Appendix 1
to the Order N 134
of the Moscow Health Care
Department
dated February 16, 2023

PROCEDURE FOR
the Experiment on the use of innovative computer vision technologies for the analysis of medical images and further application in the Moscow healthcare system

1. Terms, abbreviations and definitions

1.1. For the purpose of this Procedure for the Experiment on the use of innovative computer vision technologies for the analysis of medical images and further application in the Moscow healthcare system (hereinafter, the Procedure) the following definitions and abbreviations are used:

- DD – diagnostic device connected to the Unified Radiological Information Service of the Unified Medical Information and Analytical System of Moscow and owned by a medical facility subordinate to the Moscow public healthcare system;
- URIS UMIAS – Unified Radiological Information Service of the Unified Medical Information and Analytical System of Moscow;
- USNEI – Unified System of Notifications for External Interactions;
- Application – an application for the participation in the Experiment from a legal entity that has developed and/or has the right to operate the Service;
- CS – comprehensive Service;
- CT – calibration testing of the software based on computer vision technologies designed to analyse medical images obtained by means of designated imaging modalities;
- Committee – the Committee of the Moscow Health Care Department for reviewing grant applications from legal entities, and evaluating the operation of the Services based on computer vision technologies;
- MF – medical facility subordinate to the Moscow public healthcare system;
- Monitoring – monitoring over the operational parameters of the software based on computer vision technologies designed to analyse medical images obtained by means of designated imaging modalities;

- IHSC URIS UMIAS – Industrial Hardware-Software Complex URIS UMIAS;
- Applicant – a legal entity that has developed and/or has the right to operate the software based on computer vision technologies designed to analyse medical images obtained by means of designated imaging modalities;
- Service – a piece of software based on computer vision technologies designed to analyse medical images obtained by means of designated imaging modalities;
- THSC URIS UMIAS – Testing Hardware-Software Complex URIS UMIAS;
- Participant – an Applicant, whose Service successfully passed the functional and calibration testing and has been integrated into IHSC URIS UMIAS;
- FT – functional testing of the Service;

- Target pathology – signs of a disease or a pathology that the Service is expected to detect during the analysis of radiological studies.

2. General provisions

2.1. This Procedure defines the terms and conditions for conducting the Experiment on the use of innovative computer vision technologies for the analysis of medical images and further application in the Moscow healthcare system (hereinafter, the Experiment).

2.2. The organization and support of the Experiment is carried out by the Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department (hereinafter – the Center for Diagnostics and Telemedicine).

2.3. The Experiment is to assess the methods designed to support medical decision-making in the Moscow healthcare system by means of advanced innovative data analysis.

2.4. The Experiment is a prospective research and practical study whose goal is to assess and compare the Services for the following parameters:

- diagnostic accuracy (estimated, among other things, using identical samples of studies for the Services that operate simultaneously)
- applicability
- ease of use
- operation efficiency (including the impact on labor productivity and the reading time).

2.5. The Moscow Health Care Department (hereinafter, the Department) shall hold the exclusive rights to the data and the results obtained during the Experiment (including the rights to publications and reports in mass media, the academic periodicals, conferences, and the execution of the paperwork for the intellectual property rights, and not including the rights to the intellectual property of the Participants).

2.6. The Experiment includes the following experimental setups:

- computed tomography of the brain to detect ischemic stroke and intracranial hemorrhage;
- computed tomography and / or low-dose computed tomography of the chest to detect malignant neoplasms of the lungs, lung involvement in COVID-19, impairment of lung airness, compression fractures of the vertebral bodies, ischemic heart disease (coronary calcium, paracardial fat), emphysema, thoracic aortic aneurysms which includes measurement of the thoracic aorta diameter, dilation of the pulmonary trunk which includes measurement of the pulmonary trunk diameter, free fluid (effusion) in the pleural cavities, enlarged lymph nodes (lymphadenopathy), pulmonary tuberculosis, sarcoidosis, bronchiectasis, adrenal lesions, rib(s) fracture, focal lesions in the chest bones;
- computed tomography of the abdomen to detect urolithiasis, adrenal lesions, liver lesions, renal lesions, compression fractures of the vertebral bodies, aneurysm of the abdominal aorta which includes measurement of the abdominal aorta diameter, focal lesions in the abdominal and pelvic bones, gallstone disease (calculous gallbladder);
- magnetic resonance imaging of the brain to detect malignant neoplasms and multiple sclerosis;
- magnetic resonance imaging of the cervical spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis;

- magnetic resonance imaging of the thoracic spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis;
- magnetic resonance imaging of the lumbosacral spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis;
- magnetic resonance imaging of the knee joint to detect articular cartilage damage (chondromalacia);
- mammography to detect breast cancer;
- head X-ray to detect sinusitis;
- chest X-ray to detect various pathologies;
- photofluorography of the lungs to detect various pathologies;
- X-ray of the musculoskeletal system to detect various pathologies of the spine (osteochondrosis, scoliosis, spondylolisthesis, fracture of the vertebral bodies), fractures of the bones of the upper and lower extremities (wrist, shoulder, hip and ankle joints), arthrosis (hip, knee), longitudinal and transverse flat foot;
- automated routine measurements during CT of the brain and abdomen (kidneys, liver, pancreas and spleen); automated routine MRI of the brain and pelvis (uterus and prostate gland).

