PUBLIC HEALTH CENTER
OF THE MOH OF UKRAINE

Report on the results of the study

«Evaluating the performance of computer detection software implementation in triaging chest X-ray images in TB screening program in Ukraine»

Kyiv – 2024

Authors: Mariia Moshura, Yana Terleeva, Olena Nesterova

Suggested citation: *Moshura M., Terleeva Y., Nesterova O. (2024) REPORT on the study results «Evaluating the performance of computer detection software implementation in triaging chest X-ray images in TB screening program in Ukraine»*

In accordance with the Charter of the State Institution «Public Health Center of the Ministry of Health of Ukraine», approved by the Order of the Ministry of Health of Ukraine No. 224 of February 09, 2024,

The State Institution «Public Health Center of the Ministry of Health of Ukraine» *is a sanitary and preventive health care institution established by the Order of the Ministry of Health of Ukraine No. 604 of 18.09.2015 pursuant to the Order of the Cabinet of Ministers of Ukraine No. 909-p of 02.09.2015.*

PHC performs the functions of the main institution of the Ministry of Health of Ukraine in the field of public health and combating dangerous, especially dangerous, infectious and non-communicable diseases, biological protection and biological safety, immunoprophylaxis, laboratory diagnosis of infectious diseases, hygienic and microbiological aspects of human health related to the environment, national coordinator of the International Health Regulations, and coordination of emergency response in the public health sector.

This study was conducted as part of the project «Support to NTP activities for using digital technologies to support the TB response in Ukraine» with financial and technical assistance from «TDR. For research on diseases of poverty», which is co-sponsored by the United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP), the World Bank and the World Health Organization (WHO).

Re-publication of this text, images from it and/or data, in whole or in part, under the name of a person other than the author, infringes copyright and constitutes plagiarism, which is subject to three types of liability: civil (compensation for damages and/or non-pecuniary damage; recovery of income), administrative (payment of a fine) and criminal (imprisonment for up to 2-5 years) liability.

Plagiarism includes 1) using the text or graphic data of another author without changes, without citation and without attribution (copy & paste plagiarism), 2) combining different fragments of texts or sentences to form a new text without citation, thus presenting it as one's own thoughts (shake & paste plagiarism), 3) presenting the ideas of another author in one's own words without citing the source (idea plagiarism), and 4) translating an original text from another language without citing the source (translation plagiarism).

All quotes in the report are presented in the original language and translated into English.

Contents

Acronyms

CAD	computer-aided detection
Sen	sensitivity
Spe	specificity
WHO	World Health Organization
EHIF	Estonian Health Insurance Fund
HCF	healthcare facility
CXR	chest X-ray
TP	true positive
TN	true negative
MOH	Ministry of Health of Ukraine
NPV	negative predictive value
NHSU	National Health Service of Ukraine
NGO	non-governmental organization
PPV	positive predictive value
TB	tuberculosis
FN	false negative
FP	false positive
PHC	Public Health Center of the Ministry of Health of Ukraine

Introduction

As part of the global effort to fight tuberculosis, Ukraine has programs and activities for prevention, diagnosis and treatment. Medical facilities provide free antibiotic therapy to patients with TB, which helps control the spread of the disease. However, despite these efforts, TB management and control remain important challenges for the Ukrainian healthcare system. Early detection strategies and treatment adherence are key aspects in the fight against the disease.

Systematic screening for TB in the healthcare sector improves the epidemiological situation with regard to TB by increasing the number of active cases detected, reducing the prevalence of the disease, reducing the spread of infection, and preventing new cases and relapses of TB.

Chest X-ray called a CXR is a TB screening tool. This method is used to screen for TB among people at high risk of contracting the disease, including those who work in healthcare and among populations with high prevalence of TB. X-rays are also used to screen people who have certain symptoms or diseases that may cause TB or TB contacts.

In large medical institutions and clinics in Ukrainian cities, it is possible to use CAD systems can be used to diagnose tuberculosis at early stages and monitor treatment progress. However, it is important to take into account the availability of technology and the financial capacity of medical institutions to implement such systems. In general, CAD systems can be used to automate the analysis of medical images, detect pathologies, and plan treatment. The use of CAD systems for medical purposes can improve the quality of diagnosis and treatment of tuberculosis and other diseases, but it requires investment in technology and training of medical staff.

Modern diagnostic technologies and effective TB treatment regimens contribute to the fight against the disease in a number of ways. In general, modern diagnostic and treatment technologies increase the chances of early detection of TB and effectively treat TB, contributing to a reduction in disease and an increase in quality of life.

WHO recommends[1] replacing CAD reading human results in two broad contexts: for screening and for medical triage. In both cases, the ultimate goal is the same, namely to use CAD to determine whether a patient should undergo confirmatory diagnostic tests.

In the context of screening, CAD can be a valuable tool for screening people with or without TB symptoms or significant TB risk factors, both in prevalence studies and in active case-finding situations. Screening often involves identifying early-stage disease among populations with low TB risk and/or inadequate access to health care. In the context of medical triage, triage tests are used in people with symptoms, signs, risk markers, and/or test results suggestive of TB, typically among those who seek care at or are referred to health facilities for screening or contact investigations.[2]

---

[1] World Health Organization. Global Tuberculosis Report 2020.; 2020. http://apps.who.int/bookorders. Accessed March 26, 2021.

[2] Implementation guide of computer aided detection (CAD) technology with ultraportable X-ray devices to screen and triage TB,

Therefore, CAD systems use artificial intelligence to analyze CXRs for any abnormality that may indicate pulmonary TB. CAD assesses these abnormalities and determines whether additional diagnostic tests are needed to confirm TB, given a certain threshold.

CAD technology can improve the efficiency and performance of X-ray when used for TB screening and triage. In the Ukrainian context, it can be helpful in a TB control program, enhancing TB screening capabilities. This technology can replace or complement human reading of X-ray images during TB screening and help to avoid differences in interpretation between assessors, as well as reduce delays in X-ray image assessment, especially in situations where there is settings leccking qualified personnel.

Currently, Ukraine is considering introducing innovative technologies to reduce TB in the country. To this end, a study was implemented to determine the optimal models for implementing CAD in pilot Ukrainian regions.

1. Healthcare facilities survey component of the study

In general, the assessment of the radiologist workload is important for improving working conditions, ensuring the quality of medical diagnosis and improving the overall quality of healthcare services. The situation with the assessment of healthcare facilities may vary depending on many factors. Here are some key aspects of this situation:

- Patient flow: The number of patients requiring X-ray examinations can vary depending on the time of day, day of the week, and season. Referrals of patients for further examination can also affect this flow.
- Equipment and infrastructure: The availability and condition of equipment, accessibility of workstations and offices, technical staff to maintain the equipment, and the speed and efficiency of X-ray diagnostic centers can affect the radiologist workload.
- Staffing: The number of radiologists in a healthcare facility and their skills play an important role. Inadequate staffing or inexperienced doctors can lead to an increased workload.
- Legislation and regulations: Rules and regulations related to the use of X-ray equipment can affect the workflow of radiologists and the workload they experience.
- Management and scheduling system: An effective planning and time management system for radiologists can help optimize workload and distribute working hours based on patient flow and ensure that the appropriate amount of time is spent on each examination.
- Technological innovations: The introduction of new technologies, such as digital X-ray diagnosis or artificial intelligence, can improve the efficiency and accuracy of diagnosis, but also require physician adaptation and additional training.

Accordingly, the overall situation with the assessment of the radiologist workload requires constant monitoring, planning and adaptation to ensure quality medical diagnosis and safety for both patients and healthcare professionals.

The objective of this study component was to assess current radiologist workload and identify gaps in the TB screening process. Additionally, the study aimed at identifying barriers and enablers for CAD implementation in Ukraine. With these goals in mind, two healthcare facilities in Lviv and Sumy Regions, respectively, were selected to collect data.

Lviv oblast is one of the leaders in the fight against tuberculosis in Ukraine. For many years, innovations and new approaches have been developed by the example of the Lviv region for further implementation of systemic solutions and policies at the level of the national program to combat tuberculosis. Currently, Lviv oblast is the region that has sheltered the largest number of internally displaced persons and individuals who require radiological examination as a risk group for the development of TB. So, in Lviv Region, the Lviv Regional Phthisiopulmonology Clinical Treatment and Diagnostic Center «Pulmonary Health Center» of the Lviv Region Council was involved, which is one of the leading facilities in the implementation of systemic policies of the TB program.

As for Sumy region, this oblast has sufficient organizational potential and focus on improving the quality of medical services. Fierce fighting has been going on for a long time, and most of the territories have been occupied, causing an acute humanitarian crisis. A significant number of patients with tuberculosis were not able to continue treatment due to difficulties in visiting a health care facility, changing their place of stay or restrictions in the work of the TB service. The issue of rapid and high-quality diagnosis of TB has become urgent in oblast, thus the utilization of artificial intelligence algorithms will be helpful for that. So, in the Sumy Region, the «Regional Clinical Phthisiopulmonology Medical Center of the Sumy Region Council» was selected for the study, which also has sufficient capacity and quality improvement in healthcare services.

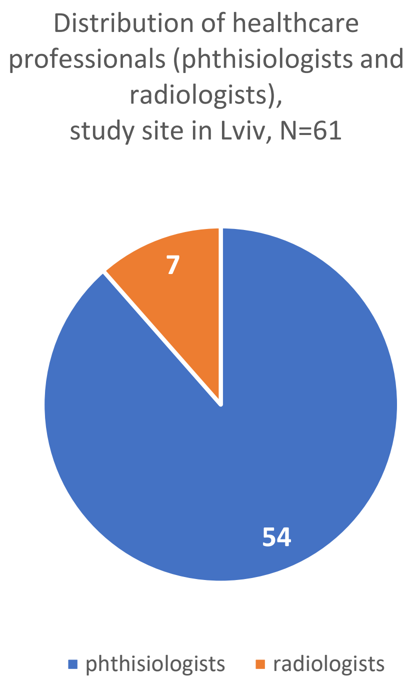
To achieve the above goal, data was collected using a questionnaire developed for a healthcare facility, which includes questions about:

1. Human resources of the facility:
    a. Number of radiologists and phthisiologists in the facility;
    b. The number of hours radiologists and phthisiologists work per day;
    c. The average number of patient consultations and radiographs reviewed per day by a radiologist and a phthisiologist;
    d. The average time required for a radiologist and a phthisiologist to read, interpret and formulate a conclusion on 1 radiograph;
    e. Availability of a full-time IT specialist in the facility who can be involved in the installation of CAD;
    f. Availability of unlimited Internet access with the facility type of connection.

2. Technical equipment of the facility:
    a. Availability of computers and Internet access at the workplace;
    b. Availability of specialized software for electronic document management, patient flow management, and specialized registries;
    c. Availability of reliable sources of emergency power;
    d. Availability of specialized computerized diagnostic equipment;
    e. Planning the procurement of specialized computerized diagnostic equipment during 2023.

The questionnaire was sent to the heads of the facilities and was filled out by self-completion.

Personnel potential of the HCF



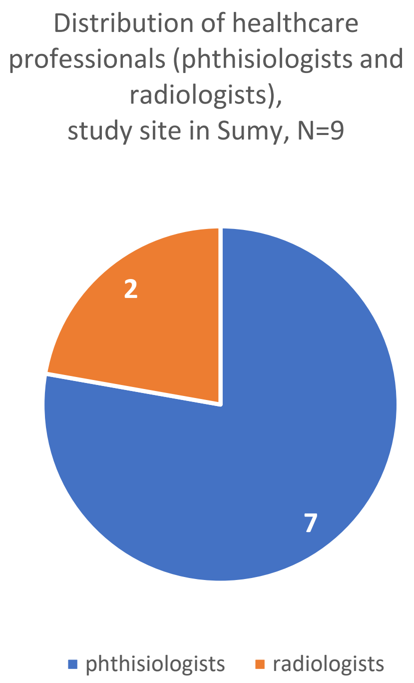
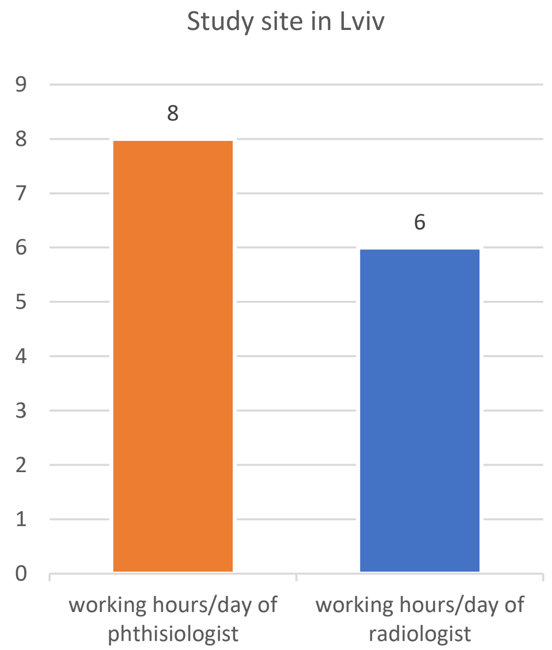


Figure 1. Number of doctors working in a facility

At the time of the study, the Pulmonary Health Center employed 54 phthisiologists, of whom 13 were of pre-retirement and retirement age. In addition, 54 phthisiologists account for 57.25 staff positions (Figure 1).

As for the Regional Clinical Phthisiopulmonology Medical Center of the Sumy Regional Council, 2 out of 17 phthisiologists are of pre-retirement age and retired. In addition, there are 19 staff positions for 17 phthisiologists.



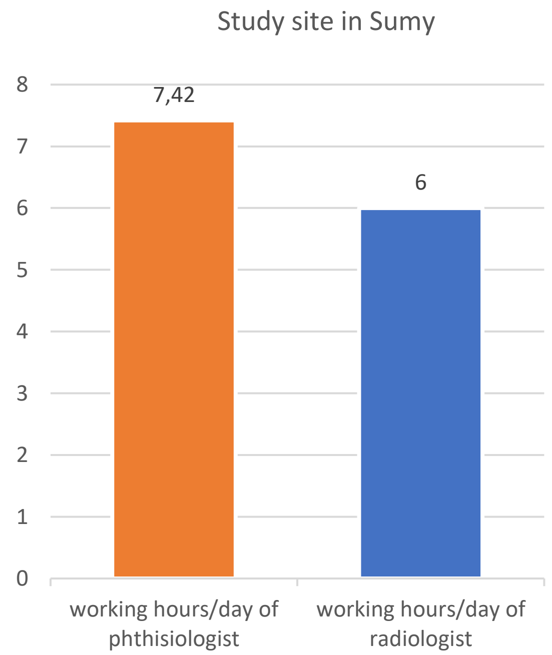


Figure 2. Number of working hours/day of doctors

Regarding the number of hours of work of doctors, we can see that the distribution of hours is almost the same in both study sites. Only for phthisiologists there is a slight difference - phthisiologists in the Sumy study site have a working day of 7.42 hours (Figure 2).

