for community-based health promotion methods and disease prevention intervention methods.

Current organization

- An independent, non-federal government committee comprising 15 public health and prevention specialists. It has specialized knowledge on a broad range of scientific, practical, and policy issues related to regional prevention services, public health, health insurance, and disease prevention.
- There are 4 permanent committees that conduct discussions with CPSTF (annual meeting reports, health equity, methods, and prioritization), and CDC staff provide support
- *Organization scale and budgets currently being confirmed

Governance/cooperative relationships

- 15 public health and prevention specialist are appointed by the head of the Centers for Disease Control and Prevention (CDC)
- Guidelines are created by 6-10 members gathered from operating agencies

Description of activities

Creation of guidelines

- Through systematic reviews, provide guidelines and perspectives on programs, services, and other interventions, and provide evidence demonstrating effectiveness and economic impact.

Preparation of guideline creation methods

- Supervise and participate in the review process, and prioritize topic areas and intervention methods to be reviewed
- Promote equity in health
- Create annual reports summarizing the details of operations, identify gaps in research, and provide suggestions on priority fields for future research

Key activities

Public meetings

- Issue verified guidelines that take into account the current status and details of reviews, and hold 3 public meetings each year to study fields for future investigations

Disclosure of conflicts of interest

- CPSTF members disclose potential conflicts of interest in advance

Relationships with other organizations

- Build collaborative relationships with 32 government agencies
- Collaborating agencies communicate the progress of community guide activities as well as the perspectives, concerns, and needs of the organizations and their component members; select and prioritize topics; conduct reviews; proliferate suggestions and results, and provide feedback on gaps between suggestions and needs

Source: CPSTF

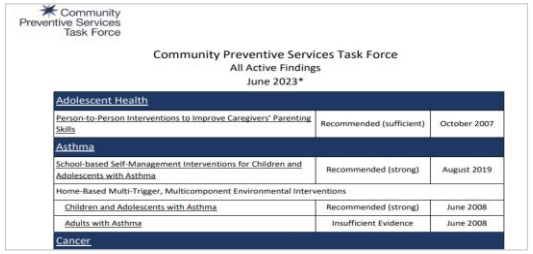
CPSTF provides easy-to-understand descriptions of items clarified through systematic reviews of some 230 interventions, and presents these as guidelines

Overview of CPSTF guidelines

Overview of guidelines

Appearance

All 230 guidelines (interventions) posted


Community Preventive Services Task Force

Community Preventive Services Task Force
All Active Findings
June 2023*

Adolescent Health		
Person-to-Person Interventions to Improve Caregivers' Parenting Skills	Recommended (sufficient)	October 2007
Asthma		
School-based Self-Management Interventions for Children and Adolescents with Asthma	Recommended (strong)	August 2019
Home-Based Multi-Trigger, Multicomponent Environmental Interventions		
Children and Adolescents with Asthma	Recommended (strong)	June 2008
Adults with Asthma	Insufficient Evidence	June 2008
Cancer		

- CPSTF identifies 230 guidelines, and inventories 21 topic fields

To 21 topic fields are: Health during early adolescence, asthma, cancer, diabetes, excessive alcohol ingestion, health communication and health information technologies, prevention of heart disease and strokes, HIV, STIs and teenage pregnancies, mental health, injuries from auto accidents, nutrition, physical activity, obesity, oral health, pregnancy, health related prevention and response, social factors affecting health, cigarettes, vaccines, violence, and workplace health

Category and numbers Topic area	# of interventions	%
Health during early adolescence	1	0%
Asthma	3	1%
Cancer	49	22%
Diabetes	7	3%
Excessive alcohol ingestion	10	4%
Health communication	3	1%
Equity in health	13	6%
Heart diseases	9	4%
HIV and STIs; teenage pregnancy	11	5%
Mental health	8	4%
Injuries from auto accidents	16	7%
Nutrition	5	2%
Physical activities	17	8%
Obesity	19	9%
Oral health	5	2%
Pregnancy	3	1%
Preparations for emergencies	1	0%
Social factors affecting health	-	-
Cigarettes	12	5%
Vaccine	18	8%
Violence	8	4%
Workplace health	5	2%

Highlight: grading system

CPSTF recommendations and results categories

Grade	Number	Grade category
Recommend	**168**	Recommended intervention has strong or sufficient evidence
Not recommended	**2**	If risks are significantly greater than benefits, CPSTF will not recommend, even when strong or sufficient evidence exists
Insufficient evidence	**55**	If there is not enough evidence to identify effectiveness of intervention, or if there is disparity in the evidence, CPSTF handles these cases as "insufficient evidence for intervention" This does not mean that an intervention would not function; only that the impact is unknown.

