|  |  |  |
| --- | --- | --- |
|  | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2022-2024 | FG-AI4H-R-028-A01 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | Cambridge, 21-24 March 2023 |
| **DOCUMENT** |
| **Source:** | TG-TM Topic Driver |
| **Title:** | Att.1 – TDD update (TG-TM) |
| **Contact:** | Saketh Ram THRIGULLATG-AI for Traditional MedicineMinistry of Ayush, INDIA | Email: dr.saketram@gmail.com, saketram74@gov.in |

|  |  |
| --- | --- |
| **Abstract:** | This topic description document (TDD) specifies a standardized benchmarking for AI-based Traditional Medicine. It covers all scientific, technical, and administrative aspects relevant for setting up this benchmarking (and follows the template structure defined in document FGAI4H-J-105). The creation of this TDD is an ongoing iterative process until it is approved by the Focus Group on AI for Health (FG-AI4H) as deliverable No. [DEL.10.23] This draft will be a continuous input- and output document. |
| **Change notes:** | **Version 3 (submitted as FG-AI4Health to Meeting R, Cambridge, USA, 21-24 March 2023)**Details of official sharepoint folder, working document updated.Details of Table 1-Topic group output documents updated.Details added for 2.2.1. “Status update for meeting” Details under 2.3. Topic group participation have been updated.3.D. Existing AI solutions - updated**Version 2 (submitted as FG-AI4Health, to Meeting Q, Douala, Cameroon, 6-9 Dec 2022)**Information w.r.t. Abstract, Topic driver has been updated.Topic description has been updated.**Version 1 (submitted as FG-AI4H-P-049 to meeting Plenary in location Helsinki, 20-22 September 2022)** |

Contributors

|  |  |
| --- | --- |
| Saketh Ram THRIGULLA Ministry of AyushINDIA | Tel: +919849995719Email: dr.saketram@gmail.com |

**CONTENTS**

|  Page |
| --- |
| [1 Introduction 5](#_Toc120596569)[2 About the FG-AI4H topic group on Traditional Medicine 5](#_Toc120596570)[2.1 Documentation 6](#_Toc120596571)[2.2 Status of this topic group 6](#_Toc120596572)[2.2.1 Status update for meeting [MEETING LETTER] 7](#_Toc120596573)[2.3 Topic group participation 7](#_Toc120596574)[3 Topic description 7](#_Toc120596575)[3.A. Definition of the AI task 8](#_Toc120596576)[3.B. Current gold standard 8](#_Toc120596577)[3.C. Relevance and impact of an AI solution 8](#_Toc120596578)[3.D. Existing AI solutions 9](#_Toc120596579)[3.1 Subtopic [A] 11](#_Toc120596580)[3.1.1 Definition of the AI task 11](#_Toc120596581)[3.1.2 Current gold standard 11](#_Toc120596582)[3.1.3 Relevance and impact of an AI solution 11](#_Toc120596583)[3.1.4 Existing AI solutions 11](#_Toc120596584)[3.2 Subtopic [B] 12](#_Toc120596585)[4 Ethical considerations 12](#_Toc120596586)[5 Existing work on benchmarking 12](#_Toc120596587)[5.1 Subtopic [A] 12](#_Toc120596588)[5.1.1 Publications on benchmarking systems 12](#_Toc120596589)[5.1.2 Benchmarking by AI developers 13](#_Toc120596590)[5.1.3 Relevant existing benchmarking frameworks 13](#_Toc120596591)[5.2 Subtopic [B] 14](#_Toc120596592)[6 Benchmarking by the topic group 14](#_Toc120596593)[6.1 Subtopic [A] 14](#_Toc120596594)[6.1.1 Benchmarking version [Y] 15](#_Toc120596595)[6.1.2 Benchmarking version [X] 23](#_Toc120596596)[6.2 Subtopic [B] 23](#_Toc120596597)[7 Overall discussion of the benchmarking 23](#_Toc120596598)[8 Regulatory considerations 24](#_Toc120596599)[8.1 Existing applicable regulatory frameworks 24](#_Toc120596600)[8.2 Regulatory features to be reported by benchmarking participants 25](#_Toc120596601)[8.3 Regulatory requirements for the benchmarking systems 25](#_Toc120596602)[8.4 Regulatory approach for the topic group 25](#_Toc120596603)[9 References 26](#_Toc120596604)[Annex A: Glossary 26](#_Toc120596605)[Annex B: Declaration of conflict of interests 28](#_Toc120596606) |

**List of Tables**

|  Page |
| --- |
| [Table 1: Topic group output documents 7](#_Toc55950345) |

**List of Figures**

|  Page |
| --- |
| [Figure 1: Example of a figure 25](#_Toc55950354) |

FG-AI4H Topic Description Document

Topic group TG-AI for Traditional Medicine

# Introduction

This topic description document specifies the standardized benchmarking for Traditional Medicine systems. It serves as deliverable No. [DEL.10.23] of the ITU/WHO Focus Group on AI for Health (FG-AI4H).

# About the FG-AI4H topic group on Traditional Medicine

88% of all countries are estimated to use traditional medicine, such as herbal medicines, acupuncture, yoga, indigenous therapies and others. One hundred seventy Member States report the use of traditional medicine (TM), and their priority request to World Health Organization (WHO) is for evidence and data to inform policies, standards and regulatory frameworks for safe, cost-effective and equitable use. Traditional medicine has been an integral resource for health for centuries in communities around the world, and it is still a mainstay for some with inequities in access to conventional medicine. The sociocultural practice and biodiversity heritages of traditional medicine are invaluable resources to evolve inclusive, diverse sustainable development. Traditional medicine is also part of the growing trillion-dollar global health, wellness, beauty, and pharmaceutical industries. Over 40% of pharmaceutical formulations are based on natural products and landmark drugs, including aspirin and artemisinin, originated from traditional medicine.

Recently, the World Health Organization (WHO) and the Government of India signed an agreement to establish the WHO Global Centre for Traditional Medicine at Jamnagar, India. This global knowledge centre for traditional medicine aims to harness the potential of traditional medicine from across the world through modern science and technology to improve the health of people and the planet. One of the mandates of the centre is to focus on innovation and technology, including an artificial intelligence project which has already started global mapping of patents and research, as well as links to the new WHO Innovation Hub.

Dr Tedros Adhanom Ghebreyesus, DG WHO, during host country agreement signing said that there has been a rapid modernization of the ways traditional medicine is being studied. Artificial intelligence is now used to map evidence and trends in traditional medicine and to screen natural products for pharmacokinetic properties. In addition, functional magnetic resonance imaging is used to study brain activity and the relaxation response that is part of some traditional medicine therapies such as meditation and yoga, which are increasingly drawn on for mental health and wellbeing in stressful times.

WHO’s Traditional Medicine Strategy 2014-23 aims to help countries develop policies and plans to strengthen the role of traditional medicine, based on solid scientific evidence. Where appropriate, countries can integrate traditional medicine into modern health systems, especially at the primary health care level, and to promote its use by individuals and communities in holistic efforts to promote wellbeing.

In recent years, the TM sector has grown in various dimensions and invited considerable amount of global attraction. To cater to new developments and demands, there is an increase in modern technologies like Artificial Intelligence (AI), Decision Support Systems (DSS), in the field of prominent TM diagnosis viz., Tongue pattern diagnosis, pulse diagnosis, Individual constitution analysis, Standardization of Traditional medicine products, etc.,. in various countries across the world. Hence there is a requirement to understand the ground situation and develop benchmarking framework for the AI used in Traditional medicine domain.

To develop this benchmarking framework, FG-AI4H decided to create the TG-AI for Traditional Medicine at the meeting FG-AI4H-P-049 in Helsinki, 20-22 September 2022.

FG-AI4H assigns a *topic driver* to each topic group (similar to a moderator) who coordinates the collaboration of all topic group members on the TDD.During FG-AI4H meeting FG-AI4H-P-049 Helsinki, 20-22 September 2022, Saketh Ram THRIGULLA from Ministry of Ayush, INDIA was nominated as topic driver for the TG-AI for Traditional Medicine.

## Documentation

This document is the TDD for the TG-Traditional Medicine. It introduces the health topic including the AI task, outlines its relevance and the potential impact that the benchmarking will have on the health system and patient outcome, and provides an overview of the existing AI solutions for Traditional Medicine. It describes the existing approaches for assessing the quality of Traditional Medicine systems and provides the details that are likely relevant for setting up a new standardized benchmarking. It specifies the actual benchmarking methods for all subtopics at a level of detail that includes technological and operational implementation. There are individual subsections for all versions of the benchmarking. Finally, it summarizes the results of the topic group’s benchmarking initiative and benchmarking runs. In addition, the TDD addresses ethical and regulatory aspects.

The TDD will be developed cooperatively by all members of the topic group over time and updated TDD iterations are expected to be presented at each FG-AI4H meeting.

The final version of this TDD will be released as deliverable “[DEL.10.23] TG-AI for Traditional Medicine.” The topic group is expected to submit input documents reflecting updates to the work on this deliverable **(Table 1)** to each FG-AI4H meeting.

Table 1: Topic group output documents

| Number | Title |
| --- | --- |
| FGAI4H-Q-029-A01 | Latest update of the Topic Description Document of the TG-AI for Traditional Medicine  |
| FGAI4H-Q-029-A02 | Latest update of the Call for Topic Group Participation (CfTGP) |
| FGAI4H-Q-029-A03 | The presentation summarizing the latest update of the Topic Description Document of the TG-AI for Traditional Medicine |

The working version of this document can be found in the official topic group SharePoint directory.

* [https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/Forms/221206.aspx]

Select the following link:

* [https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/docs/FGAI4H-Q-029-A01.docx]

## Status of this topic group

The following subsections describe the update of the collaboration within the TG-AI for Traditional Medicine for the official focus group meetings.