The estimated working day of radiologists and phthisiologists is established under Order No. 33 dated 23/02/2000 «On staffing standards and standard staffing levels of healthcare facilities» and Order of the MOH No. 340 of 28.11.1997 «On Improving the Organization of Radiation Diagnostics and Radiation Therapy Services».

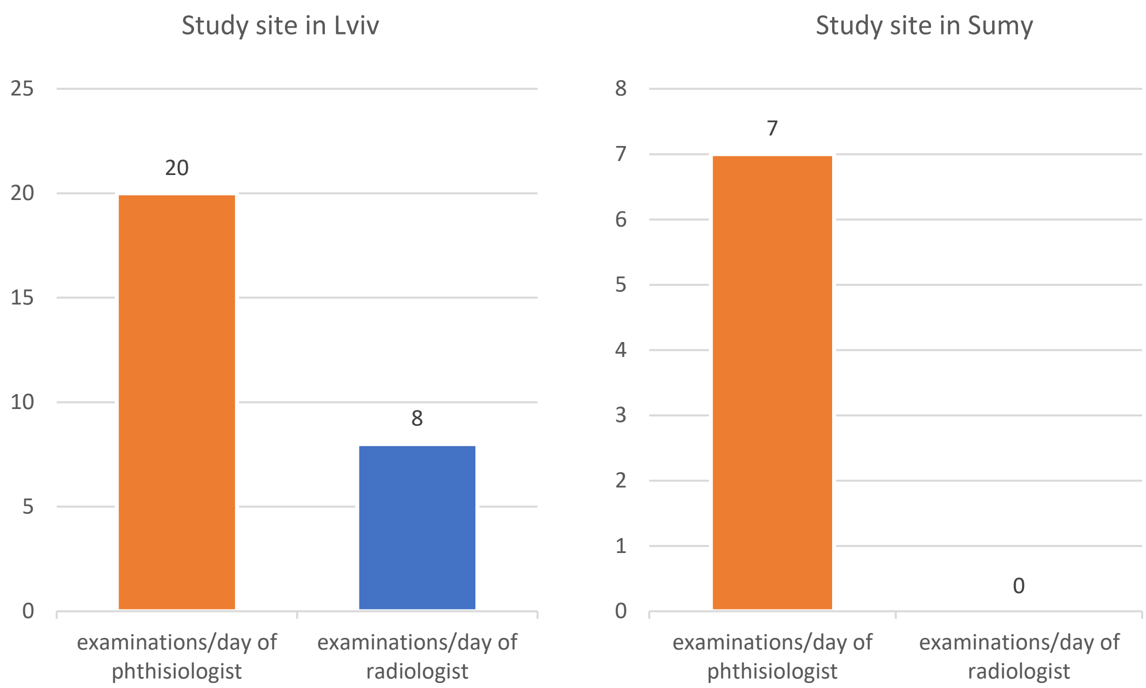


Figure 3. The average number of examinations (consultations) of patients

The average number of patients (per doctor) seen by phthisiologist in the Outpatient Department of the Pulmonary Health Center is 20 patients per day. As for the Sumy study site, the average number of examinations or consultations is lower – only 7 patients per day. This indicates a certain level of patient flow and workload on healthcare workers in different regions of the country (Figure 3). It is worth noting here, that in Ukraine, there is an order of the Ministry of Health of Ukraine No. 427 dated 11.05.2016 «On Amendments to the Order of the Ministry of Health of Ukraine No. 33 dated 23.02.2000 «On Staffing Standards and Standard Staffing of Healthcare Institutions»», according to which the estimated workload of a radiologist per shift in the office with a working day of 6 hours (five-day week) is 7-8 patients.

This flow may be influenced by the full-scale invasion of Ukraine, which has provoked mass migration to the western part of the country. In 2022, Lviv region had an 11.0% increase in the incidence of TB compared to 2021. Taking into account data from departmental services and deaths that were not registered, the incidence rate of registered new cases + relapses was 55.2 per 100 thousand people, compared to 49.6 in 2021. In addition, since the beginning of the war, Lviv region has received the largest number of internally displaced persons, which could also lead to an increase in the flow of patients in need of X-ray diagnostics.

As for the investigation of the Sumy study site, a full-scale invasion of the Sumy region could also have a significant impact on the provision of medical services in the region, particularly with regard to the use of X-ray examinations. In connection with the significant displacement of residents and the arrival of the wounded, many medical facilities became overcrowded. This led to difficulties in providing intramural care, as well as to an increase in the risk of infection of relevant diseases due to close contacts between patients. Patients who have undergone treatment for certain diseases, such as TB, have become aware of difficulties in continuing treatment due to displacement, restrictions on health services and security threats. This led to the risk of exacerbation of diseases and the possibility of becoming a source of infection for other people.

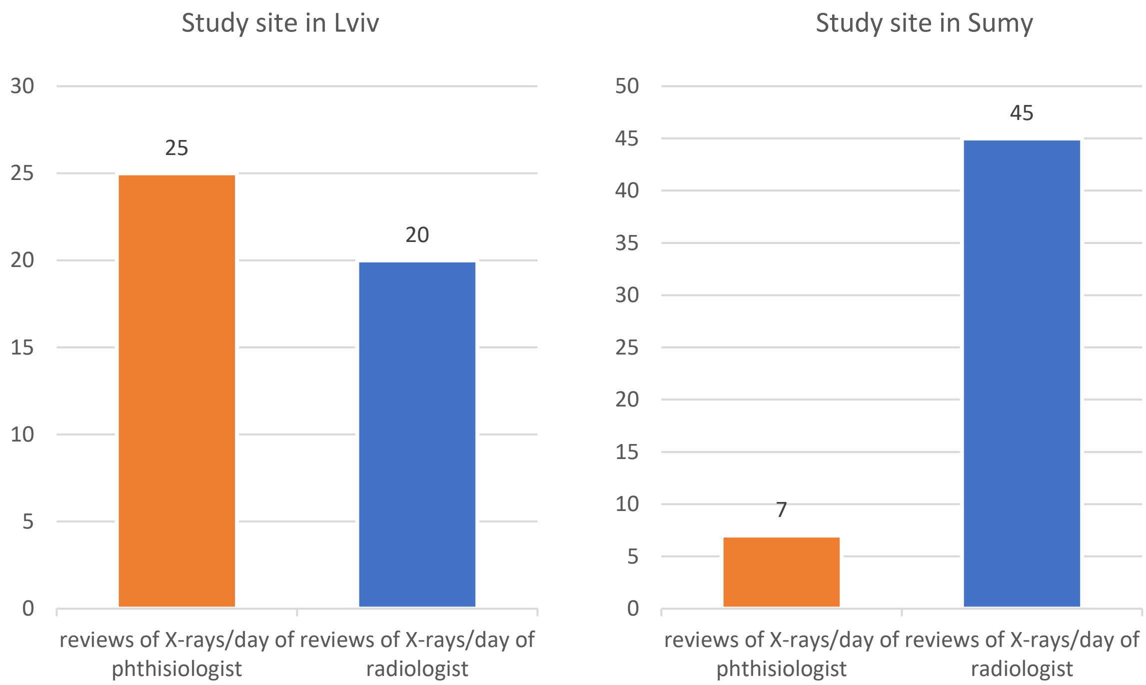


Figure 4. The average number of reviews of radiographs

However, regarding the average number of X-ray examinations, we can see that the radiologists of the Sumy study site process more radiographs per day than at the Lviv study site (45 X-rays compared to 20 per day). In addition, the phthisiologists of the Sumy study site process only 7 X-rays out of 45 per day on average (Figure 4).

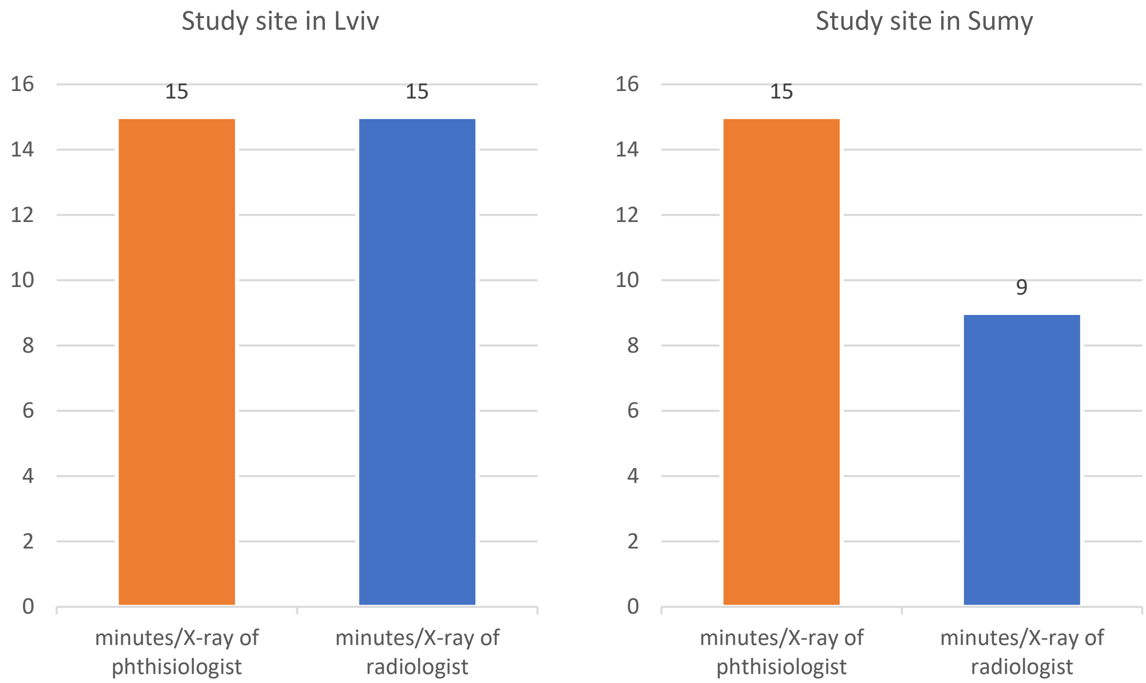


Figure 5. The average time required for a doctor to read, interpret and form an opinion based on one X-ray

The average time required for a physician to read, interpret, and form an opinion based on a single X-ray image did not differ significantly between study sites. Although we can see that the average time taken by a radiologist to process one X-ray in the Sumy study site is less than that of the phthisiologist at this site, and less than that of the radiologist and phthisiologist at the Lviv study site (Figure 5).

Most of the time is spent on forming a conclusion based on the results of an X-ray.

For the period of data collection of the study, over the past 6 months, both survey sites (Lviv and Sumy) didn't have any cases when a facility didn't have a single doctor on site for at least one day, namely a phthisiologist or radiologist.

Also, both study sites have a responsible IT-specialist on staff. The study sites also employ software engineers who are responsible for the information security of facilities and everything related to computer equipment and software, and may well be involved in the implementation and installation of CAD at the facilities.

Technical equipment of facilities

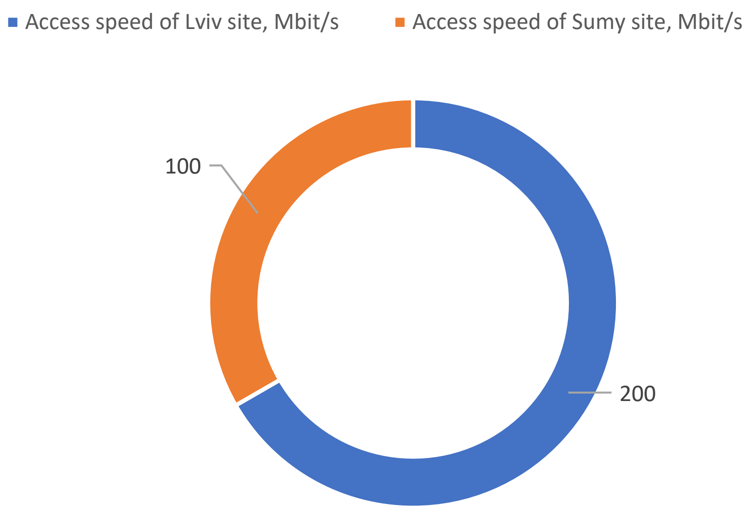


Figure 6. Availability of unlimited Internet access in the facility

Unlimited Internet access is available at both study sites. However, the chart demonstrates the difference in access speed.

Despite the fact that both study sites have the same Internet connection types - dedicated line (PPPoE, etc.) and satellite (Starlink, etc.), the access speed at the Pulmonary Health Center is twice as higher (Figure 6).

All phthisiologists and radiologists are also provided with Internet access at their workplaces and have work computers, including laptops. This ensures a continuous workflow. Indeed, all phthisiologists and radiologists (at their own expense and at the expense of the projects) are provided with the necessary modern computer equipment, namely laptops, monoblocks, multifunctional devices, routers, and uninterruptible power supplies. The study sites also have a powerful generator that, in the event of a power outage or other emergencies, quickly turns on and provides power to the entire institution, including X-ray equipment and a computed tomography scanner. There are also no problems with Internet access, as the sites are equipped with the Starlink system.

The study sites in Lviv and Sumy are equipped with specialized software, including medical information systems, to maintain electronic document management, manage patient flows, and maintain specialized registries. The Pulmonary Health Center has information systems such as RIS (software and hardware complex) – an X-ray information system, thanks to which all images are stored in one place, and all doctors at their workplaces have access to the database of X-rays and CT scans online around the clock, which significantly improves the workflow. The main medical information system Askep.net is also in operation to maintain all medical records (medical history, patient records in the institution, etc.) for interaction with the EHIF and the NHSU. In addition, the study sites have an information system for socially significant diseases, which has been in operation since May 2023, replacing the Tuberculosis Register to record patients with TB and HIV, Megapolis, an electronic document management system between institutions, MedControl, to maintain statistical information on inpatients, and MIS Kashtan, to keep records of medicines in the institution.

To ensure continuous operation, the study sites in Lviv and Sumy have reliable sources of emergency power supply (generators, high-capacity portable energy storage devices, etc.). For example, the Pulmonary Health Center has autonomous power sources that allow for uninterrupted power supply for the institution's round-the-clock operation in the absence of centralized power supply, namely 3 generators: a main diesel generator with a capacity of 308 kW, a diesel generator with a capacity of 30 kW, and a gasoline generator with a capacity of 60 kW. Portable uninterruptible power supplies are also installed at each doctor's workplace.

As of the data collection period, the study sites are provided with specialized computer diagnostic equipment, namely (Table 1):

Table 1. Availability of specialized computerized diagnostic equipment in the facilities

Specialized computerized diagnostic equipment	Lviv study site	Sumy study site
CT (computed tomography)	✓	
MRI (magnetic resonance imaging)		
Digital radiograph	✓	✓
Digital fluorograph	✓	
Mobile (portable) digital radiograph	✓	✓

The Lviv Regional Phthisiopulmonology Clinical Treatment and Diagnostic Center planned to purchase specialized computer diagnostic equipment such as MRI in 2023, but this purchase was canceled due to lack of funds. It is planned to purchase this equipment in 2024. In turn, the «Regional Clinical Phthisiopulmonary Medical Center of Sumy Regional Council» didn’t plan to purchase specialized computer diagnostic equipment in 2023.