2.7. During the Experiment, the introduction of the setups specified in clause 2.6 of this Procedure will be carried out on a step-by-step basis. The stages and time frames for the introduction are published and updated on the Experiment's website www.mosmed.ai.

The Center for Diagnostics and Telemedicine has the right to change the stages and time frames for the introduction of the experimental setups, including after reviewing the Applications submitted by the Applicants and Participants.

3. Requirements for Participants and Services

3.1. Service developers and their legal representatives can become the Participants in the Experiment

3.2. Depending on the number of target pathologies to be detected with one of the imaging modalities, the following Services may participate in the Experiment:

3.2.1. A Service capable of analyzing medical images to identify a target pathology by means of the imaging modalities.

3.2.2. A Comprehensive Service capable of analyzing medical images to detect a set of target pathologies established by this Procedure within the framework of an experimental setup.

3.3. Requirements to the Services:

- capabilities that match the Baseline Functional Requirements approved by the Center for Diagnostics and Telemedicine and published on the Experiment's website www.mosmed.ai;
- compliance with the Baseline Diagnostic Requirements to the deliverables, approved by the Center for Diagnostics and Telemedicine and published on the Experiment's website www.mosmed.ai;
- compliance with the target diagnostic accuracy metrics specified in clause 7.1.5 hereof.

3.4. The Participant shall:

- observe the conditions of the Experiment

- observe the rules related to the confidentiality of personal data and other legally protected information
- store, process and use the data obtained during the Experiment strictly on the territory of the Russian Federation and for the purposes of the Experiment, and shall not transfer it to any third parties (which includes preventing the risks of third-party access)
- delete all image data and analysis results upon completion of the Experiment.

4. Request for participation

4.1. Participation in the Experiment is carried out on the basis of the Application.

4.2. An Applicant shall submit a formal Application (see Appendix 1) and supporting documents to the Center for Diagnostics and Telemedicine at the following address: Moscow, 24 bld. 1, Petrovka ul. The copies of the said documents must be sent to the Experiment's email ai@npcmr.ru, and/or through the Applicant's personal user account on the Experiment's website www.mosmed.ai.

4.3. The following supporting documents shall be attached to the Application:

4.3.1. Certified copies of the documents evidencing state registration of the Applicant's legal entity on the territory of the Russian Federation.

4.3.2. Documents evidencing the development and / or the right to operate the Service based on the computer vision technologies for the analysis of medical images.

4.3.3. A report on preliminary clinical and technical testing conducted using medical images acquired from the citizens of the Russian Federation and / or Caucasian and Mongoloid individuals, prepared in accordance with the Guidelines No. 43 "Clinical Acceptance of Software Based on Artificial Intelligence Technologies (Radiology)" (recommended by the Expert Council for Science of the Moscow Health Care Department, Protocol No. 8 of June 25, 2019).

The report on preliminary clinical and technical testing must be signed by an authorized person and sealed by the medical organization that prepared the report.

The report on preliminary clinical and technical testing shall contain the following information:

- classic ROC curve and area under the curve (AUC)
- sensitivity, specificity and accuracy
- a four-section spreadsheet that compares the results of index testing and the reference testing
- a reading time for one study.

4.3.4. Documentation outlining and evidencing the Service capabilities that meet the Baseline Functional Requirements specified in clause 3.3 hereof, along with the technical architecture and characteristics, use case scenarios and operational requirements to the software and hardware that power the Service.

4.3.5. Report on the clinical implementation and/or acceptance testing of the Service across medical facilities of the Russian Federation or other countries, indicating the testing duration and the number of medical facilities where the Service was tested (if available).

4.3.6. Availability of certified copies of the certificates for compliance of the Applicant's Service with the local quality standards adopted by FDA, CE, and their equivalents in other countries, or a certified copy of a certificate for state registration as a medical device (if available).

4.3.7. Copies of scientific articles investigating the accuracy and efficacy of the Service published in peer-reviewed journals indexed in Scopus and/or Web of Science (if available).

4.3.8. Valid certificate of conformity of the quality management system to the ISO standard (if available).

4.3.9. Certified copies of a certificate for state registration as a medical device (if available).

4.3.10. Other documents that the Applicant may consider appropriate.

4.4. If the original document is issued in a language other than Russian, it must be accompanied by a translation into Russian, certified by an authorized person and stamped by the Applicant.

4.5. The Center for Diagnostics and Telemedicine, within **three business days** from the date the Application is hand-delivered to the Center for Diagnostics and Telemedicine's address specified in clause 4.2 hereof, shall review the Application for compliance with the requirements set out in clauses 4.3 - 4.4 hereof. During the review, the Center for Diagnostics and Telemedicine may request clarification from the Applicant regarding the submitted Application and the documents and information attached thereto. The review of the Application shall be suspended until the clarification is provided.