### Status update for meeting [FG AI4Health R]

* After the announcement of this topic group and subsequent discussions the following works will be attempted and updated in the next focus group meeting. The same are as follows:
	+ Work on the benchmarking software
	+ Progress with data acquisition, annotation, etc.
	+ Overview of the online meetings including links to meeting minutes
		- Four internal meetings have been successfully organized so far:

|  |  |  |  |
| --- | --- | --- | --- |
| Sl.No. | Date | Meeting | Mode |
| 1. | 13,14 December 2022 | Scope and Application of artificial intelligence in Traditional Medicine, India International Centre, New Delhi | Hybrid (Online/Offline) |
| 2. | 2nd February 2023 | AI4Health "Topic Group (TG)-Traditioal Medicine (TG-TM) | Online |
| 3. | 10th March 2023 | AI4Health "Topic Group (TG)-Traditioal Medicine (TG-TM) | Online |

* + Relevant insights from interactions with other working groups or topic groups
	+ Partners joining the topic group
	+ List of current partners

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sl.No** | **Name** | Affiliation | Expertise | Country | Email Id |
|  | Dr Kenji WATANABE | Center for Kampo Medicine Keio University School of Medicine 35 Shinanomachi, Shinjuku-ku Tokyo 160-8582, Japan toyokeio@sc.itc.keio.ac.jp Phone: +81-3-5366-3824 | TM Expert | Japan | watanabekenji@keio.jp |
|  | Antonio Morandi | Ayurveda Point, Milan | Ayu Expert | Italy | dr.morandi@ayurvedicpoint.it |
|  | Roberto Scavelli | Ayurveda Point, Milan | CS Expert | Italy | roberto.scavelli@libero.it |
|  | Jeffrey D. White, MD, | National Cancer Institute | Cancer, Public Health | USA | jeffreyw@mail.nih.gov |
|  | Bala Pesala | ayur ai | AI, Ayu | India | balapesala@gmail.com |
|  | Ajit Kolatkar | Gastro labs, India | Ayu, AI | India | dr.ajit@gastrolab.com |
|  | Cang Shik YIN | Department of Acupuncture, Kyung Hee University, Seoul, 130-701 South Korea. | Korean Medicine | Republic of Korea | acuyin@khu.ac.kr |
|  | Professor Donghyun NAM | Expert TM and AI, College of Korean Medicine, Sangji University, Republic of Korea. | Korean Medicine and AI | Republic of Korea | omdnam@sangji.ac.kr, omdnam@naver.com |
|  | DR Pawan Godatwar | WHO, SEARO, New DELHI | TM Expert, Policy | - | pgodatwar@gmail.com |
|  | DR Kim Sung Chol | WHO, SEARO, New DELHI | TM Expert, Policy | - | kims@who.int |
|  | Dr LIU Qin | Technical Officer, World Health Organization, Integrated Health Services, Geneva, Switzerland | TM Expert, Policy | - | LIU, Qin <liuq@who.int> |
|  | Dr Pradeep Dua | Technical Officer, World Health Organization, Integrated Health Services, Geneva, Switzerland | TM Expert, Policy | - | duadrpradeep@gmail.com  |
|  | Sivarama Prasad Vinjamury | Professor, Research, Southern California Univ. of Health Sciences (SCU) | TM Expert, Academician | USA | vsrprasad18@gmail.com |
|  | Prof Rajani Kanth Vadigepalli | Vice Chair of Research and Professor of Pathology, Anatomy and Cell Biology, Thomas Jefferson | Cell Biology, Bioinformatics | USA |  |
|  | Dr Bhavana Parashar | IGIB, New Delhi | Ayu Expert and AI | India | bhavana.p@igib.res.in |
|  | Rama Jayasundar | AIIMS, New Delhi | Nuclear Physics, Ayu, | India | ramajayasundar@gmail.com |
|  | V Rangamannar | Bengaluru | CS Expert | India | rangamannarv@gmail.com |
|  | Dr V Ramakrishna | Hyderabad | Ayu Expert | India | rasendra22@gmail.com |
|  | Manohar Gundeti | CARI, Mumbai | Ayu Expert | India | gundetipanchakarma@gmail.com |
|  | Kishor Kumar | NIMHANS, Bengaluru | Ayu Expert | India | ayurkishore@gmail.com |
|  | Ram Kumar | Nirog Street, Delhi | Industry, AI | India | ram@nirogstreet.com |
|  | Uddhavesh Sonvane | CDAC, Pune | CS Expert, AI, TM | India | uddhaveshs@gmail.com |
|  | Rajendra Joshi | CDAC, Pune | CS Expert, AI, TM | India | rajendra@cdac.in |
|  | Vinod Jani | CDAC, Pune | CS Expert, AI, TM | India | vinodj@cdac.in |
|  | Shruti Kolugi | CDAC, Pune | CS Expert, AI, TM | India | shrutik@cdac.in |
|  | Gaur Sundar | CDAC, Pune | CS Expert, AI, TM | India | gaurs@cdac.in |
|  | Manihsa Mantri | CDAC, Pune | CS Expert, AI, TM | India | manishar@cdac.in |
|  | Achyut Avinash Patil | CDAC, Pune | CS Expert, AI, TM | India | achyutp@cdac.in |
|  | Usha Rani | NIT, Rourkela | CS Expert, AI, TM | India | usharanik.niimh@gmail.com |
|  | Pratyusha Mantena | NIIMH, Hyderabad | CS Expert, AI, TM | India | hemavallipratyusha@gmail.com |
|  | Hari Prasad | NIIMH, Hyderabad | Medical Records in TM | India | ayush.fortunehari@gmail.com |
|  | Dr Kannan | CCRS, Chennai | TM Expert | India | siddhikanna@gmail.com |
|  | Dr S Natarajan | CCRS, Chennai | TM Expert | India | drnatarajan78@gmail.com |
|  | Dr Shaista Uruz | CCRUM, New Delhi | TM Expert | India | shaistaccrum@gmail.com |
|  | Dr Usama Akram | CCRAS, New Delhi | TM Expert | India | usamakramdr@gmail.com |
|  | Dr Lalitha Sharma | CCRAS, New Delhi | TM Expert | India | lalita.id15@gmail.com |
|  | Dr Rakesh Narayanan | CCRAS, New Delhi | TM Expert | India | drrakeshccras@gmail.com |
|  | Dr Anagha Ranade | CCRAS, New Delhi | TM Expert | India | anagharanade11@gmail.com |
|  | Dr Azeem | CCRAS, New Delhi | TM Expert | India | azeem2905@gmail.com |
|  | Dr Bidhan Mahajan | CCRAS, New Delhi | TM Expert | India | bidhanmahajon@gmail.com |
|  | Dr Geetha Krishnana | TM Expert, Inernational Relations | TM Expert, Policy | India | gopalakrishnag@who.int |
|  | PRASAD PANCHANGAM | SAIGEWARE, Bengaluru, Deep phenotypingplatform for personalized healthcare | CS Expert, AI, TM | India | prasad@saigeware.com |
|  | Meher Vishal Suraj | NIIMH, Hyderabad | CS Expert, AI, TM | India | kmvsreddy7.meher@gmail.com |
|  | Dr Ajit Kakodkar | Gastro Lab Pvt.Ltd, Pune, India | TM expert, entraprenuar | India | ajitkolatkar@gmail.com |

* + Relevant next steps

## Topic group participation

The participation in both, the Focus Group on AI for Health and in a TG is generally open to anyone (with a free ITU account). For this TG, the corresponding ‘Call for TG participation’ (CfTGP) can be found here:

* [https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/tg/tm/CfP-TG-TM.pdf]

Each topic group also has a corresponding subpage on the ITU collaboration site. The subpage for this topic group can be found here:

* [https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/tg/SitePages/TG-TM.aspx]

For participation in this topic group, interested parties can also join the regular online meetings. For all TGs, the link will be the standard ITU-TG ‘zoom’ link:

* <https://itu.zoom.us/my/fgai4h>

All relevant administrative information about FG-AI4H—like upcoming meetings or document deadlines—will be announced via the general FG-AI4H mailing list fgai4h@lists.itu.int.

All TG members should subscribe to this mailing list as part of the registration process for their ITU user account by following the instructions in the ‘Call for Topic Group participation’ and this link:

* <https://itu.int/go/fgai4h/join>

In addition to the general FG-AI4H mailing list, each topic group can create an *individual mailing list:*

* [INSERT YOUR TOPIC GROUP MAILING LIST HERE]

Regular FG-AI4H workshops and meetings proceed about every two months at changing locations around the globe or remotely. More information can be found on the official FG-AI4H website:

* <https://itu.int/go/fgai4h>

# Topic description

This section contains a detailed description and background information of the specific health topic for the benchmarking of AI in Traditional Medicine and how this can help to solve a relevant ‘real-world’ problem.

Topic groups summarize related benchmarking AI subjects to reduce redundancy, leverage synergies, and streamline FG-AI4H meetings. However, in some cases different subtopic groups can be established within one topic group to pursue different topic-specific fields of expertise. The TG-AI for Traditional Medicine currently has no subtopics. Future subtopics for [A] AI for Traditional Medicine Diagnosis, [B] AI for Traditional medicine product evaluation might be introduced.

### 3.A. Definition of the AI task

This section provides a detailed description of the specific task the AI systems of this TG are expected to solve. It is *not* about the benchmarking process (this will be discussed more detailed in chapter 4). This section corresponds to [DEL03](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B7997F2C1-5A1D-4409-B2A0-CBC4E9CE8CDA%7D&file=DEL03.docx&action=default) *“AI requirements specifications*,” which describes the functional, behavioural, and operational aspects of an AI system.