Additional information was also found on the Lviv study site: annually, the site conducts about 1 200 preventive X-ray examinations for TB and 17 300 chest X-rays (including CT).

In general, there is information on about 1 062 cases of all TB forms together with departmental services in the Lviv Region for 2022. Given the workload of the doctors at the Pulmonary Health Center, the optimal number of examinations per day is 55, per month – 1 700, and per year – 20 000. The facility is actively working with TB risk groups, but, according to the facility, there are still insufficiently screened rural residents of remote areas of the region. However, over the past year, the institution has intensified its visits to the districts of Lviv Oblast with a portable X-ray machine mounted on a car chassis.

2. Qualitative component of the study

X-ray algorithm

As part of the qualitative component of the study, information was collected from medical professionals and national experts in the field of the research topic at two sites in Lviv and Sumy regions using in-depth interviews with audio recording.

The Lviv study site involved the following 6 medical workers:

- Deputy Director General for Medical Affairs;
- Deputy Director General for Outpatient Care and Disability Evaluation;
- 2 Heads of Phthisiology Departments (phthisiologists);
- 2 radiologists.

By the length of service, the shortest was 5 years and the longest was 32 years. The average length of service was 19 years.

As for the Sumy study site, 6 medical workers were also participated in it:

- Acting medical director;
- Deputy director;
- 2 radiologists;
- 2 phthisiologists.

By the length of service, the shortest was 8 years and the longest was 24 years. The average length of service was 16 years.

As for the in-depth interviews with national experts in the field of TV, 4 experts were involved:

- Chief legal advisor of the legal support department;
- Senior TB/HIV program manager;
- PATH program manager;
- TB management and control department, phthisiologist.

By the length of service, the shortest was 6 years and the longest was 15 years. The average length of service was 11 years.

Prior to data collection, study site managers were contacted to inform them of the study purpose, study procedures, and expected roles of staff in relation to the study. It was also ensured that potential participants meet the inclusion criteria. At the stage of recruiting participants, individuals with at least 1 year of experience in their current position were selected. The recruitment was aimed at respondents who could provide the most detailed, informative and relevant information necessary to achieve the research objectives.

According to the respondents, currently, according to Order of the MOH No. 102 "Standards of Medical Care «Tuberculosis» dated 19.01.2023, a preventive chest X-ray is performed once a year. Under this Order, when a patient visits a family doctor, the latter must issue a referral for an X-ray, and the patient returns to the family doctor with the results of the examination. If there are any changes in the lungs, then the family doctor refers the patient to specialist doctors: a phthisiologist, pulmonologist, oncologist, etc. It takes an average of three days from the first contact with a family doctor to the probable diagnosis of TB. For high-risk patients, preventive examinations are conducted twice a year. The patient's route differs in the appointment of an additional sputum test and other tests prescribed by specialist doctors. This happens if a possible TB case is confirmed at the stage of screening questionnaire and preventive X-ray examination.

According to the participants, including the radiologists themselves, it takes a radiologist from 5 to 20 minutes to interpret an X-ray.

Obstacles for a radiologist in interpreting an X-ray image can be:

- poor quality of the X-ray;
- lack of light to interpret the X-ray;
- low quality of technical support (digital or analog device);
- poor quality of developer;
- incorrect positioning of the patient on the X-ray diagnostic equipment;
- careless handling of the image by the patient.

Radiologists have a working day of 6, not 8 hours. As for chest X-rays, a radiologist can treat 24 patients per day during a 6-hour working day (officially, according to the Order, 15 minutes per patient, as well as technological breaks). However, according to experts, the actual workload of a radiologist can be almost twice as high, up to 40 patients per day.

An X-ray examination alone is insufficient to determine a patient's TB diagnosis. It is important to provide a comprehensive diagnostic evaluation consisting of several diagnostic tests and clinical assessments that together are highly accurate, such as taking history, determining the severity of the condition, ordering laboratory tests to detect TB mycobacteria in the patient's biological material, evaluating their results, etc. The above functions are the responsibility of a radiologist. In turn, the radiologist is responsible for the stage of diagnosis and drawing a conclusion based on the results of the radiological examination, i.e. making a differential diagnosis.

According to healthcare professionals, if the X-ray room has equipment of proper quality and qualified specialists, healthcare professionals do not have any problems with chest X-rays. The only thing that doctors noted was the patient's severe condition, when it is difficult and/or impossible to place the patient on the X-ray diagnostic equipment properly.

General perception and trust in computerized/automated diagnostic systems

The overwhelming majority of respondents have heard of computerized/automated diagnostic systems and know that CAD systems are not currently used in Ukraine.

The participants expressed the following opinions on the potential advantages and disadvantages of implementing CAD systems in healthcare (Table 2).

Table 2. Advantages and disadvantages of implementing CAD systems

Type:	Advantages of using CAD:	Disadvantages of using CAD:
Inpatient healthcare facilities with X-ray diagnosis departments	• Patients can be transferred to primary care (those without abnormalities); • acceleration of radiologist's work; • efficiency.	• There may be a lack of options for interpreting the results.
Primary healthcare centers	• The speed of research and obtaining results.	• The need for a referral to a radiologist in case of detected pathologies; • overdiagnosis.
Outpatient clinics Primary healthcare centers	• The speed of research and obtaining results.	• The need for a referral to a radiologist in case of detected pathologies.
Dispensaries	According to respondents, the use of CAD is inappropriate. Some respondents noted that there are currently no dispensaries.	
Mobile radiology teams	• Increase in the number of patients examined per day.	• Images may not have high resolution.

«The advantage, I think, will be throughput... it's the speed of the study itself, and the number of people that can be examined»

«Перевага я думаю що це буде пропускні способності… це швидкість самого дослідження, і кількість людей, яких можна обстежити»

«Well, this allows us to examine much more patients. That is, if we used to examine, in principle, how many? 25-30 mobile teams, now we can examine up to 70-80 people in one visit, in one day»

«Ну, це дає змогу обстежити значно більше пацієнтів. Тобто якщо ми раніше обстежували в принципі, скільки там? 25-30 мобільних бригад, то зараз ми можемо обстежити до 70-80 чоловік за один виїзд, за один день»

«The drawback is, as it happens with young professionals, and I had it in my time, overdiagnosis. For example, the system may perceive the same residual changes as fresh. Or vice versa, it may miss them»

«Недолік це, як буває у молодих спеціалістів, і в мене було в свій час - гіпердіагностика. Наприклад, ті ж самі залишкові зміни система може сприйняти як свіжі. Або ж навпаки, може пропустити»

According to healthcare professionals, with the successful implementation of CAD systems for interpreting X-ray images, 80-85% of patients can be transferred to primary care, which currently has no X-ray machines.

Therefore, in general, we can identify the main potential positive changes from the introduction of CAD:

- increase in the number of processed X-rays;
- increase in the number of detected cases of TB at an early stage;

- X-rays will be reviewed using CAD, but if pathology is detected, they will be additionally reviewed by a doctor;
- acceleration of the time for issuing the examination results.

Since most clinical symptoms of TB have low specificity, leading to FP clinically diagnosed cases of TB, one of the benefits of implementing CAD is a reduction in patients lost to follow-up and an increase in the proportion of bacteriologically confirmed cases of TB. Therefore, CAD systems allow for faster detection of pathological changes, which can help to start treatment in a timely manner. In addition, it should be noted that CAD systems have the advantage of reducing the impact of the human factor. After all, the CAD system does not get tired and the probability of errors due to reasons that may affect the specialists analyzing the image is much lower.

As for potential negative consequences from the introduction of CAD, healthcare professionals and national experts see only the need to properly «train» artificial intelligence, i.e., to provide it with the correct input data, norms, and pathologies.

Some of the healthcare workers suggested that there may be patient concerns about CAD systems. The reason for this may be the established habit of communication with a doctor and at least a partial understanding of the diagnosis process by a doctor. And CAD will be a new introduction in the country.

- «In your opinion, what fears might patients have about the use of automated diagnostic systems?» – «Distrust»

- «А от на вашу думку, які побоювання можуть бути у пацієнтів, щодо застосування автоматизованих діагностичних систем?» - «Недовіра»

According to the respondents, the introduction of CAD would still be more appropriate in rural areas, given that urban patients have more opportunities to access healthcare facilities.

It was suggested that CAD should also be introduced not only in remote areas but also for screening certain personnel at enterprises and facilities whose employees may need routine screening. During periods of increased workload for doctors due to many X-ray examinations, the use of CAD systems to detect suspected TB will make sense, which in turn will help the doctor analyze the patient's X-ray more accurately and, at the same time, faster. For example, some doctors have to visit enterprises, facilities, or rural areas where more than 70 patients may need an X-ray in a day. In this case, one doctor may face difficulties during such visits, which will affect the efficiency of work at this pace. Here, human factors will play role in the form of fatigue, loss of concentration, etc. At the same time, computerized/automated diagnostic systems will not «get tired» when processing tens of thousands of X-rays, which is a significant advantage in their use.

WHO recommendations

According to the WHO recommendation, which reads as follows: «Among persons aged 15 years and older in populations in which screening for TB is recommended, computer-aided detection programs may be used instead of human readers to interpret digital chest X-rays for screening and triage for TB disease», respondents also express their approval of the recommendation. Thus, there is no categorical position on the scenario of introducing CAD with the subsequent replacement of human readers or the introduction of CAD to assist physicians in identifying suspected TB.

Speaking about the potential advantages of using CAD instead of a human reader, we can highlight the following:

- the ability to interpret numerous examinations;
- fast processing of X-ray images and obtaining of conclusion;
- standardization and objectivity of the assessment;
- absence of human factors in the interpretation of examinations;
- inability to miss gross pathology.

Also, when implementing a CAD instead of a human reader, potential barriers and disadvantages may arise, such as:

- lack of proper technical support for equipment;
- lack of digital X-ray systems (not applicable for analog systems);
- possibility of overdiagnosis of TB;
- technical problems and malfunctions of CAD during diagnosis;
- lack of funding for CAD implementation;
- problems with power supply due to military operations;
- loss of work of some staff members.

Therefore, CAD can be used to screen and triage TB patients in the absence of a human reader. However, under national program guidelines and protocols, a healthcare professional may be involved in the preliminary diagnosis process and review the CAD report for confirmation – this may be a radiologist, clinician, or phthisiologist who is part of the screening program.

Advantages and disadvantages for patients

According to the majority of participants, patients would rather react positively to the introduction of computerized/automated diagnostic systems in the TB diagnostic process. At the same time, given that a certain proportion of patients with probable TB may belong to risk groups, there is a possibility that there will be a cohort of patients who will express distrust of CAD systems for the reasons listed below. In addition, technical difficulties or malfunctions may occur, potentially causing additional patient anxiety. Also, the attitude towards the system may depend on the patient's mental and emotional state, for example, whether the person is healthy or suspected of having TB. If TB is suspected and CAD confirms the diagnosis, patients may distrust the system's diagnostic results due to fear and concerns about their health.

Thus, patients' fears about computerized/automated diagnostic systems may have the following basis:

- distrust of computer systems due to behavioral patterns in healthcare;
- CAD can lead to overdiagnosis;
- there may be doubts about the correctness of the system's diagnosis.

The benefits that computerized/automated diagnostic systems can provide to patients:

- the ability to get results and start treatment quickly;
- protection against medical errors;
- standardized assessment;
- unbiased attitude to the patient;
- absence of stigma from healthcare professionals;
- no additional costs for the patient (film, etc.).

«...it will be better for the patient if the system detects this pathology at an early stage, which will lead to faster and better treatment at later stages, so I think they reacted positively to this»
«..для пацієнта буде краще, що система виявить цю патологію на ранніх стадіях, що призведе до більш швидкого і більш якісного лікування на подальших етапах, то я думаю вони позитивно до цього віднеслись»

«...patients are different. There are old believers who love zemstvo doctors. There are modern patients who trust modern technologies... it all depends on the patient»
«…пацієнті разні бувають. Бувають старовіри, які люблять земських врачів. Бувають сучасні пацієнти, які довіряють сучасним технологіям… це все індивідуально, залежить від пацієнта»

Patient groups that may benefit from the introduction of computerized/automated diagnostic systems:

- residents of rural and remote areas;
- employees of enterprises (for medical examination);
- population undergoing preventive examinations;
- patients at risk;
- cancer patients;
- HIV-infected people.

Capacity building: trainings

To ensure that doctors work properly with the CAD system, healthcare professionals involved in chest X-rays need to participate in training sessions, seminars, etc.

Training can be organized remotely in an online format for a predetermined duration of 1 hour to 2 hours, depending on the need. However, the training must necessarily include practical work with the CAD system, in particular, such steps as practice in installing the software and its use at the technical level (connecting the system to the installation, launching it, uploading X-rays, obtaining a CAD report, etc.) The number of trainees depends on the scale of CAD implementation, as well as whether training is required only for healthcare professionals directly involved in the chest X-ray and obtaining the examination report, or for all end users of the software.

Speaking about the expanded categories of healthcare professionals that can be included in the training, depending on the program's workflow at the screening sites, it is proposed to involve several of the following workers, namely

- Primary care provider – will be involved in registering the patient and recording symptoms/risk factors before referring the patient for a chest x-ray, if necessary;
- Radiologist – receives the X-ray and reviews the CAD results before referring the patient to the next steps;
- Radiologist/Physician – may be present on-site or remotely to verify the result and refer the patient for further confirmatory sputum examination;
- Program manager/administrative staff – has access to the software to be able to view data coming from different sites and use it to monitor and evaluate the program.

Radiologists and X-ray technicians may be sufficient to conduct training with healthcare professionals who are directly involved in the process of uploading and verifying CAD results.

In general, the interviewees emphasized that the implementation of CAD systems in Ukraine could potentially face difficulties due to insufficient fluency in computer systems and skills in working with office programs in general.

Also, given the negative experience of respondents in working with English or Russian-language software, training and installation of the software platform as a whole in Ukrainian will be a significant advantage to ensure the proper level of CAD work of healthcare professionals.

Legal aspects of CAD implementation

It should be noted that any new diagnostic method must be officially approved and authorized for use. As of 2023, computerized/automated diagnostic systems, in particular in the healthcare sector, have not been implemented in Ukraine. Therefore, when any new product, installation, program, etc. is introduced, it requires marketing authorization and putting into circulation and operation. Accordingly, treatment standards or clinical protocols governing its use must be approved. Without such documents, this research method cannot be used.

Regarding the need to inform patients who will be subject to a particular method of diagnosis and treatment, Ukrainian law provides for informing patients about the use of relevant methods and obtaining their informed consent or providing another type of medical care, unless it is an emergency when the patient is unable to give consent due to certain circumstances.

If the clinical protocols and regulations do not provide for such a test method, the introduction of a new test method requires the implementation of certain procedures, such as putting into circulation, putting into operation and amending the relevant protocols. Thus, the medical device must be put into circulation and operated under the technical regulations approved by the Cabinet of Ministers, namely under Regulation 554.