4.5.1. Should the Application fail to meet the requirements set out in clauses 4.3 - 4.4 hereof, the Applicant shall receive a corresponding notice to the email address specified in the Application within **three business days** upon completion of the review and issuing a decision.

4.5.2. The Applicant shall bear all the risks related to the delivery of the correspondence on all emerging issues within the framework of the Experiment to the email address specified in the Application.

4.5.3. The Applicant shall have the right to re-submit the Application after removal of the discrepancies associated with the completeness of the supporting documentation and / or compliance with the requirements specified in clauses 4.3 - 4.4 hereof.

4.5.4. Should the Application meet the requirements set out in clauses 4.3 - 4.4 hereof, the technical integration into URIS UMIAS shall begin within **one business day**. A letter on Application acceptance and initiating the technical integration shall be sent to the Applicant's email address specified in the Application.

4.5.5. The Applications approved in 2022 that match the experimental setups established by this Procedure (clause 2.6) shall be considered valid for 2023.

5. Technical integration

5.1. Technical integration of the Service shall be carried out in the THSC URIS UMIAS environment. The integration period **shall not exceed 15 calendar days** from the date of sending a letter on initiating the technical integration into URIS UMIAS to the Applicant's e-mail.

5.2. The integration period may be extended upon request from the Applicant sent to the e-mail address of the Experiment: ai@zdrav.mos.ru.

5.3. The technical integration involves setting up the access to the URIS UMIAS environment and verification of the consistency of data exchange between the Applicant's Service and URIS UMIAS. Where applicable, the Applicant shall make the necessary improvements to the Service to complete the integration process. Once the technical integration is completed, the Service shall undergo the FT and CT procedures. A

description of the technical integration process is published on the Experiment's website www.mosmed.ai.

6. Functional testing

6.1. FT of the Service operation is divided into three stages:

6.1.1. THSC URIS UMIAS provides the Service with access to up to 5 studies.

6.1.2. The Service shall analyze these studies one by one and submit the results to THSC URIS UMIAS.

6.1.3. Based on the FT outcomes, the Applicant shall complete Table 1 of Appendix 2 to this Procedure and submit it to the Center for Diagnostics and Telemedicine.

6.1.4. The experts of the Center for Diagnostics and Telemedicine shall evaluate the data from Table 1 for compliance with the parameters stated in the Application and the Baseline Functional and Diagnostic Requirements for the Service deliverables in accordance with clause 3.3 of this Procedure. The assessment results shall be documented in an FT protocol using the form provided in Appendix 3 to this Procedure.

6.1.5. The Center for Diagnostics and Telemedicine sends a completed protocol to the Applicant's email address specified in the Application and (or) through the Applicant's personal user account on the Experiment's website www.mosmed.ai. In case of any critical discrepancies, the Applicant shall receive a completed Table 2 of Appendix 2 to this Procedure.

6.1.6. The discrepancies shall be considered critical if they directly or indirectly affect the life and health of patients, or negatively affect the clinician's work.

6.1.6.1. In the absence of critical discrepancies with the declared parameters listed in Table 1 of Appendix 2 to this Procedure, the Service shall proceed to the CT procedure outlined in clause 7 hereof.

6.1.6.2. In case of any critical discrepancies with the parameters stated in the Application and the Baseline Functional and Diagnostic Requirements for the Service deliverables specified in clause 3.3 of this Procedure, the Applicant to eliminate these discrepancies and notify the Center for Diagnostics and Telemedicine on the deadlines for such elimination.

6.1.6.3. Upon the elimination, the Applicant to complete a table of eliminated discrepancies (see Table 2 of Appendix 2) and submit it to the Center for Diagnostics and Telemedicine.

6.1.7. Should the Service require improvements that neither alter its initially declared functions and the technical architecture, nor affect the diagnostic accuracy metrics, the Applicant is allowed to proceed to the CT phase outlined in clause 7 of this Procedure that involves supervision over the Service improvements.

6.1.7.1. Should the Applicant make improvements that alter the initially declared functions and technical architecture, and affect the diagnostic accuracy metrics, the Service must undergo a repeated FT similar to the that outlined in clauses 6.1.1 - 6.1.5 hereof, once the Applicant submits a notification concerning the introduction of said changes.

6.2. The Applicant is allowed to pass the FT **not more than two times**. In case the repeated FT has been unsuccessful, the Service to be sent for a revision; the Service is allowed to re-apply for the FT **no earlier than 3 months** upon receipt of the last protocol with the unsatisfactory test results.

7. Calibration testing

7.1. CT of the Service operation is divided into the following stages:

7.1.1. THSC URIS UMIAS provides the Service with access to up to 100 studies within the framework of an experimental setup.

7.1.2. The Service shall analyze these studies one by one and submit the results to THSC URIS UMIAS.

7.1.3. The reading time for each study is recorded for further comparison with the standard reading time in accordance with clause 8.4 of this Procedure.