* AI4TM is utilized to replicate the logical understanding applied in traditional medicine diagnostic methods viz., Interview, Physical Examination and other specific diagnostic equipment, techniques to arrive at pre-diagnosis, prodromes, diagnosis, prognosis, determination of transient patterns, steady states viz., individual constitution etc.,
* The AI tasks implemented are viz., classification, prediction, clustering, or segmentation task etc.,
* AI4TM utilizes the big data generated in the form of the text, sensor based data and other relevant parameters.
* The output is intended towards producing objective, reproducible and clinically relevant diagnosis.

### 3.B. Current gold standard

This section provides a description of the established gold standard of the addressed health topic.

* How is the task currently solved without AI?
	+ Currently the TM diagnosis is done on one on basis (whole system approach) involving continuous interaction between the subject and the TM practitioner.
* Do any issues occur with the current gold standard? Does it have limitations?
	+ Many of the parameters utilized in TM diagnosis are predominantly subjective which is the very important limitation.
* Are there any numbers describing the performance of the current state of the art?
	+ Further, discussions on this topic will provide this information.

### 3.C. Relevance and impact of an AI solution

This section addresses the relevance and impact of the AI solution (e.g., on the health system or the patient outcome) and describes how solving the task with AI improves a health issue.

* Why is solving the addressed task with AI relevant?
	+ Subjective parameters and other whole system related data sets utilized in the TM diagnosis can be converted to objective parameters utilizing data analytics and AI and reduce the individual bias.
* Which impact of deploying such systems is expected?
	+ Deploying such system ensures the objective approach in TM diagnosis and democratizes the knowledge system for wider reach. This certainly will have impact on the health system, overall health system cost, life expectancy, or gross domestic product
* Why is benchmarking for this topic important?
	+ Benchmarking this topic provide stakeholders with numbers for decision-making; does it simplify regulation, build trust, or facilitate adoption

### 3.D. Existing AI solutions

This section provides an overview of existing AI solutions for the same health topic that are already in operation. It should contain details of the operations, limitations, robustness, and the scope of the available AI solutions. The details on performance and existing benchmarking procedures will be covered in chapter 6.

1. AI in TM Diagnosis

| Ref # | Intended Use | Target Population | Type of AI used | Input | Performance |
| --- | --- | --- | --- | --- | --- |
| Jung et al. (2019), Yeh et al.’s (2020)[[1]](#endnote-1) | Predict acupoint patterns | Retrospective E.H.R Data  | Multivariate resting-state FCs | medical records based on symptom and disease information | (Exploratory work) |
| Chen et al., 2014; Lin et al., 2015 & Li et al. (2012) [[2]](#endnote-2) | Tongue and Lip Diagnoses | Prospective, Chronic Renal, Liver Disorders; Extended to Healthy patients | Multi-class SVM algorithm & SCM-REF feature selection | color recognition, patterning, and digitization, tongue size, teeth marks | AI Work |
| LeungYeuk-Lan Alice et al. (2021)[[3]](#endnote-3); Liu, S eta al (2018)[[4]](#endnote-4); Xu, L., eta al (2008)[[5]](#endnote-5); Xu, L. S (2007)[[6]](#endnote-6); Yeuk-Lan Alice eta al (2020)[[7]](#endnote-7) | Pulse pattern recognition (TCM) | Healthy volunteers | artificial neural networks (ANNs); pulse-sensing platform (PSP); Fuzzy neural network;  | record and classify arterial human pulses | AI Work |
| A. Joshi (2007)[[8]](#endnote-8);[[9]](#endnote-9) | Pulse pattern recognition (Ayurveda) | Healthy people and others suffering with health issues. | Acquiring radial pulse patterns in analogue, digital format and analysis. | Pre-AI work | Pre-AI work |
| Han et al., 2018; He et al., 2019; Zhou et al., 2021, Fu et al., 2013; Shen et al., 2021; Yang et al., 2019; Zhang et al., 2021  [[10]](#endnote-10) | Prescription decision supporting system using traditional contexts or explore the efficacy of herbal extracts and prescriptions | Non-human pre-clinical studies | In-silico analysis | Work based on existing herbal and other pharmacological databases | (Exploratory work) |
| Feng et al., 2021b; Song et al., 2021).  [[11]](#endnote-11); Zhang, H et al., (2020)[[12]](#endnote-12) | TCM diagnosis and TCM symptom classification | Non-human; Data driven study | Artificial neural network (ANN), data mining, and multivariate analysis |  | AI Work |

1. **AI in understanding human constitution, physiology by TM methods**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ref # | Intended Use | Target Population | Type of AI used | Input | Performance |
| Wallace, R. K. (2020). [[13]](#endnote-13) ; Prashar B et al (2017)[[14]](#endnote-14); Parashar B et al (2008)[[15]](#endnote-15); Bhargav H et al (2021)[[16]](#endnote-16); Renu Singh et al (2017)[[17]](#endnote-17) | Understanding human constitution (Prakriti etc.,) and its relations ship genetic phenotype and other related use cases | Healthy volunteers | Structured questionnaires, Genome sequencing and other related methods of data collection.  | (Pre-AI work) | (Exploratory work) |
| Madan V et al[[18]](#endnote-18), Tiwari P et al[[19]](#endnote-19), Katua D et al[[20]](#endnote-20) | Predicting Ayurveda-based prakriti analysis (phenotypic traits) using ML, Deep Learning | Healthy volunteers | ML, Deep learning | AI Work | AI Work |
| Wayne P et al[[21]](#endnote-21) | A systems biology approach to studying Tai Chi, physiological complexity and healthy aging: Design and rationale of a pragmatic randomized controlled trial | Healthy volunteers | Systems biology approach | (Pre-AI work) | (Exploratory work) |
| Anonymous[[22]](#endnote-22) | Tai Chi, Physiological Complexity, and Healthy Aging - Gait v1.0.2 | Healthy volunteers | Posture and gait analysis | (Pre-AI work) | (Exploratory work) |
| Anonymous[[23]](#endnote-23) | A multi-camera and multimodal dataset for posture and gait analysis v1.0.0 | Healthy volunteers | Posture and gait analysis |  |  |

1. **AI in safety, efficacy studies of TM-Products (Food, Medicine and others)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ref # | Intended Use | Target Population | Type of AI used | Input | Performance |
| Jayasundar R et al. (2020, 2021)[[24]](#endnote-24), Kumar D, Singh A, Jayasundar R et al. (2021)[[25]](#endnote-25) | Detection of taste of medicinal plants |  | Electronic tongue, NMR analysis | (Pre-AI work) | (Exploratory work) |

1. Miscellaneous (TM Standards, Usage, Knowledge, Aptitude, Perception, Policy etc.,)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ref # | Intended Use | Target Population | Type of AI used | Input | Performance |
| Hongmin Chu et al. [[26]](#endnote-26) | The Use of Artificial Intelligence in Complementary and Alternative Medicine: A Systematic Scoping Review | Review | Scoping review | Review | Review |
| Astin J et al[[27]](#endnote-27) | Why Patients Use Alternative Medicine | General | Review | (Pre-AI work) | (Exploratory work) |
| Jansen C[[28]](#endnote-28) | Integrative medicine: Opportunities, challenges and data analytics-based solutions for traditional medicine | General | Review | (Pre-AI work) | (Exploratory work) |
| Anonymous[[29]](#endnote-29), Ammon K[[30]](#endnote-30), Bornhöft G et Al[[31]](#endnote-31) | Health Technology Assessment (HTA) and TM | General | HTA | (Pre-AI work) | (Exploratory work) |
| Rudra S et al[[32]](#endnote-32) | Utilization of alternative systems of medicine as health care services in India: Evidence on AYUSH care from NSS 2014 | General | Review | (Pre-AI work) | (Exploratory work) |
| Saketh Ram et al[[33]](#endnote-33), Anonymous[[34]](#endnote-34) | NAMASTE Portal, Glossary of Ayurveda Terminologies (BIS) | General | Colloborative, Consultative work | (Pre-AI work) | (Exploratory work) |
| Anonymous | WHO benchmark documents for training and practice of Ayurveda, Unani, Acupuncture, Naturopathy[[35]](#endnote-35), | General | Colloborative, Consultative work | (Pre-AI work) | (Exploratory work) |
| Anonymous | WHO international standard terminologies on Ayurveda[[36]](#endnote-36), Siddha[[37]](#endnote-37), Unani[[38]](#endnote-38), Traditional Chinese Medicine[[39]](#endnote-39) | General | Colloborative, Consultative work | (Pre-AI work) | (Exploratory work) |

## Subtopic [A]

### Definition of the AI task

This section provides a detailed description of the specific task the AI systems of this TG are expected to solve. It is *not* about the benchmarking process (this will be discussed more detailed in chapter 4). This section corresponds to [DEL03](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B7997F2C1-5A1D-4409-B2A0-CBC4E9CE8CDA%7D&file=DEL03.docx&action=default) *“AI requirements specifications*,” which describes the functional, behavioural, and operational aspects of an AI system.

* What is the AI doing?
* What kind of AI task is implemented (e.g., classification, prediction, clustering, or segmentation task)?
* Which input data are fed into the AI model?
* Which output is generated?

### Current gold standard

This section provides a description of the established gold standard of the addressed health topic.

* How is the task currently solved without AI?
* Do any issues occur with the current gold standard? Does it have limitations?
* Are there any numbers describing the performance of the current state of the art?

### Relevance and impact of an AI solution

This section addresses the relevance and impact of the AI solution (e.g., on the health system or the patient outcome) and describes how solving the task with AI improves a health issue.

* Why is solving the addressed task with AI relevant?
* Which impact of deploying such systems is expected (e.g., impact on the health system, overall health system cost, life expectancy, or gross domestic product)?
* Why is benchmarking for this topic important (e.g., does it provide stakeholders with numbers for decision-making; does it simplify regulation, build trust, or facilitate adoption)?