Accordingly, the question arises: Who will be responsible for CAD system errors? In medical practice, the doctor who approved the diagnosis is responsible for making a mistake when making a diagnosis. In turn, the healthcare facility is responsible for the quality of medical care in general. Therefore, the establishment of an appropriate diagnosis is the responsibility of the facility as a whole and the relevant doctor. It follows that the healthcare facility where the CAD system is installed will be responsible for errors made by the CAD system.

In turn, Ukrainian legislation also provides that all medical devices and equipment are subject to inspection and certification. Accordingly, if CAD systems are subject to certification, they must be included in the relevant list of medical equipment. In other words, the legislative process directly relates to the certification and validation of automated diagnostic systems for the detection of TB.

In the context of martial law, the issue of transferring medical data and patient records abroad has become very relevant. Therefore, as part of this issue, the relevant rules were introduced that provide for the transfer of such data, the observance of which is necessary.

If we elaborate on which structural units and stakeholders should be involved in the legislative process for the introduction of computerized/automated diagnostic systems in the medical sector of Ukraine in particular, this is the first thing to do:

- MOH;
- WHO of Ukraine;
- State Institution PHC;
- SE "Electronic Health";
- NGO working in the field of TB control;
- Expert TB and radiology groups;
- Patient organizations.

For the introduction and successful further use of CAD systems in Ukraine, it is important to conduct pilot project, which will help to anticipate potential disadvantages or advantages of using CAD or identify possible increases in patient coverage, etc.

Resources for implementation at the country level

From the point of view of TB detection and patient management, one of the main potential benefits of introducing CAD systems in healthcare will be a reduction in the number of patients lost to follow-up at the stage of triage and screening, which in turn will increase the number of patients who will be brought to further examination for sputum collection and diagnosis or isolation. Therefore, the introduction of CAD can speed up the diagnosis and treatment of patients with TB. Accordingly, the introduction of CAD at the national level in the future is quite reasonable. From this, we can conclude that it is important to ensure consolidated work and unification of all stakeholders and partners that should be involved in the planning and implementation of nationwide integration of CAD systems in Ukraine. Therefore, one of the important resources for ensuring the implementation of CAD is cooperation with healthcare facilities, technology providers, and other relevant organizations.

Another issue raised by the experts was ensuring transparency and cost-effectiveness of CAD vendors' systems. To introduce any new procedure or diagnostic algorithm, it is necessary to study the effectiveness of this system in practice, which will determine the need to implement such a CAD system. Accordingly, all issues of cooperation should be transparent and clearly regulated to prevent any corruption and not to create any advantages for specific market players. Accordingly, another resource that the country should provide when implementing CAD systems is to ensure transparency and clarity in the selection of a CAD system supplier. The cost-effectiveness of CAD should be compared to traditional methods of screening and triaging TB patients by measuring the number of reports issued simultaneously, patient coverage, and treatment prescription. Healthcare professionals and national experts see a clear cost-effectiveness of CAD systems, especially in settings with lacking human resources.

Experts see monitoring and quality assurance, as well as the accuracy of CAD results interpretation, as the CAD system notifies of suspicion and recognizes pathological changes. However, the system cannot diagnosis, so the doctor must verify the relevant CAD conclusion by performing external quality control. Given that the CAD system only sorts images, it will not completely replace radiologists but will slightly change their functional responsibilities. CAD will free doctors from reading and sorting X-rays that do not reveal pathology and leave more time for those patients who need further examination and treatment. Thus, human resources are an integral part of properly diagnosing a patient.

According to the participants, it is necessary to ensure the appropriate level of trust of healthcare professionals in computerized/automated diagnostic systems by conducting operational research to compare and contrast the results of reading an X-ray by a CAD and a radiologist. Presenting and massively disseminating the results to a wide range of stakeholders, including healthcare professionals, will primarily increase their confidence in CAD. Therefore, planning, organizing and conducting training sessions for healthcare professionals and all end users of CAD systems is one of the important factors for the successful implementation of CAD in Ukraine.

Procurement and supply

When including CAD-related materials and supplies in the procurement plan for the National TB Program, it should be taken into account that CAD software is expensive. In the future, in the absence of international assistance, the National TB Program should include CAD software in national programs that public funds will fund.

Specifics of the procurement procedure and rules to be followed for the purchase of CAD and related components:

- For NGOs, a tender should be performed;
- Procurement for state or municipal property is carried out exclusively under the Law of Ukraine «On Public Procurement».

Currently, it is only possible to ensure a stable supply of resources related to CAD systems for a maximum of 3 years. This factor depends on the state, on the availability of funding from the state or donors. It is very difficult to determine whether this is possible in the context of the war in Ukraine.

3. CAD calibration component of the study

Computer-aided detection (CAD) system results data analysis

The goal of the CAD calibration component was to facilitate future CAD implementation by determining the optimal diagnostic performance and thresholds of the selected CAD product in the local context and use case in Lviv, Ukraine.

To this end, cooperation was organized with the study site in Lviv, the Lviv Regional Phthisiopulmonology Clinical Treatment and Diagnostic Center «Pulmonary Health Center». During the first half of 2022, 3 156 X-ray examinations were performed at the site. It was decided to collect this number of X-rays for further analysis.

At this stage, the question arose of how to depersonalize the X-rays, which is a mandatory requirement to ensure the confidentiality of the data obtained, as well as compliance with the ethical and moral principles of the study. To this end, a visit to the Pulmonary Health Center was organized and carried out to establish communication with healthcare professionals and explain the purpose of the study, in particular, the CAD calibration component. Based on the results of the visit to the study site, a patient database template was developed for further processing, and technical issues related to further processing of the available information and data depersonalization were resolved. DicomBrowser software was used for the depersonalization of X-rays, which allowed to leave the X-ray in Dicom format of the proper size and quality, which is a requirement for calibration of CAD systems. As a result, all personal data indicated on the X-ray itself was removed, namely:

- Surname, name, patronymic of the patient;
- Date of birth;
- Patient ID;
- Patient's gender;
- Date of the X-ray;
- Time of the X-ray.

When deleting personal information, the variable «Patient's name, surname and patronymic» was replaced with the serial number of the X-ray from 1 to 3 156, respectively, identical to the patient number in the database. Thus, within a month, the database of the patient X-ray repository was formed along with the patient database. The patient database included the following basic information:

- Identified patient number (a unique numeric code to identify the patient and to link the X-ray and the patient (the code displays the number on the X-ray and in the table so that each X-ray can be linked to each patient);
- X-ray confirmation of TB (conclusion of the study site radiologist based on the results of the X-ray examination);
- GeneXpert result (sputum bacterial culture result, which was entered according to the laboratory values (negative-negative, positive-negative, positive-positive or traces);

- X-ray machine code (depending on the type of machine).

In addition, additional personal information about the patients was collected:

- Patient's age;
- Date of birth;
- Preliminary diagnosis of TB (history of TB);
- HIV status;
- diabetes mellitus status;
- Smoking status.

Below is a description of the study sample based on the patient database (Table 3, 4).

Table 3. Characteristics of the sample based on the patient database

CHARACTERISTIC	N = 3,156
Age, Median (Range)	47 (10; 123)
Previous TB History, No. (%)	
Case after treatment failure	1 (0.3%)
Treatment after the break	6 (1.9%)
Treatment after failure	2 (0.7%)
A new case	213 (70%)
Recession	82 (27%)
Missing	2,852
HIV Status, No. (%)	
No	213 (70%)
Yes	91 (30%)
Missing	2,852
Diabetes Mellitus History, No. (%)	
No	292 (96%)
Yes	12 (3.9%)
Missing	2,852
Smoking Status, No. (%)	
No	245 (81%)
Yes	59 (19%)
Missing	2,852

Table 4. Characteristics of the sample based on the TB results

CHARACTERISTIC	N = 3,156
Radiologist opinion, No. (%)	
Negative	2,016 (64%)
Positive	1,140 (36%)
GeneXpert Result, No. (%)	
Negative	67 (45%)
Positive	81 (55%)
Missing	3,008
X-Ray Machine Code, No. (%)	
1	3,046 (97%)
2	110 (3.5%)

Having received a complete patient database and depersonalized the X-ray database, the process of establishing communication with CAD manufacturers was started to establish an agreement on the implementation of the CAD calibration component. A secure StopTB server was chosen to store and transfer X-ray images. The PHC uploaded 3 156 X-rays in Dicom format to the server for further calibration study by representatives of CAD manufacturers.

First, contacts and agreements on further cooperation within the calibration component of the study were established with three CAD companies (Qure.ai, Delft Imaging and Infervision). As a result of the work, a remote reading session was successfully conducted with the Qure.ai and Delft Imaging teams for X-ray images uploaded to the server, and the corresponding CAD interpretation results were obtained. Below is a descriptive characterization of the sample task of the study based on the results of the interpretation of X-ray images using CAD (Table 5).

Table 5. Characteristics of the Qure.ai model results

CHARACTERISTIC	N = 3,156
Threshold based on GeneXpert Result (0.668) for Qure.ai, No. (%)	
Missing	20
Negative	2,748 (88%)
Positive	388 (12%)
Threshold based on Radiologist Conclusion Result (0.0795) for Qure.ai, No. (%)	
Missing	20
Negative	1,998 (64%)
Positive	1,138 (36%)

The total number of samples is 3 156, of which 20 of the samples were not processed by CAD because they were invalid Dicoms (Figure 7).

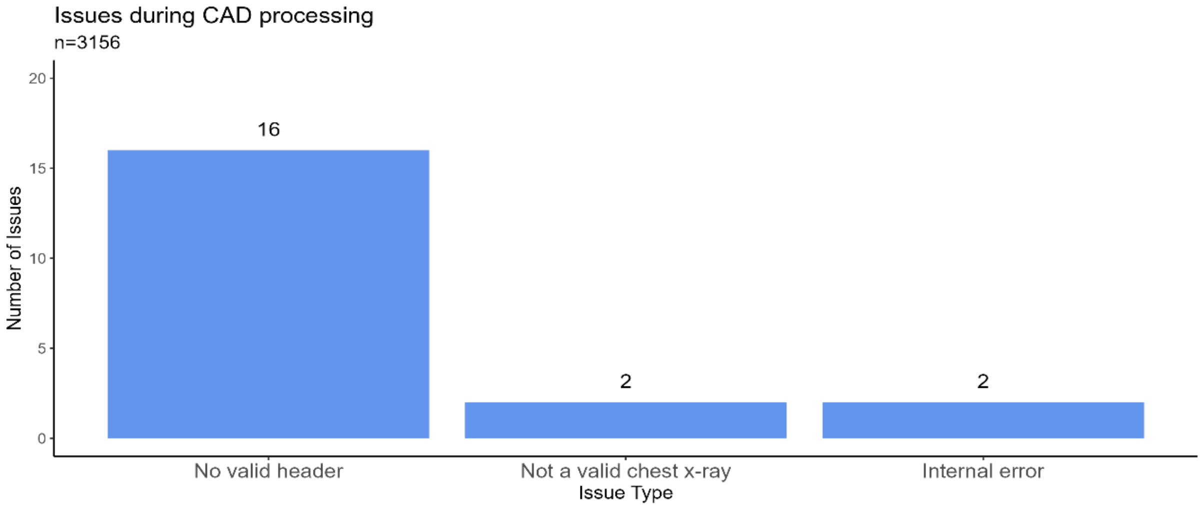


Figure 7. Types of issues during CAD system processing procedure

Below are detailed reasons and explanations for refusing to process X-rays by CAD Qure.ai (Table 6).

Table 6. Detailed reasons and explanations for non-processing

Input	Error in logs	Reason for non-processing
2	Not a valid chest X-ray	X-ray's QC model has classified it as an invalid X-ray
2	Internal error	Bad image quality - as per the logs, these are fully white Dicoms
16	No valid header	Dicom file doesn't contain the correct prefix used to identify a valid Dicom

Therefore, the total number of samples processed by CAD tool is 3 136.

Table 7. Characteristics of the Delft Imaging model results

CHARACTERISTIC	N = 3,156
Threshold based on GeneXpert Result for Delft Imaging, No. (%)	
Missing	8
Negative	2,805 (89%)
Positive	343 (11%)
Threshold based on Radiologist Conclusion Result for Delft Imaging, No. (%)	
Missing	8
Negative	2,088 (66%)
Positive	1,060 (34%)

Delft Imaging processed 3 148 images from 3 156 (99.7%), when Qure.ai processed 3 136 from 3 156 (99.4%). Only 3 images that were not processed by Delft Imaging and Qure.ai are similar, other ones are different.

Table 8. Characteristics of not processed samples by Delft Imaging model

Patient ID	Radiologist Conclusion	Age	Qure.ai Score	Delft Imaging Score
733	Positive	77		-0,01
1637	Positive	53	0,788	-0,01
2123	Negative	21		-0,03
2420	Negative	56		-0,03
2792	Negative	59	0,15	-0,01
3037	Negative	32	0,016	-0,03
3038	Negative	41	0,034	-0,03
3199	Negative	55	0,047	-0,01

In Ukraine, patients with suspected TB are identified in primary healthcare facilities and any other healthcare facilities by the staff of these facilities.

TB detection has the following algorithm: When a patient with TB symptoms presents, a screening questionnaire is conducted to identify risk factors and symptoms that may indicate TB. In case of at least 1 or more «Yes» answers, healthcare professionals provide an X-ray examination. If the X-ray examination reveals a suspected TB case, the patient is referred for a qualitative sputum specimen collection and transportation of the collected specimens to the nearest TB microbiology laboratory for testing with GeneXpert. TB diagnosis is confirmed in a specialized TB facility.

Accordingly, this is one of the reasons for the study limitation, which does not allow to collect the required number of GeneXpert results. Therefore, only 148 GeneXpert results are present (4.7% from the whole sample), of which 3 were not processed by CAD (Issue – «No valid header»). The total number of samples processed by CAD with GeneXpert result available is 145 (4.6% from the processed samples) (Figure 8).

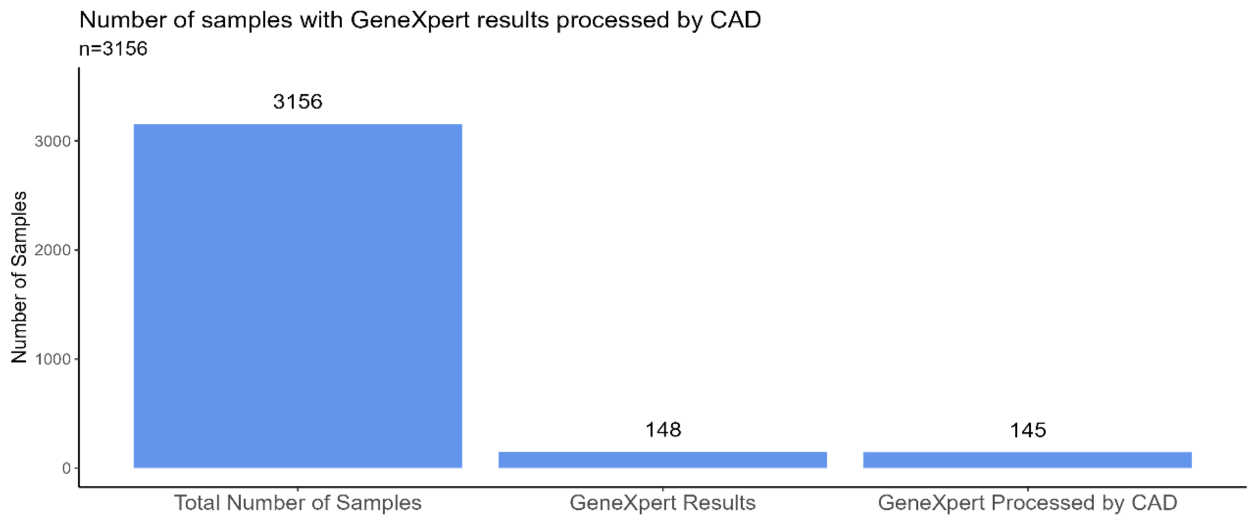


Figure 8. GeneXpert results availability cascade

During data collection, according to the TB Manager computer program, the following additional information about patients was obtained:

- previous TB history;
- smoking status;
- presence of diabetes mellitus;
- history of HIV.