7.1.4. Upon returning the analysis results to THSC URIS UMIAS, Table 3 of Appendix 2 is to be completed with the study identification numbers, probability of a pathological finding in each study and the time the Service took to read each study.

7.1.5. Based on Table 3 of Appendix 2 to this Procedure, the Center for Diagnostics and Telemedicine to assess the CT deliverables for compliance with the following requirements:

- optimal activation threshold during the processing of the studies provided by URIS UMIAS (Youden index), based on the Report on the preliminary clinical and technical tests
- reading time for one study (as per clause 8.4 of this Procedure)
- diagnostic accuracy parameters as shown on the reference dataset in THSC URIS UMIAS.

The areas under the ROC curve (AUC) obtained with the reference dataset:

- computed tomography of the brain to detect ischemic stroke and intracranial hemorrhage – at least 0.81
- computed tomography and/or low-dose computed tomography of the chest to diagnose malignant neoplasms in the lungs – at least 0.91
- computed tomography of the chest to detect lung involvement associated with COVID-19 – at least 0.90
- computed tomography and / or low-dose computed tomography of the chest to detect impairment of lung airness, compression fractures of the vertebral bodies, ischemic heart disease (coronary calcium, paracardial fat), emphysema, thoracic aortic aneurysms which includes measurement of the thoracic aorta diameter, dilation of the pulmonary trunk which includes measurement of the pulmonary trunk diameter, free fluid (effusion) in the pleural cavities, enlarged lymph nodes (lymphadenopathy), pulmonary tuberculosis, sarcoidosis, bronchiectasis, adrenal lesions, rib(s) fracture, focal lesions in the chest bones – at least 0.81
- computed tomography of the abdomen to detect urolithiasis, adrenal lesions, liver lesions, renal lesions, compression fractures of the vertebral bodies, aneurysm of the abdominal aorta which includes measurement of the abdominal aorta diameter, focal lesions in the abdominal and pelvic bones, gallstone disease (calculous gallbladder) – at least 0,81
- magnetic resonance imaging of the brain to detect intracranial neoplasms and multiple sclerosis – at least 0.81
- magnetic resonance imaging of the cervical spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 0.81

- magnetic resonance imaging of the chest to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 0.81
- magnetic resonance imaging of the lumbosacral spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 0.81
- mammography to detect breast cancer – at least 0.81;
- head X-ray to detect sinusitis – at least 0.81;
- chest X-ray to detect various pathologies – at least 0.86;
- photofluorography of the lungs to detect various pathologies – at least 0.86;
- X-ray of the musculoskeletal system to detect various pathologies of the spine (osteochondrosis, scoliosis, spondylolisthesis, fracture of the vertebral bodies), fractures of the bones of the upper and lower extremities (wrist, shoulder, hip and ankle joints), arthrosis (hip, knee), longitudinal and transverse flat foot – at least 0.81;
- automated routine measurements during CT of the brain and abdomen (kidneys, liver, pancreas and spleen); automated routine MRI of the brain and pelvis (uterus and prostate gland) – at least 0.81.
- evaluation of chest CT images to detect various pathologies by a comprehensive Service – at least 0.81.

The reduction of diagnostic accuracy parameters as compared to the reference dataset must not exceed 10% of the values specified in the Report on preliminary clinical and technical tests.

For the numerical indicators evaluated by the Service, additional metrics can be applied to measure the diagnostic accuracy of the Service.

7.1.6. The assessment results shall be documented in an FT protocol using the form provided in Appendix 4 to this Procedure.

7.1.7. The Center for Diagnostics and Telemedicine shall submit a completed protocol to the Applicant's email address specified in the Application and (or) through the Applicant's personal user account on the Experiment's website www.mosmed.ai.

7.1.8. If the requirements outlined in clause 7.1.5 of this Procedure are met, the Applicant to receive a Participant status; a respective notification shall be sent to the email address specified in the Application and (or) through the Applicant's personal user account on the Experiment's website www.mosmed.ai. Afterwards, the Participant to proceed with the integration of the Service into the IHSC URIS UMIAS environment. Be advised, that IHSC URIS UMIAS to integrate the same version and/or modification of the Service as that indicated in the CT protocol. Altering the diagnostic accuracy and functionality of the integrated Service is possible only after finishing the procedure outlined in Section 12 hereof.

7.1.9. In case of any discrepancies with the requirements described in clause 7.1.5 hereof, the Center for Diagnostics and Telemedicine to send a notification to the Applicant's email address specified in the Application addressing possible improvements to the Service.

7.1.10. Upon revision of the Service, the Applicant to notify the Center for Diagnostics and Telemedicine by e-mail and then to proceed to a repeated CT procedure similar to the that outlined in clauses 7.1.1 - 7.1.8 hereof.

7.2. The Applicant is allowed to pass the CT **not more than two times**. If the repeated CT fails, the Service is to be sent for a revision. The Service is allowed to re-apply

for the CT **no earlier than 3 months** upon receipt of the last CT protocol with the unsatisfactory test results.