### Existing AI solutions

This section provides an overview of existing AI solutions for the same health topic that are already in operation. It should contain details of the operations, limitations, robustness, and the scope of the available AI solutions. The details on performance and existing benchmarking procedures will be covered in chapter 6.

* Description of the general status and the maturity of AI systems for the health topic of your TG (e.g., exclusively prototypes, applications, and validated medical devices)
* Which are the currently known AI systems and their inputs, outputs, key features, target user groups, and intended use (if not discussed before)? This can also be provided as a table.
* What are the common features found in most AI solutions that might be benchmarked?
* What are the relevant metadata dimensions characterizing the AI systems in this field and with relevance for reporting (e.g., systems supporting offline functions, availability in certain languages, and the capability to process data in a specific format)?
* Description of existing AI systems and their scope, robustness, and other dimensions.

## Subtopic [B]

# Ethical considerations

The rapidly evolving field of AI and digital technology in the fields of medicine and public health raises a number of ethical, legal, and social concerns that have to be considered in this context. They are discussed in deliverable [DEL01](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B0505B020-362C-45B2-94BF-215D2EBBD8F5%7D&file=DEL01.docx&action=default) “*AI4H ethics considerations,”* which was developed by the working group on “Ethical considerations on AI4H” (WG-Ethics). This section refers to DEL01 and should reflect the ethical considerations of the TG-AI for Traditional Medicine.

==Discussion initiated==

* What are the ethical implications of applying the AI model in real-world scenarios?
* What are the ethical implications of introducing benchmarking (having the benchmarking in place itself has some ethical risks; e.g., if the test data are not representative for a use case, the data might create the illusion of safety and put people at risk)?
* What are the ethical implications of collecting the data for benchmarking (e.g., how is misuse of data addressed, is there the need for an ethics board approval for clinical data, is there the need for consent management for sharing patient data, and what are the considerations about data ownership/data custodianship)?
* What risks face individuals and society if the benchmarking is wrong, biased, or inconsistent with reality on the ground?
* How is the privacy of personal health information protected (e.g., in light of longer data retention for documentation, data deletion requests from users, and the need for an informed consent of the patients to use data)?
* How is ensured that benchmarking data are representative and that an AI offers the same performance and fairness (e.g., can the same performance in high, low-, and middle-income countries be guaranteed; are differences in race, sex, and minority ethnic populations captured; are considerations about biases, when implementing the same AI application in a different context included; is there a review and clearance of ‘inclusion and exclusion criteria’ for test data)?
* What are your experiences and learnings from addressing ethics in your TG?

# Existing work on benchmarking

This section focuses on the existing benchmarking processes in the context of AI and Traditional Medicine for quality assessment. It addresses different aspects of the existing work on benchmarking of AI systems (e.g., relevant scientific publications, benchmarking frameworks, scores and metrics, and clinical evaluation attempts). The goal is to collect all relevant learnings from previous benchmarking that could help to implement the benchmarking process in this topic group.

==Discussion initiated==

## Subtopic [A]

### Publications on benchmarking systems

While a representative comparable benchmarking for Traditional Medicine does not yet exist, some work has been done in the scientific community assessing the performance of such systems. This section summarizes insights from the most relevant publications on this topic. It covers parts of the deliverable DEL07 *“AI for health evaluation considerations,”* [DEL07\_1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B565EEC0A-D755-41C8-AC68-37B4C38C953F%7D&file=DEL07_1.docx&action=default) *“AI4H evaluation process description,”* [DEL07\_2](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B58679341-C738-40F0-A822-3AC2B24DD09F%7D&file=DEL07_2.docx&action=default) *“AI technical test specification*,*”* [DEL07\_3](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BA3088882-F82B-493B-B1C5-49CFF0EEEFA8%7D&file=DEL07_3.docx&action=default) *“Data and artificial intelligence assessment methods (DAISAM),”* and [DEL07\_4](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BB846B260-373A-41FC-A892-EE5BBCFE3CF8%7D&file=DEL07_4.docx&action=default) *“Clinical Evaluation of AI for health”*.

* What is the most relevant peer-reviewed scientific publications on benchmarking or objectively measuring the performance of systems in your topic?
* State what are the most relevant approaches used in literature?
* Which scores and metrics have been used?
* How were test data collected?
* How did the AI system perform and how did it compare the current gold standard? Is the performance of the AI system equal across less represented groups? Can it be compared to other systems with a similar benchmarking performance and the same clinically meaningful endpoint (addressing comparative efficacy)?
* How can the utility of the AI system be evaluated in a real-life clinical environment (also considering specific requirements, e.g., in a low- and middle-income country setting)?
* Have there been clinical evaluation attempts (e.g., internal and external validation processes) and considerations about the use in trial settings?
* What are the most relevant gaps in the literature (what is missing concerning AI benchmarking)?

### Benchmarking by AI developers

All developers of AI solutions for Traditional Medicine implemented internal benchmarking systems for assessing the performance. This section will outline the insights and learnings from this work of relevance for benchmarking in this topic group.

* What are the most relevant learnings from the benchmarking by AI developers in this field (e.g., ask the members of your topic group what they want to share on their benchmarking experiences)?
* Which scores and metrics have been used?
* How did they approach the acquisition of test data?

### Relevant existing benchmarking frameworks

Triggered by the hype around AI, recent years have seen the development of a variety of benchmarking platforms where AIs can compete for the best performance on a determined dataset. Given the high complexity of implementing a new benchmarking platform, the preferred solution is to use an established one. This section reflects on the different existing options that are relevant for this topic group and includes considerations of using the assessment platform that is currently developed by FG-AI4H and presented by deliverable [DEL07\_5](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B8BFCFF21-3908-4BAD-AB9C-9814EB3F9B36%7D&file=DEL07_5.docx&action=default) *“FG-AI4H assessment platform”* (the deliverable explores options for implementing an assessment platform that can be used to evaluate AI for health for the different topic groups).

* Which benchmarking platforms could be used for this topic group (e.g., EvalAI, AIcrowd, Kaggle, and CodaLab)?
* Are the benchmarking assessment platforms discussed, used, or endorsed by FG-AI4H an option?
* Are there important features in this topic group that require special attention?
* Is the reporting flexible enough to answer the questions stakeholders want to get answered by the benchmarking?
* What are the relative advantages and disadvantages of these diverse solutions?

## Subtopic [B]

# Benchmarking by the topic group

This section describes all technical and operational details regarding the benchmarking process for the Traditional Medicine AI task including subsections for each version of the benchmarking that is iteratively improved over time.

It reflects the considerations of various deliverables: [DEL05](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B2012357A-941E-44BD-B965-370D7829F52C%7D&file=DEL05.docx&action=default) *“Data specification”* (introduction to deliverables 5.1-5.6), [DEL05\_1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B19830259-F63B-42D4-A408-48C854D6C124%7D&file=DEL05_1.docx&action=default)*“Data requirements”* (which lists acceptance criteria for data submitted to FG-AI4H and states the governing principles and rules), [DEL05\_2](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B25141F77-E59A-45F1-B081-185C2194FE67%7D&file=DEL05_2.docx&action=default) *“Data acquisition”*, [DEL05\_3](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B05D8938E-BC2A-4A62-BCB0-1FD46AA72235%7D&file=DEL05_3.docx&action=default) *“Data annotation specification”*, [DEL05\_4](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BF267A95C-4C5B-4D63-A135-58AF487C3AD3%7D&file=DEL05_4.docx&action=default) *“Training and test data specification”* (which provides a systematic way of preparing technical requirement specifications for datasets used in training and testing of AI models), [DEL05\_5](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B71FE8B9D-ACB3-48CE-AA3F-136409B550A4%7D&file=DEL05_5.docx&action=default) *“Data handling”* (which outlines how data will be handled once they are accepted), [DEL05\_6](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B5C95327E-96A5-4175-999E-3EDB3ED147C3%7D&file=DEL05_6.docx&action=default) *“Data sharing practices”* (which provides an overview of the existing best practices for sharing health-related data based on distributed and federated environments, including the requirement to enable secure data sharing and addressing issues of data governance), [DEL06](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BF5967277-90C8-4252-A0B9-43A5692F35E2%7D&file=DEL06.docx&action=default) *“AI training best practices specification”* (which reviews best practices for proper AI model training and guidelines for model reporting), [DEL07](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B47E77197-F87B-49F4-80B3-2DD949A5F185%7D&file=DEL07.docx&action=default)*“AI for health evaluation considerations”* (which discusses the validation and evaluation of AI for health models, and considers requirements for a benchmarking platform), [DEL07\_1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B565EEC0A-D755-41C8-AC68-37B4C38C953F%7D&file=DEL07_1.docx&action=default) *“AI4H evaluation process description”* (which provides an overview of the state of the art of AI evaluation principles and methods and serves as an initiator for the evaluation process of AI for health), [DEL07\_2](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B58679341-C738-40F0-A822-3AC2B24DD09F%7D&file=DEL07_2.docx&action=default) *“AI technical test specification”* (which specifies how an AI can and should be tested *in silico*), [DEL07\_3](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BA3088882-F82B-493B-B1C5-49CFF0EEEFA8%7D&file=DEL07_3.docx&action=default) *“Data and artificial intelligence assessment methods (DAISAM)”* (which provides the reference collection of WG-DAISAM on assessment methods of data and AI quality evaluation), [DEL07\_4](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BB846B260-373A-41FC-A892-EE5BBCFE3CF8%7D&file=DEL07_4.docx&action=default)*“Clinical Evaluation of AI for health”* (which outlines the current best practices and outstanding issues related to clinical evaluation of AI models for health), [DEL07\_5](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B8BFCFF21-3908-4BAD-AB9C-9814EB3F9B36%7D&file=DEL07_5.docx&action=default) *“FG-AI4H assessment platform”* (which explores assessment platform options that can be used to evaluate AI for health for the different topic groups), [DEL09](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B3E940987-8D75-44B8-85E4-F0E475964F15%7D&file=DEL09.docx&action=default) *“AI for health applications and platforms”* (which introduces specific considerations of the benchmarking of mobile- and cloud-based AI applications in health), [DEL09\_1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B1A2EC8D5-53CA-4C8C-9B09-B61CA6F428C5%7D&file=DEL09_1.docx&action=default) *“Mobile based AI applications,”* and [DEL09\_2](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B3B5A31DE-D3B1-4EC1-A261-2C2E19F73810%7D&file=DEL09_2.docx&action=default) *“Cloud-based AI applications”* (which describe specific requirements for the development, testing and benchmarking of mobile- and cloud-based AI applications).