This personal information about patients is available for 304 individuals (9.6%). At the time of the study, the necessary information could be obtained from the TB Registry, which contains relevant risk factors and information on comorbidities. As for all other cases (90.4%), the Pulmonary Health Center does not have information on the required number of patients. The reasons for this are that some patients had separate X-ray examinations and came to the healthcare professionals of the study site with ready-made film X-rays. At the same time, risk factors were not recorded in the facility's medical information systems for patients who were not confirmed to have TB. Patients could also undergo routine diagnostic X-ray examinations and be registered in other medical information systems to which the study site healthcare professionals do not have access (especially to sensitive data such as HIV infection or diabetes), and such data are not recorded in the logs. In addition, a new server was installed on the study site, which resulted in the loss of some data.

Therefore, out of the total number of samples processed by CAD, the number of images with personal data is 300 (9.6% of the processed samples) (Figure 9).

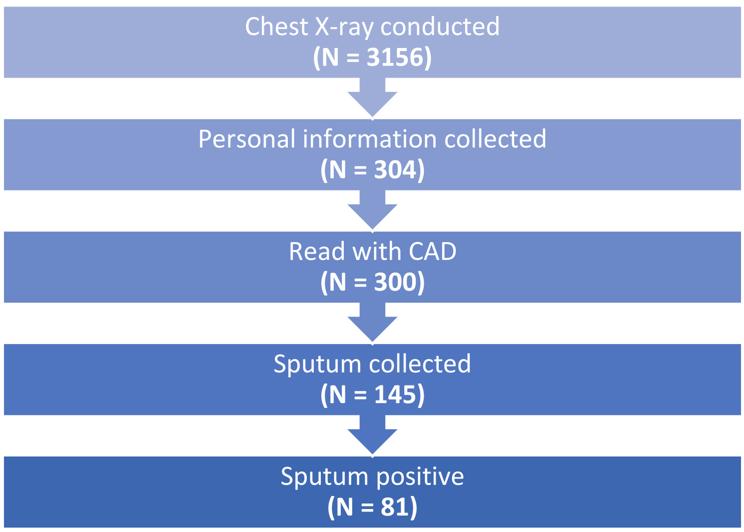


Figure 9. Samples with personal information cascade

Below is a graphical representation of the distribution of CAD scores by example Qure.ai for all 3 136 successfully processed samples to visualize how the CAD scores are distributed among the samples (Figure 10).

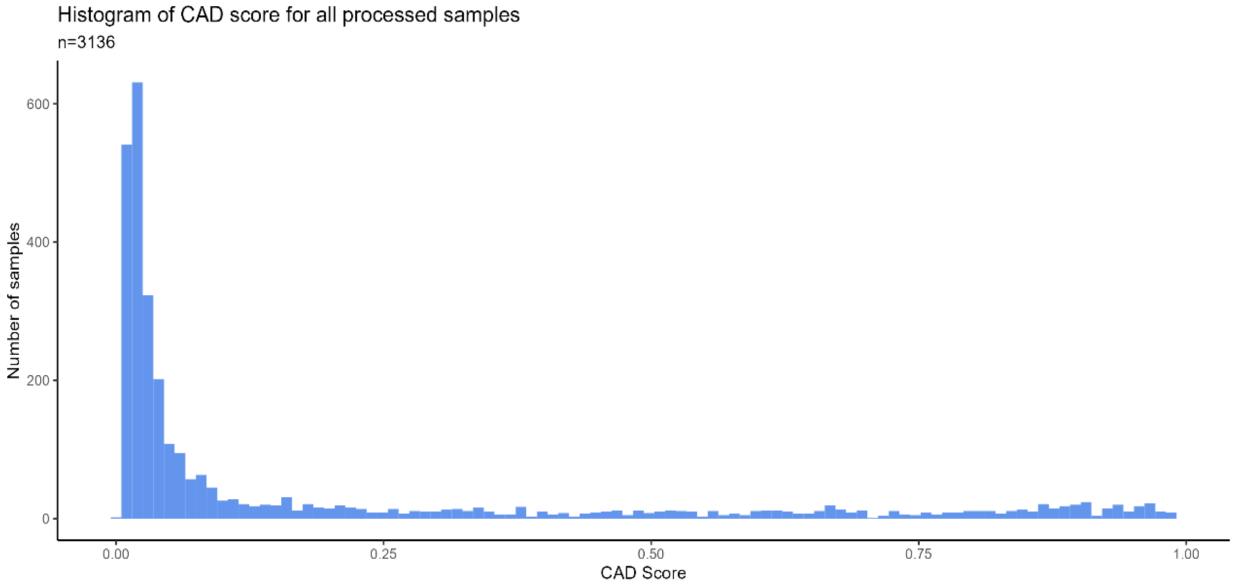


Figure 10. Histogram of CAD score for all processed samples

Figure 4 shows the distribution of the CAD score for all processed samples. According to the figure, the majority of samples got a score 0 – 0.1, which makes the distribution right skewed. According to the distribution, it is possible to assume that every change of the CAD score threshold above ~ 0.1 for positive/negative result will not have a significant impact on the distribution of positive/negative cases according to the CAD, as the proportion of negative cases will not change significantly.

GeneXpert versus CAD system results comparison

GeneXpert versus CAD system results comparison: CAD system predefined threshold

According to «A toolkit to support the effective use of CAD for TB screening», prepared by WHO, it is recommended to plan the appropriate sample size according to the study design and sensitivity level.

Considering the indicator «sensitivity» = 90%, the number of confirmed TB cases (assuming TB prevalence of 100%) should be 138 cases and the number of confirmed non-TB cases required (assuming the same precision and similar specificity as above) should be 138 cases (based on +-5% precision). Accordingly, the overall enrolment size required (assuming the same % and precision for Sp) should be 276 cases.

Considering the realized sample based on the available results for GeneXpert (N = 145, positive = 78, negative = 67), the number doesn't meet the minimum sample according to the WHO protocol.

Table 9. GeneXpert vs. CAD system by Qure.ai contingency table for pre-defined CAD system threshold (0.5)

CAD (threshold: 0.5)	GeneXpert		
	Negative	Positive	Total
Negative	48	20	68
Positive	19	58	77
Total	67	78	145

The CAD system, whose calibration estimates were used in the study, suggests using a threshold of 0.5 as a threshold point for classifying TB screening results into two categories that define a positive and negative label for the likelihood of a patient having TB.

Having received the results of the CAD system evaluation according to the available GeneXpert results, the following calculations were made:

- Sensitivity (Sen) – the proportion of TB cases correctly identified as positive (TP), calculated as the number of correctly diagnosed TB cases divided by the total number of true TB cases – **74.4%** =TP/(TP+FN), if FN - false negative cases;
- Specificity (Spe) – the proportion of non-TB cases correctly classified as true negative (TN), calculated as the number of correctly diagnosed non-TB cases divided by the total number of true non-TB cases – **71.6%** =TN/(TN+FP), if FP – false positive cases;
- Positive Predictive Value (PPV) – represents the probability that a positive test result represents a TP, calculated as the number of correctly diagnosed TB cases divided by the total number of TB positive diagnoses – **75.3%** =TP/(TP+FP);
- Negative Predictive Value (NPV) – is used to describe the performance of a test and represents the probability that a negative test result represents a TN, calculated as the number of correctly diagnosed non-TB cases divided by the total number of TB negative diagnoses – **70.6%** =TN/(TN+FN).
- Accuracy – is a statistical measure used to assess the overall correctness or reliability of a classification or diagnostic test – **73.1%** = (TP+TN)/TP.
- Percent agreement: **73.1%**

Accordingly, when comparing two tests, where one of the tests is the gold standard, the value of Percent agreement = the value of Accuracy.

Cohen's kappa (κ): 0.46 (95% CI: 0,31 to 0.60) – fair-to-moderate level of agreement. F1 Score – 0.71 (the F-score or F-measure is a measure of a test's accuracy. It is calculated from the precision and recall of the test, where the precision is the number of TP results divided by the number of all positive results, including those not identified correctly, and the recall is the number of TP results divided by the number of all samples that should have been identified as positive[1]).

Looking at the calculations, we can conclude that the frequency of true-positive patients, i.e., the indicator of CAD sensitivity to GeneXpert results, doesn't show a high enough value.

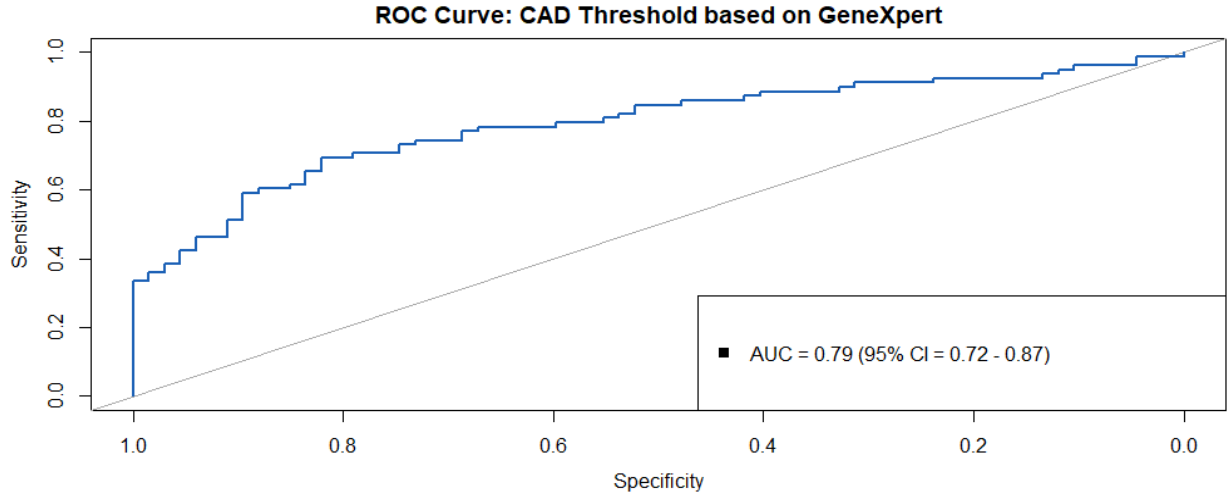


Figure 11. ROC-Curve for different CAD system by Qure.ai threshold based on GeneXpert results

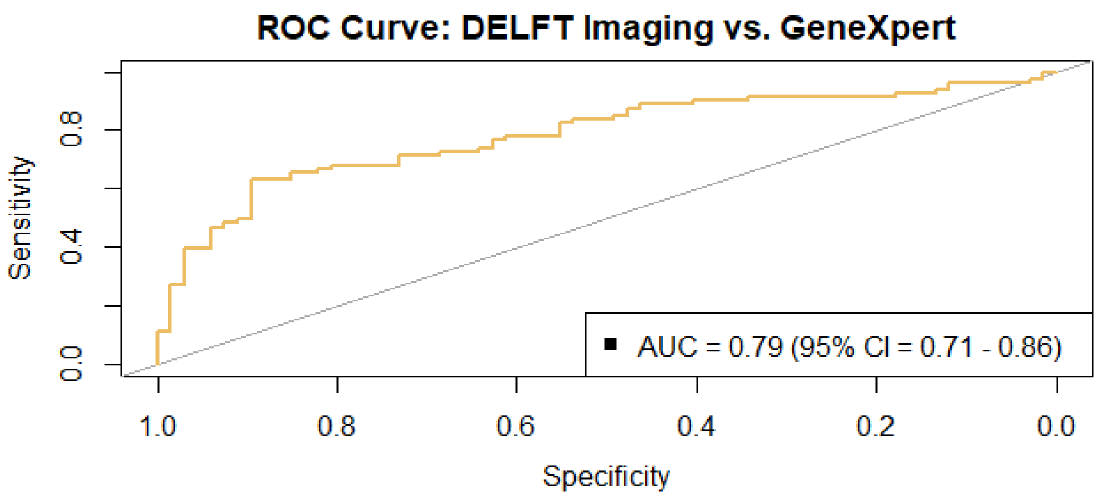


Figure 12. ROC-Curve for different CAD system by Delft Imaging threshold based on GeneXpert results

AUC provides an aggregate measure of performance across all possible classification thresholds. One way to interpret AUC is as the probability that the model ranks a random positive example more highly than a random negative one.

AUC ranges in value from 0 to 1. A model whose predictions are 100% wrong has an AUC of 0.0. One whose predictions are 100% correct has an AUC of 1.0.[2]

According to the data obtained, the AUC = 0.79 can be interpreted as a measure of the model's effectiveness in recognizing two classes based on the features or input data. The resulting model with AUC = 0.79 has a high chance of correctly classifying objects based on the given characteristic features, but did not achieve perfect resolution. Accordingly, the model shows good performance, but can be improved if higher recognition accuracy is important.

The choice of a threshold value, such as 0.5, can significantly impact the sensitivity, specificity, and overall performance of the test.

Having obtained the calculations, we can determine the Younden's Index. The maximum value of the index may be used as a criterion for selecting the optimum threshold point. The index is represented graphically as the height above the chance line, and it is also equivalent to the area under the curve (Figures 11,12).

Qure.ai: Youden's J statistic (Youden's index) = Sensitivity + Specificity - 1

Youden's Index – **0.668**
Specificity – **82%**
Sensitivity – **69%**

Youden's Index = 0.668 indicates a fairly high balance between Sensitivity and Specificity. This means that the resulting model distinguishes well between objects of two classes and detects TP and TN with high accuracy. Youden's Index indicates the high ability of the classifier to recognize and distinguish objects of different classes effectively.

Delft Imaging: Youden's J statistic (Youden's index) = Sensitivity + Specificity - 1

Youden's Index – **0.558**
Specificity – **89%**
Sensitivity – **63%**

Consider the «Area Under the ROC Curve» (AUC), which measures the entire two-dimensional area under the entire ROC curve from (0;0) to (1;1).