Strength of evidence is further judged as strong, sufficient, or insufficient

Benefits are also judged according to cost benefits. These cost benefits are judged according to standard values for the costs required to improve the "quality-adjusted life year" (QALY)

1. As of June 2023. Does not include inactive or archived reviews

Source: CPSTF

Inventorying of methods for systematically creating multiple guidelines functions effectively, and multiple DTx guidelines have been created

Examples of CPSTF guidelines

Examples

Selection criteria

Filters

External standards:

- Recommend
- 2010s onwards

Target patients: Lifestyle diseases and mental health (similar to themes selected by R4' AMED)

Interventions related to DTx: 7 cases

From these, select 2 items that are new/easy to visualize

Selection results and details

Title	Recommendation	Audience	Topic	Strategy	Intervention	CPSTF Finding and Rationale Statement
Nutrition and Physical Activity: Digital Health and Telephone Interventions to Increase Healthy Eating and Physical Activity Among Students at Institutions of Higher Education (2021)	Recommended (Sufficient Evidence) *These are only reviews of effectiveness; no economic reviews have been conducted	Adolescents and young adults	Nutrition	Counselling Health education Tech	Digital Health and interventions by phone Websites, mobile apps, text messages, telephone consultations Professional coaching and counseling Self-monitoring of meal contents, exercise, weight, and goal setting Computer generated, but feedback is customized	Rationale behind impact of interventions Eating habits (eating food and vegetables, improving ingestion of fatty foods) Exercise (improving exercise time) Weight (improving BMI and body weight) Inventory the above evidence, and indicate the number of academic papers showing evidence, and median values in the data presented in those papers
Obesity Prevention and Control: Digital Health Interventions for Adolescents with Overweight or Obesity (2019)	Recommended (Sufficient Evidence) *These are only reviews of effectiveness; no economic reviews have been conducted	Adolescents and young adults	Obesity	Counselling Health education Tech	Self-monitoring and goal setting using digital health tools (websites, mobile apps, and wearable devices) Trained staff provided subjects with awareness education on weight management, under the supervision of doctors, counselors, nutritionists, nurses, and pediatricians Subjects self-monitored weight, eating habits, and exercise over 2 months Subjects were also provided with various forms of support (e.g., personal feedback and how to use tools)	Rationale behind impact of interventions Improvements in BMI-z Improvements in body fat percentage BMI improvements Inventory the above evidence, and indicate the number of academic papers showing evidence, and median values in the data presented in those papers

Suggestions for Japan

The following may be of use in improving Japan's guidelines

How guidelines are developed

- All guidelines are formulated according to a consistent process based on established editing rules
- Over 200 guidelines have been issued; version updates are managed with older ones flagged as "inactive"

How details are provided

- Provide clear indication of targets, intervention methods, and evidence discovered
- How to present rationale:
  - Number of academic papers
  - Median value
  - Other additional information

Source: CPSTF

CPSTF guidelines are created based on 2 perspectives: effectiveness and economics. The effectiveness review is created through a 10-step process

Recommendation Development Process

	Step 1 Select topics	Step 2 Build a team	Step 3 Choose intervention methods	Step 4 Finalize concept	Step 5 Research	Step 6 Filter	Step 7 Extract	Step 8 Analyze	Step 9 Recommendation	Step 10 Adoption
Development process	Select topics based on "CPSTF prioritization"	Form teams to develop guidelines; generally 6-10 persons	1. Team creates a comprehensive list 2. CPSTF signs off	1. Define intervention methods 2. Study analysis frameworks 3. Identify outcomes	1. Systematically investigate evidence	Narrow down listed papers	1. Evaluate the quality of papers 2. Gather data from papers	Analyze collected data	Create CPSTF recommendations and findings from evidence investigated	Adopt guidelines
Stakeholders	**15 CPSF members**									
	Public comments are made when "CPSTF prioritization" is created	• Community Guide Program staff: Senior scientists lead and provide support in all aspects of guideline creation • Subject Matter Experts: Participation by persons with detailed knowledge of the topics in question from among academia and persons inside and outside of federal government; e.g., researchers and doctors • CPSTF member: At least 1 participating • CPSTF Lieasons: At least 1 person participates from a related cooperating organization, as a representative of the community using the guidelines in question								Considerable efforts in adoption ✓ Form adoption team ✓ Post on website, create 1 set, and create materials using shared language ✓ After creating the document, circulate among committee members, create a website, post academic papers, and present with cooperating organizations
Considerations	✓ Create "CPSTF prioritization" every 5 years ✓ Select nine topics for the 2020-25 plan ✓ Selection criteria include compatibility, balance, and coverage ✓ CPSTF makes the final decision through deliberations and votes, following discussions with various related parties	✓ Select team members who can represent diverse perspectives	✓ Select methods based on importance, feasibility of prevention, interest expressed by operating organizations, and resources ✓ Possible intervention methods include services for individual patients, behavioral changes in social programs, maintaining environments, and policies	5 approaches to finalizing concepts: 1. Define interventions 2. Incorporation/exclusion criteria 3. Research question 4. Analysis framework 5. Applicability	1. Preliminary survey; e.g., existing systematic review 2. Search document databases (e.g., Chocrane, PubMed)	1. Title screening 2. Abstract screening 3. Text screening 4. Create PRISMA flow diagram	1. Investigate related information 2. Evaluate the quality of papers 3. Evaluate compatibility of survey design in evaluation of intervention impact	✓ Using the analysis framework, analyze impact, benefits, harm, and evidence gaps	Grade from Decision Table 	