==Discussion initiated==

## Subtopic [A]

The benchmarking of Traditional Medicine is going to be developed and improved continuously to reflect new features of AI systems or changed requirements for benchmarking. This section outlines all benchmarking versions that have been implemented thus far and the rationale behind them. It serves as an introduction to the subsequent sections, where the actual benchmarking methodology for each version will be described.

* Which benchmarking iterations have been implemented thus far?
* What important new features are introduced with each iteration?
* What are the next planned iterations and which features are they going to add?

### Benchmarking version [Y]

This section includes all technological and operational details of the benchmarking process for the benchmarking version [Y] (latest version, chronologically reversed order).

#### Overview

This section provides an overview of the key aspects of this benchmarking iteration, version [Y].

* What is the overall scope of this benchmarking iteration (e.g., performing a first benchmarking, adding benchmarking for multi-morbidity, or introducing synthetic-data-based robustness scoring)?
* What features have been added to the benchmarking in this iteration?

#### Benchmarking methods

This section provides details about the methods of the benchmarking version [Y]. It contains detailed information about the benchmarking system architecture, the dataflow and the software for the benchmarking process (e.g., test scenarios, data sources, and legalities).

##### Benchmarking system architecture

This section covers the architecture of the benchmarking system. For well-known systems, an overview and reference to the manufacturer of the platform is sufficient. If the platform was developed by the topic group, a more detailed description of the system architecture is required.

* How does the architecture look?
* What are the most relevant components and what are they doing?
* How do the components interact on a high level?
* What underlying technologies and frameworks have been used?
* How does the hosted AI model get the required environment to execute correctly? What is the technology used (e.g., Docker/Kubernetes)?

##### Benchmarking system dataflow

This section describes the dataflow throughout the benchmarking architecture.

* How do benchmarking data access the system?
* Where and how (data format) are the data, the responses, and reports of the system stored?
* How are the inputs and the expected outputs separated?
* How are the data sent to the AI systems?
* Are the data entries versioned?
* How does the lifecycle for the data look?

##### Safe and secure system operation and hosting

This section addresses security considerations about the storage and hosting of data (benchmarking results and reports) and safety precautions for data manipulation, data leakage, or data loss.

In the case of a manufactured data source (vs. self-generated data), it is possible to refer to the manufacturer’s prescriptions.

* Based on the architecture, where is the benchmarking vulnerable to risk and how have these risks been mitigated (e.g., did you use a threat modelling approach)? A discussion could include:
* Could someone access the benchmarking data before the actual benchmarking process to gain an advantage?
* What safety control measures were taken to manage risks to the operating environment?
* Could someone have changed the AI results stored in the database (your own and/or that of competitors)?
* Could someone attack the connection between the benchmarking and the AI (e.g., to make the benchmarking result look worse)?
* How is the hosting system itself protected against attacks?
* How are the data protected against data loss (e.g., what is the backup strategy)?
* What mechanisms are in place to ensure that proprietary AI models, algorithms and trade-secrets of benchmarking participants are fully protected?
* How is it ensured that the correct version of the benchmarking software and the AIs are tested?
* How are automatic updates conducted (e.g., of the operating system)?
* How and where is the benchmarking hosted and who has access to the system and the data (e.g., virtual machines, storage, and computing resources, configurational settings)?
* How is the system’s stability monitored during benchmarking and how are attacks or issues detected?
* How are issues (e.g., with a certain AI) documented or logged?
* In case of offline benchmarking, how are the submitted AIs protected against leakage of intellectual property?

##### Benchmarking process

This section describes how the benchmarking looks from the registration of participants, through the execution and resolution of conflicts, to the final publication of the results.

* How are new benchmarking iterations scheduled (e.g., on demand or quarterly)?
* How do possible participants learn about an upcoming benchmarking?
* How can one apply for participation?
* What information and metadata do participants have to provide (e.g., AI autonomy level assignment (IMDRF), certifications, AI/machine learning technology used, company size, company location)?
* Are there any contracts or legal documents to be signed?
* Are there inclusion or exclusion criteria to be considered?
* How do participants learn about the interface they will implement for the benchmarking (e.g., input and output format specification and application program interface endpoint specification)?
* How can participants test their interface (e.g., is there a test dataset in case of file-based offline benchmarking or are there tools for dry runs with synthetic data cloud-hosted application program interface endpoints)?
* Who is going to execute the benchmarking and how is it ensured that there are no conflicts of interest?
* If there are problems with an AI, how are problems resolved (e.g., are participants informed offline that their AI fails to allow them to update their AI until it works? Or, for online benchmarking, is the benchmarking paused? Are there timeouts?)?
* How and when will the results be published (e.g., always or anonymized unless there is consent)? With or without seeing the results first? Is there an interactive drill-down tool or a static leader board? Is there a mechanism to only share the results with stakeholders approved by the AI provider as in a credit check scenario?
* In case of online benchmarking, are the benchmarking data published after the benchmarking? Is there a mechanism for collecting feedback or complaints about the data? Is there a mechanism of how the results are updated if an error was found in the benchmarking data?

#### AI input data structure for the benchmarking

This section describes the input data provided to the AI solutions as part of the benchmarking of Traditional Medicine. It covers the details of the data format and coding at the level of detail needed to submit an AI for benchmarking. This is the only TDD section addressing this topic. Therefore, the description needs to be complete and precise. This section does *not* contain the encoding of the labels for the expected outcomes. It is only about the data the AI system will see as part of the benchmarking.

* What are the general data types that are fed in the AI model?
* How exactly are they encoded? For instance, discuss:
	+ The exact data format with all fields and metadata (including examples or links to examples)
	+ Ontologies and terminologies
	+ Resolution and data value ranges (e.g., sizes, resolutions, and compressions)
	+ Data size and data dimensionality

#### AI output data structure

Similar to the input data structure for the benchmarking, this section describes the output data the AI systems are expected to generate in response to the input data. It covers the details of the data format, coding, and error handling at the level of detail needed for an AI to participate in the benchmarking.

* What are the general data output types returned by the AI and what is the nature of the output (e.g., classification, detection, segmentation, or prediction)?
	+ How exactly are they encoded? Discuss points like:
		- The exact data format with all fields and metadata (including examples or links to examples)
		- Ontologies and terminologies
* What types of errors should the AI generate if something is defective?

#### Test data label/annotation structure *.*

While the AI systems can only receive the input data described in the previous sections, the benchmarking system needs to know the expected correct answer (sometimes called ‘labels’) for each element of the input data so that it can compare the expected AI output with the actual one. Since this is only needed for benchmarking, it is encoded separately. The details are described in the following section.

* What are the general label types (e.g., expected results, acceptable results, correct results, and impossible results)?
* How exactly are they encoded? Discuss points like:
	+ The exact data format with all fields and metadata (including examples or links to examples)
	+ Ontologies and terminologies
* How are additional metadata about labelling encoded (e.g., author, data, pre-reviewing details, dates, and tools)?
* How and where are the labels embedded in the input data set (including an example; e.g., are there separate files or is it an embedded section in the input data that is removed before sending to the AI)?

#### Scores and metrics

Scores and metrics are at the core of the benchmarking. This section describes the scores and metrics used to measure the performance, robustness, and general characteristics of the submitted AI systems.

* Who are the stakeholders and what decisions should be supported by the scores and metrics of the benchmarking?
* What general criteria have been applied for selecting scores and metrics?
* What scores and metrics have been chosen/defined for robustness?
* What scores and metrics have been chosen/defined for medical performance?
* What scores and metrics have been chosen/defined for non-medical performance?
	+ Metrics for technical performance tracking (e.g., monitoring and reporting when the performance accuracy of the model drops below a predefined threshold level as a function of time; computational efficiency rating, response times, memory consumption)
* What scores and metrics have been chosen/defined for model explainability?
* Describe for each aspect
	+ The exact definition/formula of the score based on the labels and the AI output data structures defined in the previous sections and how they are aggregated/accumulated over the whole dataset (e.g., for a single test set entry, the result might be the probability of the expected correct class which is then aggregated to the average probability of the correct class)
	+ Does it use some kind of approach for correcting dataset bias (e.g., the test dataset usually has a different distribution compared to the distribution of a condition in a real-world scenario. For estimating the real-world performance, metrics need to compensate this difference.)
	+ What are the origins of these scores and metrics?
	+ Why were they chosen?
	+ What are the known advantages and disadvantages?
	+ How easily can the results be compared between or among AI solutions?
	+ Can the results from benchmarking iterations be easily compared or does it depend too much on the dataset (e.g., how reproducible are the results)?
* How does this consider the general guidance of WG-DAISAM in [DEL07\_3](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BA3088882-F82B-493B-B1C5-49CFF0EEEFA8%7D&file=DEL07_3.docx&action=default) “Data and artificial intelligence assessment methods (DAISAM)”?
* Have there been any relevant changes compared to previous benchmarking iterations? If so, why?

#### Test dataset acquisition

Test dataset acquisition includes a detailed description of the test dataset for the AI model and, in particular, its benchmarking procedure including quality control of the dataset, control mechanisms, data sources, and storage.