GeneXpert versus CAD system results comparison: CAD system threshold based on Youden's Index

Table 10. GeneXpert vs. CAD system by Qure.ai contingency table for Youden's Index CAD system threshold (0.668)

CAD (threshold: 0.668)	GeneXpert		
	Negative	Positive	Total
Negative	55	24	79
Positive	12	54	66
Total	67	78	145

Given the resulting Youden's Index = 0.668, we can make recalculations and see the changes (Table 10):

- Sensitivity – **69.2%,** this indicator decreased with the recalculation of the Youden's Index (compared to the previous one 74, 4%).
- Specificity – **82.1%,** this indicator increased with the recalculation of the Youden's Index (compared to the previous one 71, 6%).

Accordingly, we have the following changes in the following indicators:

- Positive Predictive Value (PPV) – 81.8%.
- Negative Predictive Value (NPV) – 69.6%.
- Accuracy – 75.2%.
- Cohen's kappa (κ): 0.51 (95% CI: 0.37 to 0.64) – fair-to-moderate level of agreement.
- F1 Score – 0.75.

Table 11. GeneXpert vs. CAD system by Delft Imaging contingency table for Youden's Index CAD system threshold (0.558)

CAD (threshold: 0.558)	GeneXpert		
	Negative	Positive	Total
Negative	60	30	90
Positive	7	51	58
Total	67	81	148

Given the resulting Youden's Index = 0.558, we can make recalculations and see the changes (Table 11):

- Sensitivity – **62.9%,** this indicator decreased with the recalculation of the Youden's Index.
- Specificity – **89.5%,** this indicator increased with the recalculation of the Youden's Index.

Accordingly, we have the following changes in the following indicators:

- Positive Predictive Value (PPV) – 87.9%.
- Negative Predictive Value (NPV) – 66.7%.
- Accuracy – 75%.
- Kohen's kappa (κ): 0.51 (95% CI: 0.38 to 0.64) – fair-to-moderate level of agreement.

Table 12. Different CAD score by Qure.ai thresholds for specificity > 0.8

CAD Threshold	Specificity	Sensitivity
0.662	0.806	0.692
0.668	0.821	0.692
0.671	0.821	0.667
0.674	0.821	0.654
0.679	0.836	0.654
0.686	0.836	0.641
0.694	0.836	0.628
0.698	0.836	0.615
0.700	0.851	0.615
0.710	0.851	0.603

According to the distributions of values in Table 12, Youden's index defines the best balance between Specificity and Sensitivity indeed, as a higher CAD score threshold leads to a significant drop in Sensitivity without a significant increase in Specificity.

Radiologist versus CAD system results comparison

Radiologist versus CAD system results comparison: CAD system predefined threshold

The calibration study has been designed based on the following assumptions:

- ✓ Potential users do not plan to use CAD to replace a human reader completely.
- ✓ CAD can be used as an initial screen, with any abnormal result by CAD (i.e., an abnormality score above the determined threshold) referred to a human reader for final interpretation before being referred for diagnostic evaluation.
- ✓ There is a need to determine the most appropriate thresholds based on local conditions and usage scenarios.

The feasibility of comparing CAD with a radiologist depends on many factors. The calibration component of the study was conducted to conduct a meaningful and informative comparison.

According to «A toolkit to support the effective use of CAD for TB screening» prepared by WHO, it is not entirely appropriate to compare CAD with a radiologist.

The appropriateness of the comparison should also consider how they can complement each other. CAD can assist radiologists with patient screening and triage, allowing radiologists to focus on more complex or critical cases.

Table 13. Radiologist vs. CAD system by Qure.ai contingency table for predefined CAD system threshold (0.5)

CAD by Qure.ai	**Radiologist**		
	Negative	Positive	Total
Negative	1,994	602	2,596
Positive	14	526	540
Total	2,008	1,128	3,136

For comparison of CAD scores and radiologist's results of X-ray examination, taking into account the predefined threshold of the CAD system by Qure.ai, we have the following indicators (Table 13):

Sensitivity – **46.6%**
Specificity – **99.3%**
Positive Predictive Value (PPV) – 97.4%
Negative Predictive Value (NPV) – 76.8%
Accuracy – 80.4%
Cohen's kappa (κ): 0.52 (95% CI: 0.49 to 0.55) – moderate level of agreement
F1 Score – 0.87

Table 14. Radiologist vs. CAD system by Delft Imaging contingency table for predefined CAD system threshold (0.5)

CAD by Delft Imaging	**Radiologist**		
	Negative	Positive	Total
Negative	1,771	317	2,088
Positive	239	821	1,060
Total	2,010	1,138	3,148

For comparison of CAD scores and radiologist's results of X-ray examination, taking into account the predefined threshold of the CAD system by Delft Imaging, we have the following indicators (Table 14):

Sensitivity – **72.1%**
Specificity – **88.1%**
Positive Predictive Value (PPV) – 77.5%
Negative Predictive Value (NPV) – 84.8%
Accuracy – 82.1%
Kohen's kappa (κ): 0.61 (95% CI: 0.58 to 0.64) – moderate level of agreement

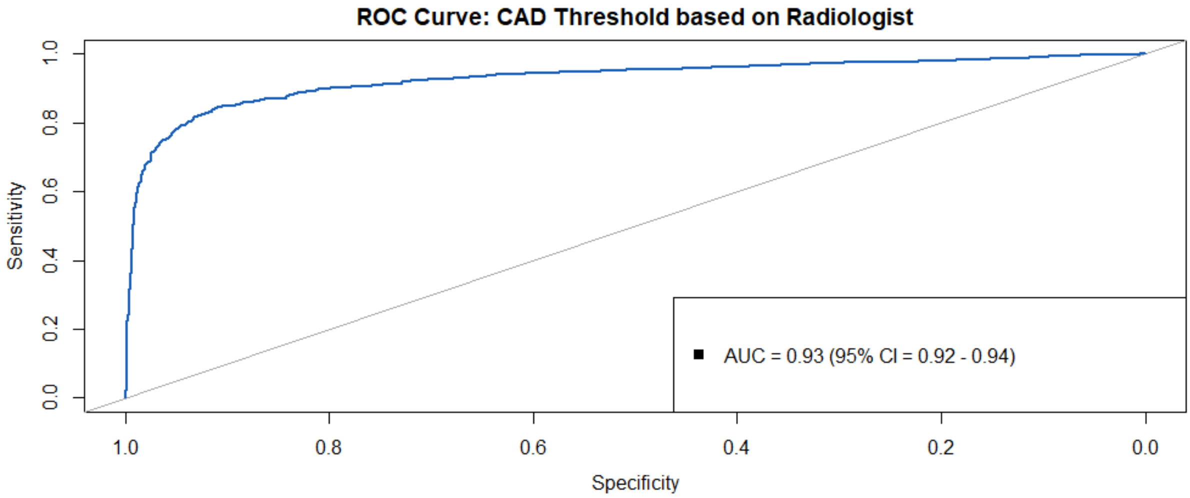


Figure 13. ROC-Curve for different CAD system by Qure.ai threshold based on radiologist results

According to the results obtained by the radiologist from the X-rays, we can build the following ROC curve (Figure 13).
Younden's Index – **0.0795**
Specificity – **90.8**
Sensitivity – **84.6**
Instead, we obtained an **AUC = 0.93**, demonstrating the model's high efficiency. Accordingly, the use of Younden's Index = 0.0795 results in higher specificity and sensitivity.

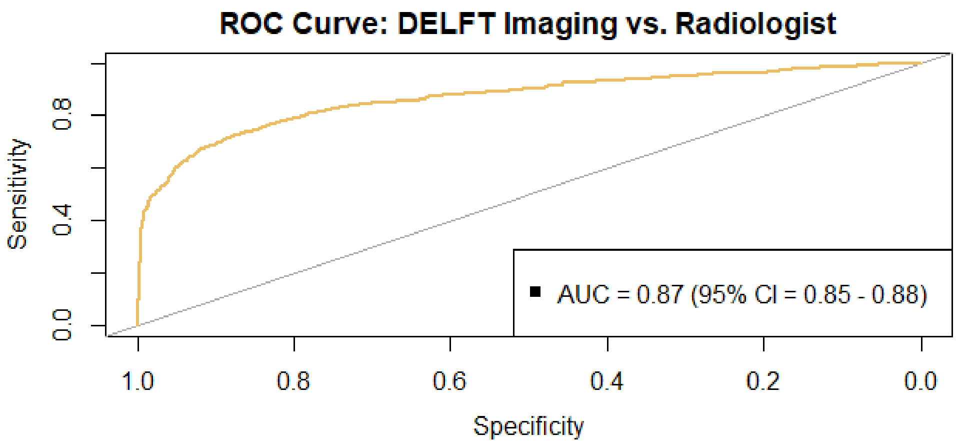


Figure 14. ROC-Curve for different CAD system by Delft Imaging threshold based on radiologist results

According to the results obtained by the radiologist from the X-rays, we can build the following ROC curve (Figure 14).

Younden's Index – **0.2145**
Specificity – **88.1**
Sensitivity – **72.1**
Instead, we obtained an **AUC = 0.87**

Radiologist versus CAD system results comparison: CAD system threshold based on Youden's Index

Table 15. Radiologist vs. CAD system contingency table for predefined CAD system threshold (0.0795)

CAD (threshold: 0.0795)	Radiologist		
	Negative	Positive	Total
Negative	1,824	174	1,998
Positive	184	954	1,138
Total	2,008	1,128	3,136

If we accept the threshold value and consider the CAD system threshold to be 0.0795, we can make recalculations and see the changes (Table 15):

Sensitivity – **84.6%**
Specificity – **90.8%**
Positive Predictive Value (PPV) – 83.8%
Negative Predictive Value (NPV) – 91.3%
Accuracy: 88.6%
Cohen's kappa (κ): 0.75 (95% CI: 0.73 to 0.78) – substantial level of agreement
F1 Score – 0.84

According to the data obtained, it can be seen that the sensitivity index has increased significantly compared to the previous value, with a CAD system threshold ratio of 0.5 (46.6%). In turn, the Specificity index slightly decreased compared to the previous value, with a CAD system threshold value of 0.5 (99.3%).

Radiologist versus GeneXpert

Table 16. GeneXpert vs. Radiologist contingency table

Radiologist	GeneXpert		
	Negative	Positive	Total
Negative	20	4	24
Positive	47	74	121
Total	67	78	145

If we compare the CAD score with the radiologist's results, we have the following indicators:

Sensitivity – **94.9%**
Specificity – **29.9%**
Positive Predictive Value (PPV) – 61.2%
Negative Predictive Value (NPV) – 83.3%
Accuracy – 64.8%
Cohen's kappa (κ): 0.26 (95% CI: 0.13 to 0.39) – slight-to-fair level of agreement

According to the obtained results and further comparison with GeneXpert results, it becomes clear, that even when Radiologist's conclusion has very high Sensitivity (94.9%) when compared with GeneXpert as a gold standard; however, the Specificity remains at the low level (only 29.9%), what potentially leads to significant amount of FP results and thus to high amount of unnecessary further diagnostic procedures, mostly culture and GeneXpert, what is either long or relatively expensive. Even when CAD sensitivity is lower compared to radiologists (69.2% vs. 94.9%, CAD score threshold based on Youden's index – 0.668), the Specificity of CAD is significantly higher (82.1% and 29.9%). This increase in Specificity, which can be achieved with the implementation of a CAD system, can help decrease the number of unnecessary diagnostic procedures, potentially saving both the time of healthcare professionals and the money of healthcare facilities that provide TB testing.

Also, the overall performance of CAD system might be considered better than Radiologist, when compared with GeneXpert as a gold standard, as the accuracy of CAD and Radiologist is 75.2% and 64.8% respectively.

However, it is worth noting, when compared Cohen's kappa for CAD vs. GeneXpert and Cohen's kappa for Radiologist vs. GeneXpert, even Cohen's kappa for CAD system vs. GeneXpert is higher (Cohen's kappa (κ): 0.51 (95% CI: 0.37 to 0.64), CAD score threshold 0.668) is higher than the Cohen's kappa for Radiologist vs. GeneXpert (Cohen's kappa (κ): 0.26 (95% CI: 0.13 to 0.39)), as 95% Confidence intervals for both Cohen's kappa are overlapping, the difference between Cohen's kappa values can't be treated as statistically significant ($p>0.05$).

Comparison of Qure.ai and Delft Imaging models

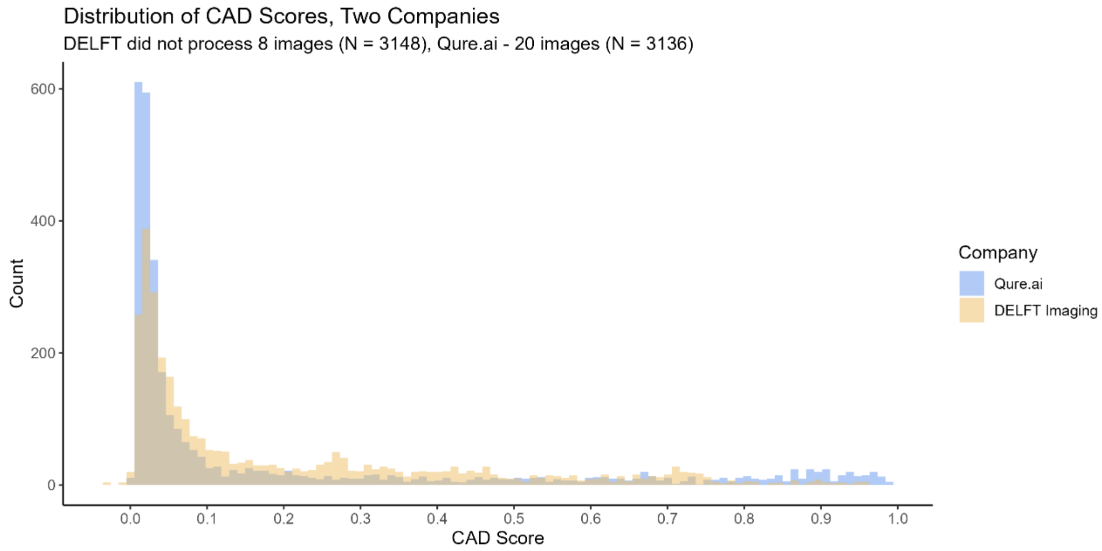


Figure 15. Distribution of CAD scores, Delft Imaging and Qure.ai

Delft Imaging has more 'smooth' distribution of CAD scores, with less sharp transition from 0.0 to 0.1 and lower number of values after 0.8.