Source: CPSTF

Priority topics are decided in 5 years, and are selected based on discussions with a wide range of stakeholders

Selecting topics for guideline creation

Priority topics

- “CPSTF prioritization,” which indicates priority topics, created every 5 years
- Nine topics selected during the 2020-25 period
  - ① Preventing heart disease and strokes
  - ② Preventing harm
  - ③ Mental health
  - ④ Nutrition, physical activity, obesity
  - ⑤ Preparations and responses
  - ⑥ Social determining factors affecting health
  - ⑦ Use of substances
  - ⑧ Use of tobacco
  - ⑨ Violence prevention

How priority topics are selected

1. Start from topics included in "Healthy People 2020" (the national plan on improving health in the United States, created every 10 years)
2. Increase input through discussions with various stakeholders (regular citizens, CDC, CPSTF cooperating organizations (e.g., state and municipal public health agencies, military, federal government agencies, specialist organizations, and national organizations))
3. Narrow down topics based on the following criteria
   - ① **Compatibility**: Degree to which the selected candidate intervention methods are compatible with the activities of the federal government and citizens
   - ② **Balance**: Balance with public health topics overall, and the degree to which evidence gaps can be filled
   - ③ **Burden:** The degree to which those topics reflect a high burden of disease or severity of symptoms
   - ④ **Coverage:** CPSTF's ability to create robust recommendations based on sufficient evidence overall
   - ⑤ **Disparity:** Existence of important health disparity that could be resolved through intervention
   - ⑥ **Impact**: Relationship with survey results and degree of effectiveness in the clinical front lines
   - ⑦ **Potential for prevention:** Degree to which prevention results can be achieved through population-based interventions on this topic
   - ⑧ **Interest expressed by partners:** Priority or degree to which key partners express an interest in the topics
4. Decide priority topics through deliberations and votes

Hints

In topic selection, the CPSTF is aware of "which stakeholders to discuss" and "what perspectives to consider" in deciding topics, and the method of deciding is by voting, which clarifies the "method of selection" itself

Source: CPSTF

Task force member conflict of interest information is disclosed; feasible items and limitations are decided based on the details of those conflicts of interest, and in some cases, those members will not participate in the creation of guidelines

"Conflict of Interest" prevention framework

Overview

- The creation of CPSTF guidelines requires broad trust from citizens, so the reputation of CPSTF members is extremely important
- CPSTF members disclose potential interests in advance
- If any members decline, the chairman announces this at the start of the meeting
- If there are any changes in interest in the case of topics where discussions are ongoing, responses are updated

Levels and how they are handled

The CPSTF Chairman and CDC leaders confirm disclosure details in accordance with the specified standards, and take 1 of the following actions based on the nature and importance of potential conflicts of interest

Action	Description	What is possible	Restrictions
A	No particular conflicts of interest	All	No particular restrictions
B	Depends on details of conflicts of interest	Participation in systematic review teams Discussions and votes regarding specified intervention approaches	No restrictions, but information disclosure is required The details of conflicts of interest must be disclosed to CPSTF members
C	Depends on details of conflicts of interest	Discussions and votes regarding specified intervention approaches	Some restrictions; information disclosure is also required Restrictions on participation in systematic review teams The details of conflicts of interest must be disclosed to CPSTF members
D	Depends on details of conflicts of interest	N1	Substantial restrictions; information disclosure is also required Discussions and votes cannot be conducted with regard to participation on the systematic review team or specified intervention approaches The details of conflicts of interest must be disclosed to CPSTF members

Examples of potential interests

E.g., financial interests, business/employment related interests, intellectual property

Source: CPSTF

Reference: List of abbreviations

Category	Abbreviation	Formal name
Term	CBT	Cognitive Behavior Therapy
Term	DTAC	Digital Technology Assessment Criteria
Term	HIPAA	Health Insurance Portability and Accountability Act
Term	HITECH	The Health Information Technology for Economic and Clinical Health
American agency	APA	American Psychological Association
American agency	CDC	Centers for Disease Control and Prevention
American agency	CMS	Centers for Medicare & Medicaid Services
American agency	CPSTF	Community Preventive Services Task Force
American agency	FDA	U.S. Food and Drug Administration
American agency	USPSTF	U.S. Preventive Services Task Force
British agency	AHSN	Academic Health Science Networks
British agency	NHS	National Health Service
British agency	NICE	National Institute for Health and Care Excellence
German agency	BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
German agency	DiGA	Digitale Gesundheitsanwendungen
German agency	GDPR	General Data Protection Regulation