* How does the overall dataset acquisition and annotation process look?
* How have the data been collected/generated (e.g., external sources vs. a process organized by the TG)?
* Have the design goals for the benchmarking dataset been reached (e.g., please provide a discussion of the necessary size of the test dataset for relevant benchmarking results, statistical significance, and representativeness)?
* How was the dataset documented and which metadata were collected?
	+ Where were the data acquired?
	+ Were they collected in an ethical-conform way?
	+ Which legal status exists (e.g., intellectual property, licenses, copyright, privacy laws, patient consent, and confidentiality)?
	+ Do the data contain ‘sensitive information’ (e.g., socially, politically, or culturally sensitive information; personal identifiable information)? Are the data sufficiently anonymized?
	+ What kind of data anonymization or deidentification has been applied?
	+ Are the data self-contained (i.e., independent from externally linked datasets)?
	+ How is the bias of the dataset documented (e.g., sampling or measurement bias, representation bias, or practitioner/labelling bias)?
	+ What addition metadata were collected (e.g., for a subsequent detailed analysis that compares the performance on old cases with new cases)? How was the risk of benchmarking participants accessing the data?
* Have any scores, metrics, or tests been used to assess the quality of the dataset (e.g., quality control mechanisms in terms of data integrity, data completeness, and data bias)?
* Which inclusion and exclusion criteria for a given dataset have been applied (e.g., comprehensiveness, coverage of target demographic setting, or size of the dataset)?
* How was the data submission, collection, and handling organized from the technical and operational point of view (e.g., folder structures, file formats, technical metadata encoding, compression, encryption, and password exchange)?
* Specific data governance derived by the general data governance document (currently [F-103](https://www.itu.int/en/ITU-T/focusgroups/ai4h/Documents/FGAI4H-F-103-DataPolicy.pdf) and the deliverables beginning with [DEL05](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B2012357A-941E-44BD-B965-370D7829F52C%7D&file=DEL05.docx&action=default))
* How was the overall quality, coverage, and bias of the accumulated dataset assessed (e.g., if several datasets from several hospitals were merged with the goal to have better coverage of all regions and ethnicities)?
* Was any kind of post-processing applied to the data (e.g., data transformations, repackaging, or merging)?
* How was the annotation organized?
	+ How many annotators/peer reviewers were engaged?
	+ Which scores, metrics, and thresholds were used to assess the label quality and the need for an arbitration process?
	+ How have inter-annotator disagreements been resolved (i.e., what was the arbitration process)?
	+ If annotations were part of the submitted dataset, how was the quality of the annotations controlled?
	+ How was the annotation of each case documented?
	+ Were metadata on the annotation process included in the data (e.g., is it possible to compare the benchmarking performance based on the annotator agreement)?
* Were data/label update/amendment policies and/or criteria in place?
* How was access to test data controlled (e.g., to ensure that no one could access, manipulate, and/or leak data and data labels)? Please address authentication, authorization, monitoring, logging, and auditing
* How was data loss avoided (e.g., backups, recovery, and possibility for later reproduction of the results)?
* Is there assurance that the test dataset is undisclosed and was never previously used for training or testing of any AI model?
* What mechanisms are in place to ensure that test datasets are used only once for benchmarking? (Each benchmarking session will need to run with a new and previously undisclosed test dataset to ensure fairness and no data leakage to subsequent sessions)

#### Data sharing policies

This section provides details about legalities in the context of benchmarking. Each dataset that is shared should be protected by special agreements or contracts that cover, for instance, the data sharing period, patient consent, and update procedure (see also [DEL05\_5](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B71FE8B9D-ACB3-48CE-AA3F-136409B550A4%7D&file=DEL05_5.docx&action=default) on *data handling* and [DEL05\_6](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B5C95327E-96A5-4175-999E-3EDB3ED147C3%7D&file=DEL05_6.docx&action=default) on *data sharing practices*).

* Which legal framework was used for data sharing?
* Was a data sharing contract signed and what was the content? Did it contain:
	+ Purpose and intended use of data
	+ Period of agreement
	+ Description of data
	+ Metadata registry
	+ Data harmonization
	+ Data update procedure
	+ Data sharing scenarios
		- Data can be shared in public repositories
		- Data are stored in local private databases (e.g., hospitals)
	+ Rules and regulation for patients’ consent
	+ Data anonymization and de-identification procedure
	+ Roles and responsibilities
		- Data provider
		- Data protection officer
		- Data controllers
		- Data processors
		- Data receivers
* Which legal framework was used for sharing the AI?
* Was a contract signed and what was the content?

#### Baseline acquisition

The main purpose of benchmarking is to provide stakeholders with the numbers they need to decide whether AI models provide a viable solution for a given health problem in a designated context. To achieve this, the performance of the AI models needs to be compared with available options achieving the same clinically meaningful endpoint. This, in turn, requires data on the performance of the alternatives, ideally using the same benchmarking data. As the current alternatives typically involve doctors, it might make sense to combine the test data acquisition and labelling with additional tasks that allow the performance of the different types of health workers to be assessed.

* Does this topic require comparison of the AI model with a baseline (gold standard) so that stakeholders can make decisions?
* Is the baseline known for all relevant application contexts (e.g., region, subtask, sex, age group, and ethnicity)?
* Was a baseline assessed as part of the benchmarking?
* How was the process of collecting the baseline organized? If the data acquisition process was also used to assess the baseline, please describe additions made to the process described in the previous section.
* What are the actual numbers (e.g., for the performance of the different types of health workers doing the task)?

#### Reporting methodology

This section discusses how the results of the benchmarking runs will be shared with the participants, stakeholders, and general public.

* What is the general approach for reporting results (e.g., leader board vs. drill down)?
* How can participants analyse their results (e.g., are there tools or are detailed results shared with them)?
* How are the participants and their AI models (e.g., versions of model, code, and configuration) identified?
* What additional metadata describing the AI models have been selected for reporting?
* How is the relationship between AI results, baselines, previous benchmarking iterations, and/or other benchmarking iterations communicated?
* What is the policy for sharing participant results (e.g., opt in or opt out)? Can participants share their results privately with their clients (e.g., as in a credit check scenario)?
* What is the publication strategy for the results (e.g., website, paper, and conferences)?
* Is there an online version of the results?
* Are there feedback channels through which participants can flag technical or medical issues (especially if the benchmarking data was published afterwards)?
* Are there any known limitations to the value, expressiveness, or interpretability of the reports?

#### Result

This section gives an overview of the results from runs of this benchmarking version of your topic. Even if your topic group prefers an interactive drill-down rather than a leader board, pick some context of common interest to give some examples.

* When was the benchmarking executed?
* Who participated in the benchmarking?
* What overall performance of the AI systems concerning medical accuracy, robustness, and technical performance (minimum, maximum, average etc.) has been achieved?
* What are the results of this benchmarking iteration for the participants (who opted in to share their results)?

#### Discussion of the benchmarking

This section discusses insights of this benchmarking iterations and provides details about the ‘outcome’ of the benchmarking process (e.g., giving an overview of the benchmark results and process).

* What was the general outcome of this benchmarking iteration?
* How does this compare to the goals for this benchmarking iteration (e.g., was there a focus on a new aspect to benchmark)?
* Are there real benchmarking results and interesting insights from this data?
	+ How was the performance of the AI system compared to the baseline?
	+ How was the performance of the AI system compared to other benchmarking initiatives (e.g., are the numbers plausible and consistent with clinical experience)?
	+ How did the results change in comparison to the last benchmarking iteration?
* Are there any technical lessons?
	+ Did the architecture, implementation, configuration, and hosting of the benchmarking system fulfil its objectives?
	+ How was the performance and operational efficiency of the benchmarking itself (e.g., how long did it take to run the benchmarking for all AI models vs. one AI model; was the hardware sufficient)?
* Are there any lessons concerning data acquisition?
	+ Was it possible to collect enough data?
	+ Were the data as representative as needed and expected?
	+ How good was the quality of the benchmarking data (e.g., how much work went into conflict resolution)?
	+ Was it possible to find annotators?
	+ Was there any relevant feedback from the annotators?
	+ How long did it take to create the dataset?
* Is there any feedback from stakeholders about how the benchmarking helped them with decision-making?
	+ Are metrics missing?
	+ Do the stakeholders need different reports or additional metadata (e.g., do they need the “offline capability” included in the AI metadata so that they can have a report on the best offline system for a certain task)?
* Are there insights on the benchmarking process?
	+ How was the interest in participation?
	+ Are there reasons that someone could not join the benchmarking?
	+ What was the feedback of participants on the benchmarking processes?
	+ How did the participants learn about the benchmarking?

#### Retirement

This section addresses what happens to the AI system and data after the benchmarking activity is completed. It might be desirable to keep the database for traceability and future use. Alternatively, there may be security or privacy reasons for deleting the data. Further details can be found in the reference document of this section [DEL04](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BC68833D1-9B31-4E8E-8A4A-3939D7DEA56F%7D&file=DEL04.docx&action=default) “*AI software lifecycle specification”* (identification of standards and best practices that are relevant for the AI for health software life cycle).

* What happens with the data after the benchmarking (e.g., will they be deleted, stored for transparency, or published)?
* What happens to the submitted AI models after the benchmarking?
* Could the results be reproduced?
* Are there legal or compliance requirements to respond to data deletion requests?

### Benchmarking version [X]

This section includes all technological and operational details of the benchmarking process for the benchmarking version [X].

## Subtopic [B]

# Overall discussion of the benchmarking

This section discusses the overall insights gained from benchmarking work in this topic group. This should not be confused with the discussion of the results of a concrete benchmarking run (e.g., in 4.2.11).