7.3. In case the repeated testing was unsuccessful (clauses 6.2, 7.2 of this Procedure), the Applicant may be offered to consider an alternative scenario for research and practical collaboration. Within the framework of such collaboration, the Applicant will be able to use the database containing de-identified (anonymized) study results, while the Center for Diagnostics and Telemedicine shall have the right to use the Service free-of-charge (with all its updates, improvements and modifications) on the territory of the Russian Federation on the terms of a simple non-exclusive license.

8. Processing of diagnostic studies in IHSC URIS UMIAS

8.1. Upon successful completion of FT and CT, the Service shall be integrated into IHSC URIS UMIAS. During the integration phase the following factors are taken into account:

- results of the THSC testing completed after 2022 that meet the requirements of clauses 3.3, 7.1.5 hereof
- Service status in IHSC in 2022.

8.2. The Center for Diagnostics and Telemedicine shall notify the Committee members about successful completion of the testing before integrating the Service into IHSC URIS UMIAS.

8.3. URIS UMIAS receives diagnostic studies from DDs. Using the URIS UMIAS technical capacities, the studies shall be de-identified, routed to the Service for the processing and then sent back to URIS UMIAS for reverse de-identification. Each study may contain a various number of images or scan sections depending on the imaging modality.

8.4. The standard reading time for one study is the time elapsed from the notification form USNEI that the study has become available for download and processing by the Service and until the Service alerts USNEI that the reading results have been uploaded to URIS UMIAS. The standard reading time must not exceed 6.5 minutes.

8.5. Upon completion of the integration into the IHSC URIS UMIAS environment, the Participant proceeds to the acceptance testing.

8.6. During the acceptance testing, each Participant shall receive enough studies to evaluate the Service's fitness for the operation testing. For each Service this stage shall take **at least three months** not including the periods of suspension outlined in clauses 9.10, 9.10.1, 9.10.4 hereof.

8.7. During the acceptance testing within each experimental setup, the studies shall be routed from all DDs to the Service using the "chessboard routing" approach.

"Chessboard routing" means automated submission of the studies to the Service to alternate interaction between the Service and all DDs within the framework of the corresponding experimental setup. To execute this scenario across all experimental setups, MFs shall be grouped depending the total number of DDs operating in a particular experimental setup. The time frame and order of the Service interaction with each group shall be determined by the Center for Diagnostics and Telemedicine.

8.8. Once the three-month-period of the acceptance testing is over, the automated submission of the studies is suspended until the Committee decides the Service is ready to proceed to the operational testing. The approval of the acceptance testing results will

depend on the THSC testing outcomes, the technical monitoring reports and the experts' clinical assessment of the Service performance.

Until the Committee makes a decision, each study shall be routed to the Service upon request from radiologists.

8.9. During the operation testing, the studies shall be routed to the Services from the DDs based on the following:

- the results of a quarterly survey of radiologists addressing their Service preferences for each experimental setup
- what Services the medical facilities pick through their user account on the official website of the Experiment www.mosmed.ai. A medical facility is allowed to switch to another Service within the framework of a particular experimental setup no more than once per month
- submission of an individual study to the Service at the radiologist's request.

8.10. In 1-2Q2023, during the operation testing the routing of the studies to the Services is carried out in accordance with the rules outlined in Appendix 6 to this Procedure; in 3-4Q2023, the routing shall be governed by the rules outlined in Appendix 7 to this Procedure.

8.11. The Center for Diagnostics and Telemedicine may consider engaging the Services that have made it to the operation testing into information, technological, scientific, practical and other forms of collaboration within the Experiment on the terms and conditions established by the respective agreements. Moreover, the same version and modification of the Service is allowed to participate in said collaborations within the Experiment.

9. Monitoring the technical parameters of the Service operation

9.1. During the acceptance and operation testing phases, the Center for Diagnostics and Telemedicine monitors the technical parameters of the Service operation.

9.2. The monitoring shall be carried out by means of retrospective review of the studies analyzed by the Service in accordance with the data received from URIS UMIAS. A monitoring report shall be drawn up based on the monitoring results (Appendix 5 to this Procedure).

9.3. Monitoring of the technical parameters of the Service operation within the framework of information, technological, scientific, practical and other forms of collaboration within the Experiment shall be carried out based on the data generated by the HUB Telemed software (hereinafter referred to as "HUB Telemed") developed by the Center for Diagnostics and Telemedicine. The monitoring results shall provide the basis for the decision-making with regard to the criteria established by the respective collaboration agreements.

The monitoring for the following defects shall be carried out for all studies analyzed by the Service over the reporting period:

"a": the reading time for one study exceeds 6.5 minutes;

"b": no analyzed studies uploaded to IHSC URIS UMIAS according to the URIS UMIAS data.