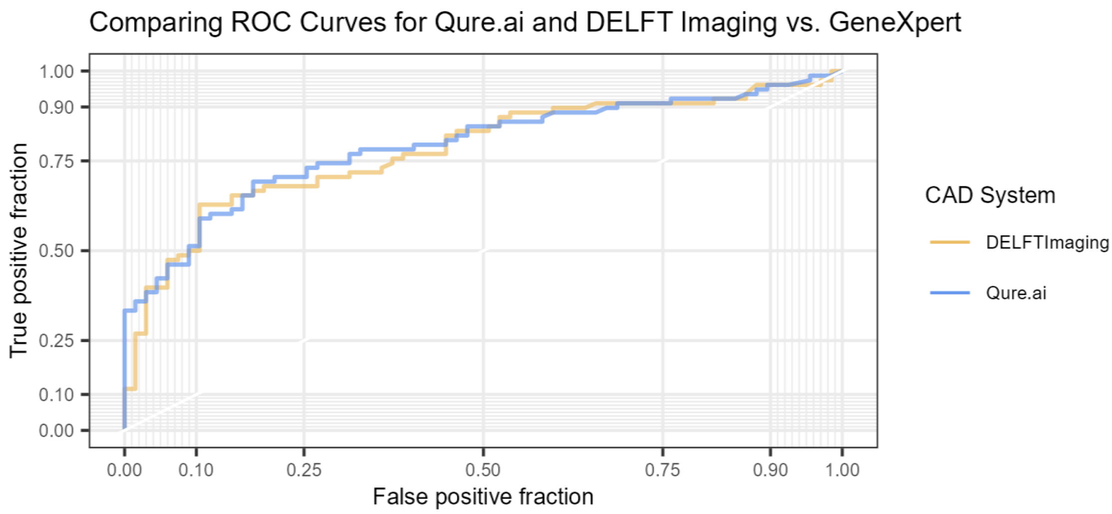


Figure 16. Comparing ROC Curves for Qure.ai and Delft Imaging vs. GeneXpert

Delft Imaging AUC: 0.79 (95% CI = 0.71 – 0.86).
Qure.ai AUC: 0.79 (95% CI = 0.72 – 0.87).
Percent agreement between Delft Imaging and Qure.ai when compared with GeneXpert: 94.9%
Cohen's data value for two models based on GeneXpert threshold: 0.75 (95% CI: 0.72 to 0.79) – substantial level of agreement.

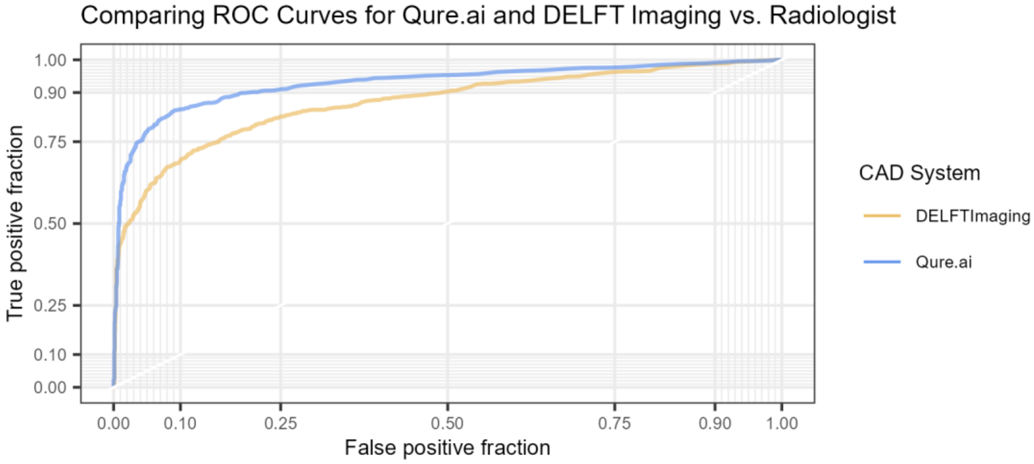


Figure 17. Comparing ROC Curves for Qure.ai and Delft Imaging vs. Radiologist

Delft Imaging AUC: 0.87 (95% CI = 0.85 – 0.88)
Qure.ai AUC: 0.93 (95% CI = 0.92 – 0.94)
Percent agreement between Delft Imaging and Qure.ai when compared with Radiologist: 84.8%
Cohen's data value for two models based on Radiologist threshold: 0.67 (95% CI: 0.64 to 0.69) – substantial level of agreement.

When compared with GeneXpert, the two systems have approximately the same indicators and perform at the same level. However, when compared to a radiologist, Qure.ai system performs slightly better (statistically significant higher AUC in Qure.ai).

In general, if we take into account the fact that the comparison of CAD systems with a radiologist is not entirely correct, the systems show almost the same efficiency, with the only difference that Delft Imaging processed more images.

Fleiss kappa to compare 3 methods: radiologist opinion, GeneXpert results and CAD system results.

After performing calculations and analyzing the proposed CAD thresholds, we need to quantify the reliability between the above approaches.

Fleiss kappa is used to assess the level of agreement between several methods when categorizing or evaluating items.

In practice, a positive value of Fleiss Kappa indicates agreement beyond chance, and the closer the value is to 1, the higher the level of agreement among raters. Conversely, a negative value indicates disagreement beyond what would be expected by chance. We can use Fleiss kappa to assess the reliability of ratings or classifications produced by multiple methods, and this can help determine the consistency of the ratings:

- Fleiss kappa (κ), CAD threshold – 0.5 (default): **0.326** (95% CI: 0.232 to 0.42) – fair-to-moderate agreement.

- Fleiss kappa (κ), CAD threshold – 0.668 (based on GeneXpert ROC): **0.305** (95% CI: 0.211 to 0.399) – fair agreement.
- Fleiss kappa (κ), CAD threshold – 0.0795 (based on Radiologist ROC): **0.361** (95% CI: 0.267 to 0.455) – fair-to-moderate agreement. This approach provides the highest Fleiss kappa among all three thresholds.

Even though Fleiss kappa (κ) for CAD system threshold based on Youden's index obtained from the comparison with GeneXpert is lower (0.305 compared to 0.361 for Youden's index based on comparison with Radiologist and 0.326 based on default threshold), this difference might be explained by the presence of relatively large difference of conclusions between GeneXpert results and Radiologist opinions (Cohen's kappa (κ) for GeneXpert and Radiologist – 0.26). Thus, the CAD system's threshold based on GeneXpert results makes the difference in results between CAD and Radiologist bigger, which in turn decreases Fleiss kappa. Indeed, when comparing CAD system vs. Radiologist, Cohen's kappa (κ) for the CAD system threshold based on Youden's index obtained from the comparison with Radiologist is 0.75 (95% CI: 0.73 to 0.78), and Cohen's kappa value for CAD system threshold based Youden's index obtained from the comparison CAD system vs. GeneXpert is 0.39 (95% CI: 0.36 to 0.42), and the difference is statistically significant ($p<0.05$).

4. CAD implementation cost assessment component of the study

Estimating the cost of implementing computerized diagnostic systems for TB diagnosis is an important step in the process of developing and implementing new medical technologies that help to improve treatment efficiency and reduce the spread of this disease. Here are some of the main reasons for this:

- Budget planning: Finding out the costs of implementing such systems helps budget planning for organizations, hospitals or government agencies. This allows to determine the availability of the necessary resources and financial commitments.
- Cost-effectiveness: Cost-effectiveness assessments help determine the effectiveness of implementing computerized diagnostic systems in the fight against TB. It helps to assess whether the costs are justified in terms of improved diagnostic accuracy and speed, reduced treatment time, and prevention of disease spread.
- Quality of health services: Computerized diagnostic systems can improve the quality of TB diagnosis and provide more accurate and reliable detection of the disease. A cost assessment helps to determine how this can affect the overall quality of healthcare services and how beneficial it is for patients.
- Decision support: Knowing the value helps hospitals, organizations, and government agencies make informed decisions about implementing such systems. It allows them to prioritize and develop implementation plans based on limited resources.
- The spread of TB: TB remains a serious global health problem, and accurate and rapid diagnosis is key to controlling it. Computerized diagnostic systems can help identify and isolate patients in a timely manner, thus preventing the spread of the infection.

Evolving from paper to digital data management using one platform is key to making the TB screening process efficient. Given that Lviv Region has been demonstrating strong development of new approaches to timely diagnosis and appropriate treatment of TB for many years, as well as providing all possible support for innovation, development and implementation of new policies and facility solutions, a decision was made at the state level to plan the implementation of the ADS system in Lviv Region, namely in the Lviv Region Phthisiopulmonology Clinical Treatment and Diagnostic Center «Pulmonary Health Center» of the Lviv Region Department of Health. This particular HCF was involved in the costing component of the CAD system implementation within the TB program.
At the time of the survey data collection, Ukraine had not yet introduced CAD systems for TB diagnosis. However, the PHC has already started the process of purchasing CAD systems from Delft Imaging. CAD is a CE-certified AI software that supports fast, easy, highly accurate automated TB detection with cost-efficiency. CAD is available for online and offline use, enabling data-driven TB screening in resource-constrained settings. It can analyze images of individuals aged four and above.

The CAD platform enables TB programs to rapidly screen risk groups and support prevalence surveys while simultaneously collecting digitized data and information into one platform. Thus, it facilitates and simplifies data governance and optimizes the successful adoption of digital TB screening systems. The SAM functionalities contribute to data quality and integrity while avoiding data fragmentation. It also enables data access from multiple locations: by providing connectivity with other data systems at the national level, SAM can serve as the first data entry point for health workers at the community level, thereby ensuring that National TB Programs can access aggregated data and dashboards. To obtain information for the purpose of estimating the costs of CAD implementation, data collection forms were developed that covered all the indicators necessary for the assessment. These forms contained the following information:

- The cost of the CAD software
- Training of healthcare professionals to use CAD
- Common type of Xpert in Ukraine
- Time required to perform an analysis on Xpert
- Materials for sputum collection
- Cost of materials
- Transportation of sputum to the laboratory
- Areas of responsibility of doctors at the stage of diagnosis, treatment and completion of TB treatment
- Tests for the diagnosis and treatment of TB
- DOT TB
- Level of healthcare facilities in TB treatment
- Characteristics of a laboratory equipped to work with Xpert
- Assessment of wages in the public sector
- Time estimation of healthcare professionals

Characterization of cost elements

Speaking about the cost of CAD by Delft Imaging software, information about a fixed period should be provided for a subscription; it is given below (Table 17).

Table 17. CAD4TB cost

PRODUCT	DESCRIPTION	NUMBER OF UNITS PER PACK	PRICE, USD
SOFTWARE	CAD Software (Perpetual License)	1	12,750.00
EQUIPMENT	CAD4 box	1	2,750.00
SERVICES	CAD Installation and Training	1	1,150.00
	CAD Support and Maintenance Extension 1 year	1	5,100.00

The PHC made a purchase and has already received two systems in October 2023.

The service package purchased with the two CAD systems also includes a CAD box, system installation, training for medical personnel who will be involved in CAD work, as well as technical and administrative support and a 1-year extension of the system's service life.

The approximate duration of the training of healthcare professionals in the use of CAD, as well as the number of participants and categories of healthcare professionals to be trained. Depending on the scale of the implementation, a joint remote training session can be held, in which end users and program managers can participate. The training can take place via video call for a pre-determined duration of 1 hour to 2 hours, depending on the need. The number of trainees again depends on the scale of the system implementation, as well as whether the training is to be conducted for superusers only or all end users of the software.

As for the categories of health workers, depending on the workflow of the program at the screening sites, it is advisable to involve the following workers:

- Radiologist – They would acquire the x-ray and view the CAD output before sending the client for the next steps. May also print the report for records.
- Phthisiologist – They may be either present at the site or remotely to verify the output and send the client for further confirmatory testing.
- Lab Assistants – They may be preparing a person for an X-ray and involving them in the X-ray, downloading X-ray results using CAD.

According to the data received from the responsible healthcare professionals of the Lviv study site, we have the following information.

In Ukraine, the MTB/RIF Ultra or Xpert MTB/RIF technologies are commonly used for TB diagnosis. For the most part, Xpert MTB/RIF Ultra is an improved version of the original Xpert MTB/RIF test. It can detect Mycobacterium TB (Mtb) as well as resistance mechanisms to rifampicin (RIF), one of the antibiotics used to treat TB. In some cases, other test versions may also be used, such as the Xpert MTB/RIF Ultra G4, G8, or G16, but this may depend on equipment availability and the need for additional tests.

Maintenance and calibration of the Xpert system is typically performed under the recommendations of the equipment manufacturer. The accuracy and reliability of the tests are essential, so these procedures are important to ensure that the system is functioning properly.

The regularity of maintenance and calibration may depend on the specific model of equipment used in the healthcare facility and the manufacturer's recommendations. This is usually done on a quarterly or annual basis. In the case of the Lviv study site, maintenance/calibration of the Xpert is performed as the problem arises. The maintenance and calibration procedure can take 1 to 2 hours, and the time spent may vary depending on the specific system and its condition.

The time required for healthcare professionals to prepare a complete set of specimens for analysis on Xpert may vary depending on the specific procedure and the number of specimens being processed at one time. Typically, this process includes the following steps:

1. Preparation of the work area and all necessary materials: This step involves preparing a clean workspace and ensuring access to all necessary reagents and instruments. The time required for this step can vary, but usually takes between 10 and 15 minutes.

2. Sample collection: This involves collecting clinical specimens, such as saliva, sputum, or other bodily fluids, and preparing them for analysis. The time for this step can vary considerably depending on the specimen collection procedure and the number of specimens, but it typically takes between 5 and 30 minutes per specimen.
3. Sample preparation for analysis: This step involves adding the necessary reagents and preparing the samples for further analysis on the Xpert. Time for this step can also vary, but typically takes 10 to 15 minutes per sample.

Therefore, the total time to prepare a full set of samples for Xpert analysis will vary from approximately 25 minutes to an hour or more, depending on the procedures and number of samples being processed at one time. Healthcare professionals at the Lviv study site take approximately 3 hours to prepare a full set of 50 samples for Xpert analysis.

The time it takes for healthcare professionals to communicate the results of each Xpert specimen to physicians can vary depending on a number of factors, including location, notification procedures, and the results processing system. However, the overall time for this process can be minimized with efficient processing and notification systems. Typically, physicians are notified of Xpert results as soon as they are available or as soon as possible to improve the diagnosis and treatment of patients.

The transportation of biological material from all districts of Lviv Region to the Level 3 laboratory of the Pulmonary Health Center is carried out within the framework of the project «Support to the TB Control Efforts in Ukraine». Accordingly, transportation of samples is free of additional costs.

To collect sputum for laboratory analysis, using the example of the Lviv study site, the following materials and devices are usually required:

- Sputum collection cup: This container collects sputum from the patient. This can usually be a special plastic cup or container.
- Gloves: Medical gloves are used by healthcare personnel to ensure hygiene and safety during sputum collection. They help prevent contamination and transmission of infections.
- Respirators: Special masks used by healthcare personnel to collect sputum from patients with suspected infectious diseases such as TB. The function of these respirators is to prevent the spread of infections during sputum collection and to ensure safety for both healthcare professionals and other patients.
- A secure means of transportation and collection point: If the specimen needs to be transported to the laboratory, special safe containers for transporting biological specimens can be used.

It is important to note that when collecting sputum, it is critical to follow hygiene and safety standards to prevent the spread of infections, as sputum may contain pathogens such as bacteria or tubercle bacillus. Therefore, following the rules and using the appropriate equipment are important aspects of the sputum collection procedure.

Sputum collection for laboratory analysis is usually performed by nursing staff who are trained and skilled in this procedure. Nurses and medical assistants trained and equipped with the necessary skills to collect sputum from patients may be the main professionals involved in sputum collection.