==Discussion initiated==

* What is the overall outcome of the benchmarking thus far?
* Have there been important lessons?
* Are there any field implementation success stories?
* Are there any insights showing how the benchmarking results correspond to, for instance, clinical evaluation?
* Are there any insights showing the impact (e.g., health economic effects) of using AI systems that were selected based on the benchmarking?
* Was there any feedback from users of the AI system that provides insights on the effectiveness of benchmarking?
	+ Did the AI system perform as predicted relative to the baselines?
	+ Did other important factors prevent the use of the AI system despite a good benchmarking performance (e.g., usability, access, explainability, trust, and quality of service)?
* Were there instances of the benchmarking not meeting the expectations (or helping) the stakeholders? What was learned (and changed) as a result?
* What was learned from executing the benchmarking process and methodology (e.g., technical architecture, data acquisition, benchmarking process, benchmarking results, and legal/contractual framing)?

# Regulatory considerations

For AI-based technologies in healthcare, regulation is not only crucial to ensure the safety of patients and users, but also to accomplish market acceptance of these devices. This is challenging because there is a lack of universally accepted regulatory policies and guidelines for AI-based medical devices. To ensure that the benchmarking procedures and validation principles of FG-AI4H are secure and relevant for regulators and other stakeholders, the working group on *“*[*Regulatory considerations on AI for health”* *(WG-RC)*](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/wg/SitePages/WG-RC.aspx) compiled the requirements that consider these challenges.

The deliverables with relevance for regulatory considerations are [DEL02](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BF2F46A99-7457-4BC8-81A3-0E1E63D6072A%7D&file=DEL02.docx&action=default) *“AI4H regulatory considerations”* (which provides an educational overview of some key regulatory considerations), [DEL02\_1](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B6AF7C004-8BCE-4151-9F44-45F041A1EB1D%7D&file=DEL02_1.docx&action=default) *“Mapping of IMDRF essential principles to AI for health software”,* and[DEL02\_2](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7B1ED0D4D1-876C-4A0F-AEF7-06D3F445F5E6%7D&file=DEL02_2.docx&action=default) *“Guidelines for AI based medical device (AI-MD): Regulatory requirements”* (which provides a checklist to understand expectations of regulators, promotes step-by-step implementation of safety and effectiveness of AI-based medical devices, and compensates for the lack of a harmonized standard). [DEL04](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BC68833D1-9B31-4E8E-8A4A-3939D7DEA56F%7D&file=DEL04.docx&action=default) identifies standards and best practices that are relevant for the “*AI software lifecycle specification*.*”* The following sections discuss how the different regulatory aspects relate to the TG-AI for Traditional Medicine.

==Discussion initiated==

## Existing applicable regulatory frameworks

Most of the AI systems that are part of the FG-AI4H benchmarking process can be classified as *software as medical device* (SaMD) and eligible for a multitude of regulatory frameworks that are already in place. In addition, these AI systems often process sensitive personal health information that is controlled by another set of regulatory frameworks. The following section summarizes the most important aspects that AI manufacturers need to address if they are developing AI systems for Traditional Medicine.

* What existing regulatory frameworks cover the type of AI in this TDD (e.g., MDR, FDA, GDPR, and ISO; maybe the systems in this topic group always require at least “MDR class 2b” or maybe they are not considered a medical device)?
* Are there any aspects to this AI system that require additional specific regulatory considerations?

## Regulatory features to be reported by benchmarking participants

In most countries, benchmarked AI solutions can only be used legally if they comply with the respective regulatory frameworks for the application context. This section outlines the compliance features and certifications that the benchmarking participants need to provide as part of the metadata. It facilitates a screening of the AI benchmarking results for special requirements (e.g., the prediction of prediabetes in a certain subpopulation in a country compliant to the particular regional regulatory requirements).

* Which certifications and regulatory framework components of the previous section should be part of the metadata (e.g., as a table with structured selection of the points described in the previous section)?

## Regulatory requirements for the benchmarking systems

The benchmarking system itself needs to comply with regulatory frameworks (e.g., some regulatory frameworks explicitly require that all tools in the quality management are also implemented with a quality management system in place). This section outlines the regulatory requirements for software used for benchmarking in this topic group.

* Which regulatory frameworks apply to the benchmarking system itself?
* Are viable solutions with the necessary certifications already available?
* Could the TG implement such a solution?

## Regulatory approach for the topic group

Building on the outlined regulatory requirements, this section describes how the topic group plans to address the relevant points in order to be compliant. The discussion here focuses on the guidance and best practice provided by the [DEL02](https://extranet.itu.int/sites/itu-t/focusgroups/ai4h/_layouts/15/WopiFrame.aspx?sourcedoc=%7BF2F46A99-7457-4BC8-81A3-0E1E63D6072A%7D&file=DEL02.docx&action=default) *“AI4H regulatory considerations.”*

* Documentation & Transparency
	+ How will the development process of the benchmarking be documented in an effective, transparent, and traceable way?
* Risk management & Lifecycle approach
	+ How will the risk management be implemented?
	+ How is a life cycle approach throughout development and deployment of the benchmarking system structured?
* Data quality
	+ How is the test data quality ensured (e.g., the process of harmonizing data of different sources, standards, and formats into a single dataset may cause bias, missing values, outliers, and errors)?
	+ How are the corresponding processes document?
* Intended Use & Analytical and Clinical Validation
	+ How are technical and clinical validation steps (as part of the lifecycle) ensured (e.g., as proposed in the IMDRF clinical evaluation framework)?
* Data Protection & Information Privacy
	+ How is data privacy in the context of data protection regulations ensured, considering regional differences (e.g., securing large data sets against unauthorized access, collection, storage, management, transport, analysis, and destruction)? This is especially relevant if real patient data is used for the benchmarking.
* Engagement & Collaboration
	+ How is stakeholder (regulators, developers, healthcare policymakers) feedback on the benchmarking collected, documented, and implemented?

# References

*Please see the endnote.*

|  |  |  |
| --- | --- | --- |
| A |  | B |

Captions for figures use WinWord style "Figure\_No & title"

Figure 1: Example of a figure

Annex A:
Glossary

This section lists all the relevant abbreviations, acronyms and uncommon terms used in the document.

|  |  |  |
| --- | --- | --- |
| Acronym/Term | Expansion | Comment |
| TDD | Topic Description Document | Document specifying the standardized benchmarking for a topic on which the FG AI4H Topic Group works. This document is the TDD for the Topic Group [YOUR TOPIC GROUP] |
| TG | Topic Group |  |
| WG | Working Group |  |
| FGAI4H | Focus Group on AI for Health |  |
| AI | Artificial intelligence |  |
| ITU | International Telecommunication Union |  |
| WHO | World Health Organization |  |
| DEL | Deliverable  |  |
| CfTGP | Call for topic group participation |  |
| AI4H  | Artificial intelligence for health |  |
| IMDRF | International Medical Device Regulators Forum |  |
| MDR | Medical Device Regulation |  |
| ISO | International Standardization Organization |  |
| GDPR | General Data Protection Regulation |  |
| FDA | Food and Drug administration |  |
| SaMD | Software as a medical device |  |
| AI-MD | AI based medical device |  |
| LMIC | Low-and middle-income countries |  |
| GDP | Gross domestic product |  |
| API | Application programming interface |  |
| IP | Intellectual property |  |
| PII | Personal identifiable information |  |
| […] |  |  |

Annex B:
Declaration of conflict of interests

In accordance with the ITU transparency rules, this section lists the conflict-of-interest declarations for everyone who contributed to this document. Please see the guidelines in FGAI4H-F-105 “ToRs for the WG-Experts and call for experts” and the respective forms (Application form & Conflict of interest form).

Institution

Ministry of Ayush, Government of India is a dedicated government body which is working in the domain of traditional medicine in India. In collaboration with World Health Organization and other concerned agencies, the Ministry is actively contributing towards development of standards, benchmarking for TM practices, education, training, and quality control of products for the benefit of humankind. As part of the same, this topic group has been created and will be led as a collaborative effort for developing benchmarking for utilization of AI/ML for TM world over. It is hereby declared that, the Ministry and the topic driver who is coordinating this activity is a research officer working with the Ministry and has no commercial and conflict of interests out of this activity.

References:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. https://www.frontiersin.org/articles/10.3389/fphar.2022.826044/full [↑](#endnote-ref-1)
2. https://www.frontiersin.org/articles/10.3389/fphar.2022.826044/full [↑](#endnote-ref-2)
3. https://www.sciencedirect.com/science/article/pii/S258937772100001X [↑](#endnote-ref-3)
4. Liu, S., Hua, L., Lv, P., Yu, Y., Gao, Y., & Sheng, X. (2018). A Pulse Condition Reproduction Apparatus for Remote Traditional Chinese Medicine. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10984 LNAI*, 453–464. https://doi.org/10.1007/978-3-319-97586-3\_41 [↑](#endnote-ref-4)
5. Xu, L., Meng, M. Q. H., Shi, C., Wang, K., & Li, N. (2008). Quantitative analyses of pulse images in traditional Chinese medicine. *Medical Acupuncture*, *20*(3), 175–189. https://doi.org/10.1089/ACU.2008.0632 [↑](#endnote-ref-5)
6. Xu, L. S., Meng, M. Q. H., & Wang, K. Q. (2007). Pulse image recognition using fuzzy neural network. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 3148–3151. https://doi.org/10.1109/IEMBS.2007.4352997 [↑](#endnote-ref-6)
7. Yeuk-Lan Alice, L., Binghe, G., Shuang, C., Hoyin, C., Kawai, K., Wenjung, L., & Jiangang, S. (2021). Artificial intelligence meets traditional Chinese medicine: a bridge to opening the magic box of sphygmopalpation for pulse pattern recognition. *Digital Chinese Medicine*, *4*(1), 1–8. https://doi.org/10.1016/J.DCMED.2021.03.001 [↑](#endnote-ref-7)
8. A. Joshi, A. Kulkarni, S. Chandran, V. K. Jayaraman and B. D. Kulkarni, "Nadi Tarangini: A Pulse Based Diagnostic System," 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 2207-2210, doi: 10.1109/IEMBS.2007.4352762. [↑](#endnote-ref-8)
9. <div class="csl-entry">Joshi, A., Kulkarni, A., Chandran, S., Jayaraman, V. K., &#38; Kulkarni, B. D. (2007). Nadi Tarangini: A Pulse Based Diagnostic System. <i>2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society</i>, 2207–2210. https://doi.org/10.1109/IEMBS.2007.4352762</div> [↑](#endnote-ref-9)
10. https://www.frontiersin.org/articles/10.3389/fphar.2022.826044/full [↑](#endnote-ref-10)
11. https://www.frontiersin.org/articles/10.3389/fphar.2022.826044/full [↑](#endnote-ref-11)
12. Zhang, H., Ni, W., Li, J., & Zhang, J. (2020). Artificial intelligence-based traditional chinese medicine assistive diagnostic system: Validation study. *JMIR Medical Informatics*, *8*(6). https://doi.org/10.2196/17608 [↑](#endnote-ref-12)
13. Wallace, R. K. (2020). Ayurgenomics and modern medicine. *Medicina (Lithuania)*, *56*(12), 1–7. https://doi.org/10.3390/medicina56120661 [↑](#endnote-ref-13)
14. Prasher, B., Varma, B., Kumar, A., Khuntia, B. K., Pandey, R., Narang, A., Tiwari, P., Kutum, R., Guin, D., Kukreti, R., Dash, D., Mukerji, M., Aggarwal, S., Natarajan, V., Salvi, S., Aatreya, P., Unni, S., Mishra, N., Mudgal, N., … Makhija, N. (2017). Ayurgenomics for stratified medicine: TRISUTRA consortium initiative across ethnically and geographically diverse Indian populations. *Journal of Ethnopharmacology*, *197*, 274–293. https://doi.org/10.1016/j.jep.2016.07.063 [↑](#endnote-ref-14)
15. Prasher, B., Aggarwal, S., Mandal, A. K., Sethi, T. P., Deshmukh, S. R., Purohit, S. G., Sengupta, S., Khanna, S., Mohammad, F., Garg, G., Brahmachari, S. K., & Mukerji, M. (2008). Whole genome expression and biochemical correlates of extreme constitutional types defined in Ayurveda. *Journal of Translational Medicine*, *6*. https://doi.org/10.1186/1479-5876-6-48 [↑](#endnote-ref-15)
16. Bhargav H, Jasti N, More P, Kumar V, Chikkanna U, Kishore Kumar R, et al. Correlation of prakriti diagnosis using AyuSoft prakriti diagnostic tool with clinician rating in patients with psychiatric disorders. J Ayurveda Integr Med [Internet]. 2021;12(2):365–8. Available from: https://www.sciencedirect.com/science/article/pii/S0975947621000115 [↑](#endnote-ref-16)
17. .https://www.researchgate.net/publication/322899120\_Development\_of\_Standardized\_Prakriti\_Assessment\_Tool\_An\_Overview\_of\_Ongoing\_CCRAS\_Initiatives [↑](#endnote-ref-17)
18. Madaan, V., & Goyal, A. (2020). Predicting Ayurveda-based constituent balancing in human body using machine learning methods. *IEEE Access*, *8*, 65060–65070. https://doi.org/10.1109/access.2020.2985717 [↑](#endnote-ref-18)
19. Tiwari, P., Kutum, R., Sethi, T., Shrivastava, A., Girase, B., Aggarwal, S., Patil, R., Agarwal, D., Gautam, P., Agrawal, A., Dash, D., Ghosh, S., Juvekar, S., Mukerji, M., & Prasher, B. (2017). Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits. *PLoS ONE*, *12*(10). https://doi.org/10.1371/journal.pone.0185380 [↑](#endnote-ref-19)
20. Khatua, D., Sekh, A. A., Kutum, R., Mukherji, M., Prasher, B., & Kar, S. (2023). Classification of Ayurveda constitution types: a deep learning approach. *Soft Computing*, 1–9. https://doi.org/10.1007/S00500-023-07942-2/METRICS [↑](#endnote-ref-20)
21. Wayne, P. M., Manor, B., Novak, V., Costa, M. D., Hausdorff, J. M., Goldberger, A. L., Ahn, A. C., Yeh, G. Y., Peng, C. K., Lough, M., Davis, R. B., Quilty, M. T., & Lipsitz, L. A. (2013). A systems biology approach to studying Tai Chi, physiological complexity and healthy aging: Design and rationale of a pragmatic randomized controlled trial. *Contemporary Clinical Trials*, *34*(1), 21–34. https://doi.org/10.1016/J.CCT.2012.09.006 [↑](#endnote-ref-21)
22. *Tai Chi, Physiological Complexity, and Healthy Aging - Gait v1.0.2*. (n.d.). Retrieved February 11, 2023, from https://physionet.org/content/taichidb/1.0.2/ [↑](#endnote-ref-22)
23. *A multi-camera and multimodal dataset for posture and gait analysis v1.0.0*. (n.d.). Retrieved February 11, 2023, from https://physionet.org/content/multi-gait-posture/1.0.0/ [↑](#endnote-ref-23)
24. Jayasundar R, Singh A, Kumar D. Challenges in using electronic tongue to study rasa of plants: I. Finding the right tool for the right job. J Ayurveda Integr Med [Internet]. 2021;12(2):234–7. Available from: https://www.sciencedirect.com/science/article/pii/S0975947620301467 [↑](#endnote-ref-24)
25. Kumar D, Singh A, Jayasundar R. Challenges in using Electronic tongue to study rasa of plants: II. Impact of solvent and concentration on sensor response and taste ranking. J Ayurveda Integr Med [Internet]. 2021;12(2):238–44. Available from: https://www.sciencedirect.com/science/article/pii/S0975947620301455 [↑](#endnote-ref-25)
26. Chu, H., Moon, S., Park, J., Bak, S., Ko, Y., & Youn, B. Y. (2022). The Use of Artificial Intelligence in Complementary and Alternative Medicine: A Systematic Scoping Review. *Frontiers in Pharmacology*, *13*. https://doi.org/10.3389/FPHAR.2022.826044 [↑](#endnote-ref-26)
27. Astin, J. A. (1998). Why Patients Use Alternative Medicine: Results of a National Study. *JAMA*, *279*(19), 1548–1553. https://doi.org/10.1001/JAMA.279.19.1548 [↑](#endnote-ref-27)
28. Jansen, C., Baker, J. D., Kodaira, E., Ang, L., Bacani, A. J., Aldan, J. T., Shimoda, L. M. N., Salameh, M., Small-Howard, A. L., Stokes, A. J., Turner, H., & Adra, C. N. (2021). Medicine in motion: Opportunities, challenges and data analytics-based solutions for traditional medicine integration into western medical practice. *Journal of Ethnopharmacology*, *267*, 113477. https://doi.org/10.1016/J.JEP.2020.113477 [↑](#endnote-ref-28)
29. *HTA in Switzerland - SwissHTA - Swiss Health Technology Assessment*. (n.d.). Retrieved December 26, 2022, from http://www.swisshta.org/index.php/HTA\_in\_Switzerland.html [↑](#endnote-ref-29)
30. Ammon, K. von, Cardini, F., Daig, U., Dragan, S., FreiErb, Martin, Hegyi, G., Sarsina, P. R. di, Sørensen, J., Ursoniu, S., Weidenhammer, W., & Lewith, G. (n.d.). *Health Technology Assessment (HTA) and a map of CAM provision in the EU (Final Report of CAMbrella Work Package 5)*. CAMBRELLLA. Retrieved December 26, 2022, from https://cam-europe.eu/wp-content/uploads/2018/09/WP5-CAMbrella-WP5final.pdf [↑](#endnote-ref-30)
31. Bornhöft, Gudrun., & Matthiessen, P. F. (2011). *Homeopathy in healthcare -- Effectiveness, appropriateness, safety, costs : an HTA report on homeopathy as part of the Swiss Complementary Medicine Evaluation Programme*. 209. https://www.hri-research.org/resources/homeopathy-the-debate/the-swiss-hta-report-on-homeopathy/ [↑](#endnote-ref-31)
32. Rudra, S., Kalra, A., Kumar, A., & Joe, W. (2017). Utilization of alternative systems of medicine as health care services in India: Evidence on AYUSH care from NSS 2014. *PloS One*, *12*(5). https://doi.org/10.1371/JOURNAL.PONE.0176916 [↑](#endnote-ref-32)
33. Saketh Ram et al. (2017). National Ayurveda Morbidity Codes (NAMC). In *National Ayush Morbidity and Standardized Electronic (NAMASTE) Portal*. Ministry of Ayush, Government of India. http://namstp.ayush.gov.in/#/Ayurveda [↑](#endnote-ref-33)
34. Anonymous. (2021). *Glossary of Ayurvedic Terminolgy Part1-5*. https://www.services.bis.gov.in:8071/php/BIS\_2.0/bisconnect/cls\_module/Ministry\_list/ministry\_stndrds\_list?mns\_id=NTI%3D&mns\_name=TWluaXN0cnkgb2YgQVlVU0g%3D&aspect=&from=&to= [↑](#endnote-ref-34)
35. https://www.who.int/publications/ [↑](#endnote-ref-35)
36. https://www.who.int/publications/i/item/9789240064935 [↑](#endnote-ref-36)
37. https://www.who.int/publications/i/item/9789240064973 [↑](#endnote-ref-37)
38. https://www.who.int/publications/i/item/9789240064959 [↑](#endnote-ref-38)
39. https://www.who.int/publications/i/item/9789240042322 [↑](#endnote-ref-39)