At least 80 studies processed over the reporting period will be sampled to check for the following defects:

"c": the Service does not function as expected, which complicates the radiologist's job or makes it impossible to maintain its quality; subcategories:

c1 - no additional series
c2 - no DICOM SR
c3 - 2 or more DICOM SRs
c4 - Service name is not indicated
c5 - Service version is not indicated.
“d”: defects related to the display of the region of interest:
d1 - images in the additional series are cropped
d2 - brightness/contrast of the additional series does not match the original image
d3 - not all the images are analyzed
d4 - no warning label “Academic purpose only”
d5 - changes in the original series.
“e”: Other violations of the integrity and study results that make the diagnostic interpretation impossible, including:
e1 - labelling is outside the target organ
e2 - an incorrect anatomical region / projection / series was analyzed.

9.3.1. Clinical evaluation is generated by assessing the localization of findings (labelling) and drawing up a report (conclusion) for each study from the sample. The clinical evaluation of the Service's performance is an arithmetic mean of the labelling and conclusion scores. The labelling and conclusion score is an arithmetic mean of all scores in the sample expressed as a percentage, according to the criteria listed in the Table of criteria and indicators for clinical evaluation.

Criteria and parameters for clinical evaluation

Evaluation criteria	Report	Scores	
		labelling	conclusion
Full match	All target findings are labelled or “normal”	1	1
Incorrect assessment (partial/excessive)	More than 1 finding of the target pathology is observed (or finding is not seen on all images/projections). Inaccurate contouring of findings. Incorrect estimation of the volume/number of findings.	0.5	0.5
False-positive	False/odd finding. The finding is observed although the target pathology is not present.	0.25	0.25
False-negative	Missing finding Not a single finding of the target pathology was observed, when they are actually present.	0	0

The Center for Diagnostics and Telemedicine determines the thresholds for the clinical parameters of the Service operation for each experimental setup. Acceptable clinical parameters of the Service operation:

- computed tomography of the brain to detect ischemic stroke – at least 50%
- computed tomography of the brain to detect intracranial hemorrhage – at least 75%
- computed tomography and/or low-dose computed tomography of the chest to detect malignant neoplasms in the lungs – at least 55%
- computed tomography of the chest to detect lung involvement associated with COVID-19 – at least 53%
- computed tomography of the chest to detect compression fracture of vertebral bodies – at least 90%
- computed tomography of the chest to detect free pleural fluid (effusion) – at least 77%
- computed tomography of the chest to detect emphysema – at least 85%.
- computed tomography of the chest to detect adrenal lesions – at least 86%
- computed tomography and/or low-dose computed tomography of the chest to detect ischemic heart disease (coronary calcium) – at least 76%
- computed tomography and/or low-dose computed tomography of the chest to detect ischemic heart disease (paracardial fat) – at least 90%
- computed tomography and/or low-dose computed tomography of the chest to detect aneurysm and measurement of the thoracic aorta diameter – at least 90%
- computed tomography and/or low-dose computed tomography of the chest to detect dilation of the pulmonary trunk and measurement of the pulmonary trunk diameter – at least 88%
- computed tomography and/or low-dose computed tomography of the chest to detect impairment of lung airness, enlarged lymph nodes (lymphadenopathy), pulmonary tuberculosis, sarcoidosis, bronchiectasis, rib(s) fracture, focal lesions in the chest bones – at least 76%
- computed tomography of the abdomen to detect adrenal lesions – at least 86%
- computed tomography of the abdomen to detect compression fracture of vertebral bodies – at least 90%
- computed tomography of the abdomen to detect aneurysm and measurement of the abdominal aorta diameter – at least 90%
- computed tomography of the abdomen to detect urolithiasis, liver lesions, renal lesions, focal lesions in the abdominal and pelvic bones, gallstone disease (calculous gallbladder) – at least 50%
- magnetic resonance imaging of the brain to detect intracranial neoplasms and multiple sclerosis – at least 50%
- magnetic resonance imaging of the cervical spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 50%
- magnetic resonance imaging of the thoracic spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 50%
- magnetic resonance imaging of the lumbosacral spine to detect focal lesions in the vertebral body, protrusions and herniations of the intervertebral discs, and spinal stenosis – at least 59%

- mammography to detect breast cancer – at least 80%
- head X-ray to detect sinusitis – at least 50%
- chest X-ray to detect various pathologies – at least 78%
- photofluorography of the lungs to detect various pathologies – at least 77%
- X-ray of the musculoskeletal system to detect various pathologies of the spine (osteochondrosis, scoliosis, spondylolisthesis, fracture of the vertebral bodies), fractures of the bones of the upper and lower extremities (wrist, shoulder, hip and ankle joints), arthrosis of the hip, transverse flat foot – at least 50%
- X-ray of the musculoskeletal system to detect arthrosis of the knee – at least 90%
- X-ray of the musculoskeletal system to detect longitudinal flat foot – at least 90%
- automated routine measurements during CT of the brain and abdomen (kidneys, liver, pancreas and spleen); automated routine MRI of the brain and pelvis (uterus and prostate gland) – at least 52%
- evaluation of chest CT images to detect various pathologies by a comprehensive Service – at least 88%

9.3.2. At least 80 studies are sampled if there are 80 or more studies without “a” and “b” type defects over the reporting period. Otherwise, only studies with the “a” and “b” defects should be monitored.

9.4. If, upon receiving a study, the Service finds it impossible to process, the Service to return a corresponding error notification. The list of errors is available in the Baseline Functional Requirements for the Service results, published on www.mosmed.ai. These studies shall be considered unprocessed.

9.4.1. If the error is not caused by a failure in the Service's and (or) the Participant’s infrastructure, these studies shall be not be included in the list of studies with defects.

9.4.2. If the error is caused by a failure in the Service's and (or) the Participant’s infrastructure, these studies shall be considered defective; the proportion of defects is estimated relative to the total number of studies submitted to the Service over the reporting period.

9.5. The monitoring is performed on a monthly basis until the end of the Service's participation in the Experiment. The reporting period of the Monitoring lasts one calendar month. An interim monitoring report on the “a” and “b” defects is drawn up on the Days 10 and 20 of each month and forwarded to the Participant.

As an exception, the Center for Diagnostics and Telemedicine of the Moscow Health Care Department may change the reporting period due to changes in the organizational and technical conditions of the Experiment.

9.6. The studies processed over the reporting period shall be considered analyzed if they have no “a”-“e” defects as per clause 9.3 hereof and for which no error notifications were sent as per clause 9.4 hereof.

9.7. Based on the outcomes of the monthly monitoring, the Center for Diagnostics and Telemedicine to report one of the following conclusions:

- participation in the Experiment continues
- Participant is requested to improve the Service
- participation in the Experiment is suspended until the Service is improved.

9.8. Participation in the Experiment shall continue once all of the following conditions are met:

- all studies analyzed over the reporting period contain less than 10% of “a” and “b” defects
- less than 10% of studies from the sample contain “c”-“e” defects;

- clinical evaluation score meets the requirements of clause 9.3.1 hereof.

9.9. The operation of the Service should be improved if at least one of the following conditions is met:

- all studies analyzed over the reporting period contain more than 10% of "a" and "b" defects
- less than 10% of studies from the sample contain the "c"-"e" defects
- clinical evaluation score does not exceed the threshold value established for a particular experimental setup.

9.10. Participation in the Experiment shall be suspended if at least one of the following conditions is met:

- all studies analyzed over the reporting period still contain more than 10% of "a" and "b" defects in comparison with the Monitoring results for the previous reporting period
- all studies analyzed over the reporting period still contain more than 10% of "c"-"e" defects in comparison with the Monitoring results for the previous reporting period
- clinical evaluation score repeatedly failed to exceed the threshold established for the experimental setup in comparison with monitoring results for the previous reporting period.

9.10.1. Participation of the Service shall be suspended until the causes of the defects in Service operation are removed.

9.10.2. To make a decision to suspend the participation in the Experiment, the Center for Diagnostics and Telemedicine shall take the following documents into account: monitoring reports for the current and previous (if any) periods, a response from the URIS UMIAS operator confirming the number of studies with defects (if any), and a Participant's response to the inquiry regarding the causes of defects (if any).

9.10.3. If the Service is suspended during operation testing, the studies that were supposed to be routed to it shall be evenly redistributed among the other Participants within the same experimental setup.

9.10.4. After removing the causes of defects, the Service may re-join the Experiment following the review of the request for changing the Service's status as per clause 12 hereof.

9.11. The Participant has the right to approach the Center for Diagnostics and Telemedicine to explain the decision specified in clause 9.7 hereof.

10. Comprehensive Service

10.1. The provisions of this section guide the participation of a CS in the Experiment. The participation in the Experiment in respect of the "Chest XR for detecting various pathologies" and "Photofluorography of the lungs to detect various pathologies" setups shall comply with the sections 4–9 hereof.

10.2. The list of target pathologies (mandatory and optional) for the CS shall be specified in the Application for participation in the Experiment (see Appendix 1). The requirements to the CS deliverables for each target pathology are set out in the Baseline Diagnostic Requirements for Service Deliverables published on the website www.mosmed.ai.

10.3. The Participant may present only one CS for each experimental setup. If the Service is a part of a CS and analyze studies within one experimental setup, it shall be excluded from any routing except for the automated routing to the CS.

10.4. The CS Application to participate in the Experiment complies with the procedure outlined in clause 4 hereof. The CS shall be granted the Participant status if it is found capable of detecting every target pathology specified in the Application (see Appendix 1).

10.5. The Functional testing (FT) of the CS operation is performed in accordance with the FT procedure set out in clause 6 hereof. The number of studies submitted to CS depends to the set of pathologies covered by the CS.

10.6. CT of the CS operation shall be performed separately for each target pathology as per clause 7 hereof. As per clause 7.1.6, a unified protocol shall be drawn up containing an individual conclusion for each target pathology based on the CT outcomes.

10.7. The CS is transferred to the IHSC environment once the test results for all mandatory pathologies are satisfactory.

10.8. The list of the analyzed pathologies available to the Service in IHSC may be expanded after reviewing a Request for Service Improvement submitted by the Participant to the Center for Diagnostics and Telemedicine. The procedure for requesting changes in the Service is set out by clause 11 hereof.

10.9. The processing of the diagnostic studies in IHSC URIS UMIAS by a CS shall be carried out in accordance with clauses 8.1–8.10 hereof.

10.9.1. During the operation testing phase, the studies shall be submitted to the CS in accordance with the routing rules set out in Appendices 6–7 hereto.

10.9.2. The technical parameters of the CS operation are monitored in accordance with clause 9 hereof.

11. Changes to the Service

11.1. Should the Service require any changes, the Participant shall submit a formal request to the Center for Diagnostics and Telemedicine notifying about the modifications or the version change and their consequences (see Appendix 9 hereto), but not more often than once per 1,5 months. When the changes affect the measurable performance parameters, their values before and after the changes should be specified in the numerical form, where:

- modifications are any patches in the software and infoware of the Service reportable to URIS UMIAS, which affect operational characteristics only without adjusting the functionality; this also includes changes that eliminate errors and improve the Service algorithm
- version change is any change in software and infoware reportable to URIS UMIAS that allow to perform the declared or extra functions and secure the transition to new operating systems and IT environment.

11.2. This Procedure is aimed to ensure transparency of the Service application in the real-world environment. Transparency involves notifying the interested parties, including the Center for Diagnostics and Telemedicine about any changes (modifications, version changes) to the Service. The Participant must ensure, that the Service change request fully and accurately describes the proposed modifications or changes that come with the new version.

11.3. The Center for Diagnostics and Telemedicine shall review the request and then allow / forbid to change the current version or make modifications to the Service, and carry out any additional testing (one or more stages) within the Experiment to verify the

changes to the Service. The request shall be reviewed within **10 (ten) business days**. The review period may be extended if the request is incomplete or require clarification to determine how the changes should be tested.

11.4. In case of additional testing due to modifications / version change, a report presenting the outcomes and a decision on whether the changes are possible / impossible shall be issued.

12. Withdrawing from the Experiment

12.1. The Participant may abandon the Experiment on their own initiative by sending a corresponding notification to the email of the Center for Diagnostics and Telemedicine.

12.2. The participation of the Service in the Experiment may be terminated by the Moscow Health Care Department Committee if within 6 months from either the date of suspension of the Service from the Experiment or the receipt of an unsatisfactory test results protocol, the Participant / Applicant fails to inform about the deadlines for the revision of the Service or provides a written refusal to improve the Service.

13. Grants

13.1. The Participant may submit a grant application to the Department not more than once a month starting from the date when either IHSC URIS UMIAS receives the studies from DDs during the acceptance and operation testing phases, or HUB Telemed does so as part of the information, technological, scientific, practical and other forms of collaboration within the Experiment.

13.2. Based on the monitoring report, the Center for Diagnostics and Telemedicine shall recognize the studies as analysed as per clause 9 hereof. If the Participant is involved in information technological, scientific and practical, and other forms of interaction and collaboration within the Experiment, the analyzed studies shall be summarized based on monitoring reports generated following the agreements governing such interaction and collaboration.

13.3. Based on the monitoring reports issued by the Center for Diagnostics and Telemedicine, the Committee shall approve the number of analyzed studies to estimate the grant amount.

13.4. The grant amount shall be determined by multiplying the number of the analyzed studies by the analysis cost established in Appendix 5 hereto.

14. Clinical and technical testing

14.1. The Participant may request the Center for Diagnostics and Telemedicine to carry out clinical and technical testing of the Service in order the to register software as a medical device.

14.2. The terms and conditions for the clinical and technical testing, the corresponding application form and the list of supporting documents are published on www.mosmed.ai.

15. Evaluation and comparative analysis of the Services.

15.1. During the Experiment, the Center for Diagnostics and Telemedicine shall evaluate quantitative and qualitative parameters of each Service and perform a general

comparative analysis of the Services. The evaluation shall be based on the data received during the testing and monitoring of the Service. The evaluation to address changes in the certain parameters over a selected period of time.

15.2. The results of the Experiment shall be structured and analyzed using data obtained over the calendar year **no later than April 1 of the year following the reporting one**.

15.3. The evaluation shall be carried out, inter alia, in accordance with Guidelines No. 43 “Clinical Acceptance of Software Based on Artificial Intelligence Technologies (Radiology)” (recommended by the Expert Council for Science of the Moscow Health Care Department, Protocol No. 8 of June 25, 2019).

15.4. The main criteria for evaluation and comparative analysis of the Services:

- diagnostic accuracy
- applicability
- ease of use
- operation efficiency (including the impact on labor productivity and the reading time).

15.5. The criteria shall apply:

- to each Service individually
- to all Services that analyze the studies within the given experimental setup
- to all Services participating in the Experiment.

15.6. The results of evaluation and comparative analysis shall be documented by the Center for Diagnostics and Telemedicine of the Moscow Health Care Department and presented as a final report for the corresponding calendar year.