The biological material is transported to the laboratory for analysis from remote areas twice a week by two cars of the Pulmonary Health Center, which cost UAH 57 058 and UAH 77 749. On average, the vehicle travels 400 km/day to transport sputum. The average cost of fuel required for transportation is UAH 7 000 per month. If sputum is collected in a hospital, it is sent directly to the laboratory. On average, it takes about 5 hours a day to transport 100 sputum samples.

A patient diagnosed with TB usually receives diagnosis, treatment, and follow-up from a team of healthcare professionals. This team may include different specialists working together to ensure optimal treatment and follow-up. At the Lviv study site, TB diagnosis involves a TB physician who diagnoses TB and develops a treatment strategy. The doctor is also responsible for monitoring treatment and adjusting therapy as needed. Healthcare staff, including nurses, may also be responsible for providing medical care, administering medications, performing injections, and monitoring the patient's condition during treatment. In some cases, a social worker may help the patient get the support and resources needed to be successful in treatment, as well as help with various social issues that may arise during TB treatment. During the last visit after completion of TB treatment, the patient should also visit a phthisiologist to confirm cure.

When diagnosed with TB, patients are prescribed a number of laboratory tests and other tests to assess their condition and help with treatment:

- Patients with TB are tested for human immunodeficiency virus (HIV) because HIV and TB can coexist, and TB treatment may require consideration of this factor.
- A liver function test (liver function tests) is performed to determine if the patient has liver dysfunction, especially since some TB medications can affect the liver.
- A general blood test and a test for the presence of specific markers are performed to determine the patient's general condition and identify diseases.
- Microscopic analysis of sputum to detect Mycobacterium TB (TB bacilli) and cultural diagnosis to determine the type of bacteria and its sensitivity to antibiotics.
- Sputum samples for treatment resistance: If TB is suspected, the laboratory can also perform an antibiotic resistance test to help determine whether the TB bacterium is sensitive to commonly used drugs.

During the treatment of sensitive TB, a smear is performed to monitor the response at 0, 60, 90, 120 doses. For drug-resistant TB, a copy and culture are performed monthly, depending on the treatment regimen.

TB treatment is also provided under Directly Observed Therapy (DOT), as this system helps to improve treatment adherence and prevent the development of treatment resistance. The DOT principle implies that a patient takes medications under the supervision of a healthcare provider or other qualified person, instead of taking medications at home without supervision.

Considering the entire Lviv Region, the share of patients receiving DOT is about 99%. The Pulmonary Health Center also provides DOT services for TB treatment, both in inpatient and outpatient settings. Speaking about the approximate time that a patient's visit to the DOT takes, the duration of a visit to the DOT may vary depending on a number of factors, such as the patient's condition, the number of medications taken, etc. Therefore, the duration of a visit to the DOT can vary, but it usually takes about 15 minutes. Patients do not have to pay any monetary costs associated with obtaining a DOT.

The frequency of routine medical check-ups for patients during TB treatment may vary depending on the practice, standards of care, and other factors. Typically, patients being treated for TB see their doctor or other healthcare professional every month for the duration of their treatment. This makes it possible to monitor treatment progress and respond to any complications or side effects in a timely manner. Additional tests and examinations may also be performed as needed.

As for the laboratory that is equipped to work with Xpert and cooperates with the Pulmonary Health Center, according to the financial information provided, the overhead costs of the healthcare facility amounted to UAH 4 million in the first half of 2023. The laboratory has a total of 19 full-time employees, with 9 healthcare professionals working on Xpert. If we count the number of Xpert samples evaluated in this facility each month, we will have a figure of 400. Overall, this is an average workload and patient flow for a laboratory working with Xpert.

Understanding the amount of physician salaries is an important aspect of implementing CAD in medical practice. Therefore, it is worth analyzing the salary estimates for healthcare professionals in the public sector in general in Ukraine (Table 18).

Table 18. Salaries of medical personnel in the public sector

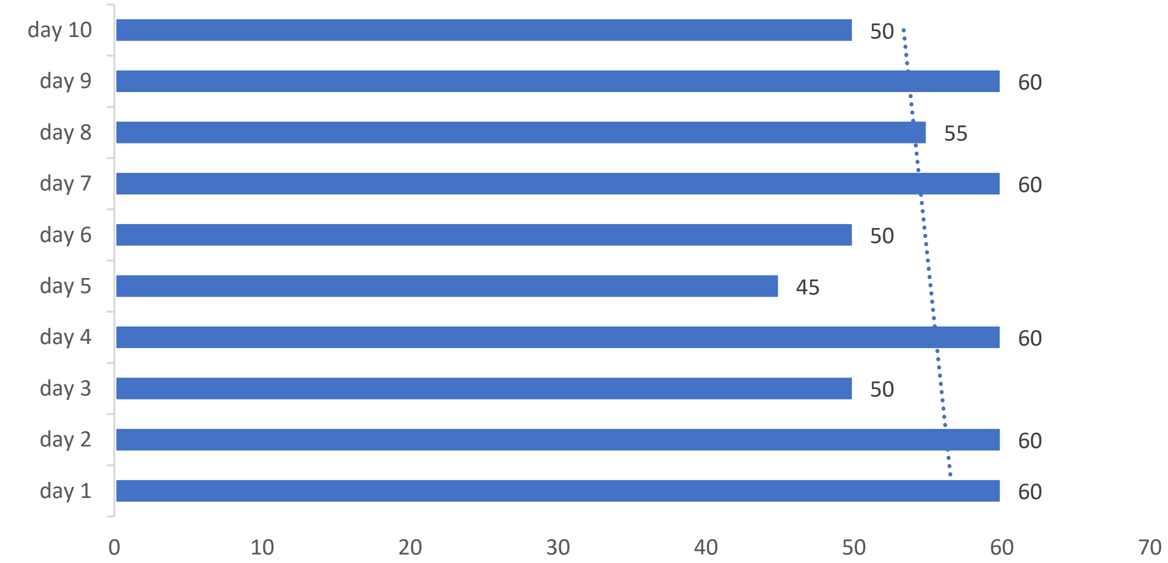
ANNUAL SALARY	UAH
Nurse	162 000,00
Phthisiologist	240 000,00
Radiologist	259 860,00
Healthcare administrative staff	186 936,00
Laboratory assistant	163 200,00

Time estimation of healthcare professionals

To be able to transparently and clearly assess the cost of implementing CAD systems, it is worth analyzing the time spent by healthcare professionals involved in the diagnosis, treatment and follow-up of TB patients.

The study collected data from a phthisiologist, nurse, radiologist, and X-ray technician at the Pulmonary Health Center on the performance of certain activities and the time allotted for the performance of the health service. From July 31 to August 11, 2023, for 10 working days, at the end of each day, the responsible healthcare professionals reviewed each of the specified tasks and noted whether they had completed it today. In addition, healthcare professionals estimated the time in minutes they spent on this task per patient. The total time spent performing the task was indicated, including any preparation and time spent with the patient. Also included was time spent interpreting results, discussing with colleagues, and any other administrative work associated with the task. The purpose of this data collection was to determine the average time required to perform various tasks related to the treatment of multidrug-resistant TB to support the appropriate allocation of funding and human resources.

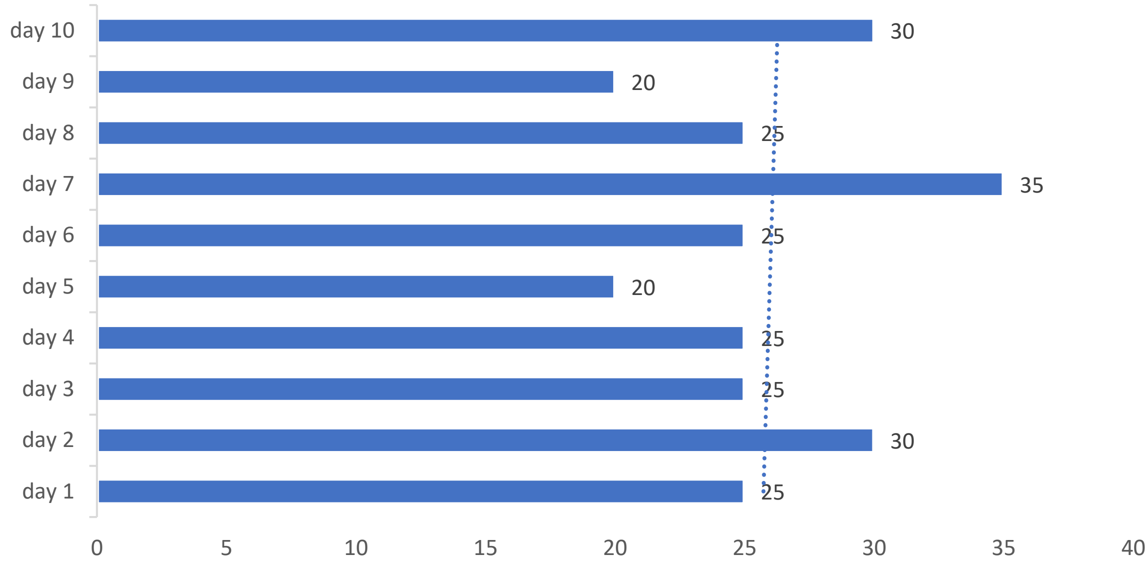
When it comes to estimating the time required to provide services by a phthisiologist, we have the following distributions.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
minutes	60	60	50	60	45	50	60	55	60	50

Figure 18. Initiating a patient on TB disease treatment (first visit) – includes conducting any tests, counselling patient, prescribing medication

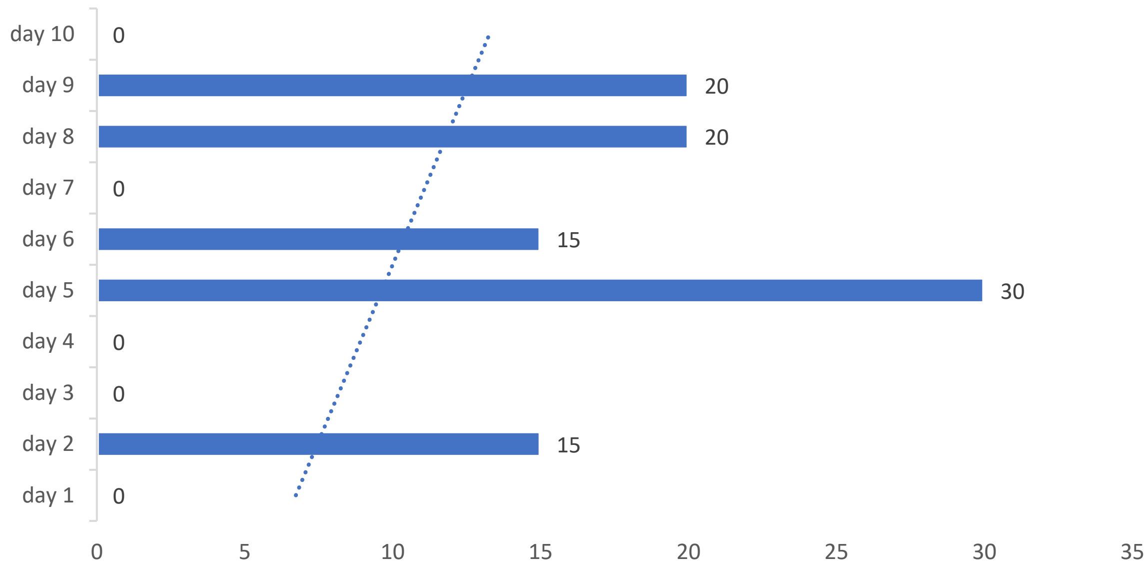
According to the distribution, we can analyze what services a phthisiologist provides when starting TB treatment. This includes conducting tests, counseling the patient, and prescribing treatment. We can see that during each day a phthisiologist performs the above activities. The average time is 55 minutes, taking into account the provision of services for one patient.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
■ minutes	25	30	25	25	20	25	35	25	20	30

Figure 19. Seeing a patient for a follow-up visit during treatment – includes conducting clinical examination, interpretation of test results, inquiring about adherence, recording any patient notes

Speaking of patient follow-up during treatment, according to the trend line, we can analyze that, on average, a phthisiologist takes about 26 minutes per patient to conduct a clinical examination, interpret test results, determine treatment adherence, and record any patient comments.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
■ minutes	0	15	0	0	30	15	0	20	20	0

Figure 20. Seeing a patient for their final visit once they complete treatment – includes conducting clinical examination, interpretation of test results, recording any patient notes, declaring a treatment outcome

The patient’s final visit after the treatment is completed includes a clinical examination, interpretation of test results, recording any patient notes, and announcement of the treatment result. During the data collection period, these activities were not performed by the phthisiologist every day. However, 15 to 30 minutes per patient were allocated, depending to the time spent on these services.

It is also worth analyzing the performance of certain services nurses provide and the amount of time allocated to these services.

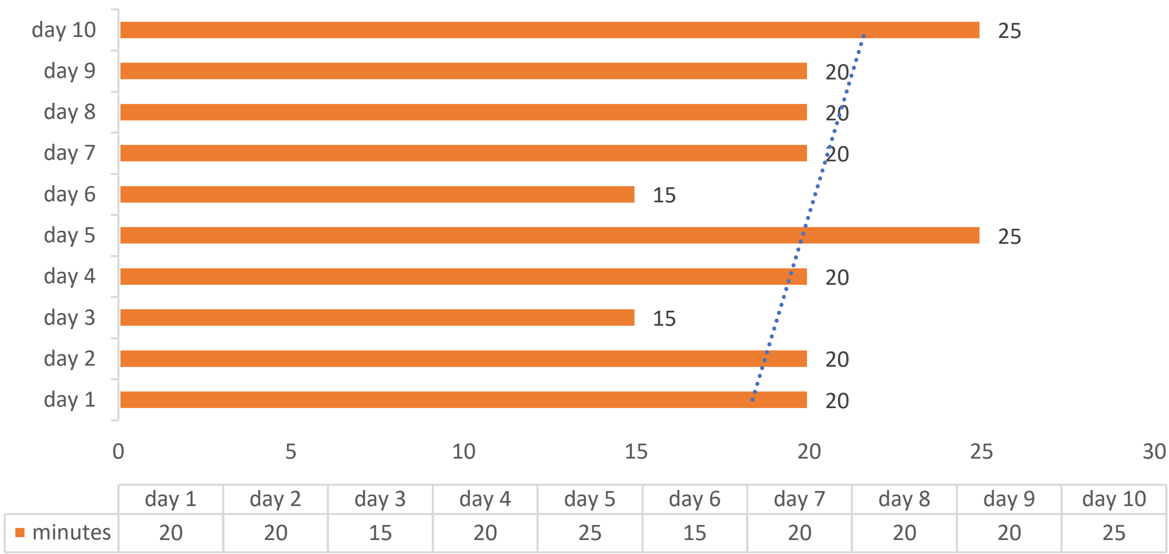


Figure 21. Initiating a patient on TB disease treatment (first visit)

At the beginning of TB treatment of a patient, during the first visit, a nurse spends an average of 20 minutes. At the same time, we can see that during each patient's first visit, the nurse is involved in communication with the patient and provides appropriate services.

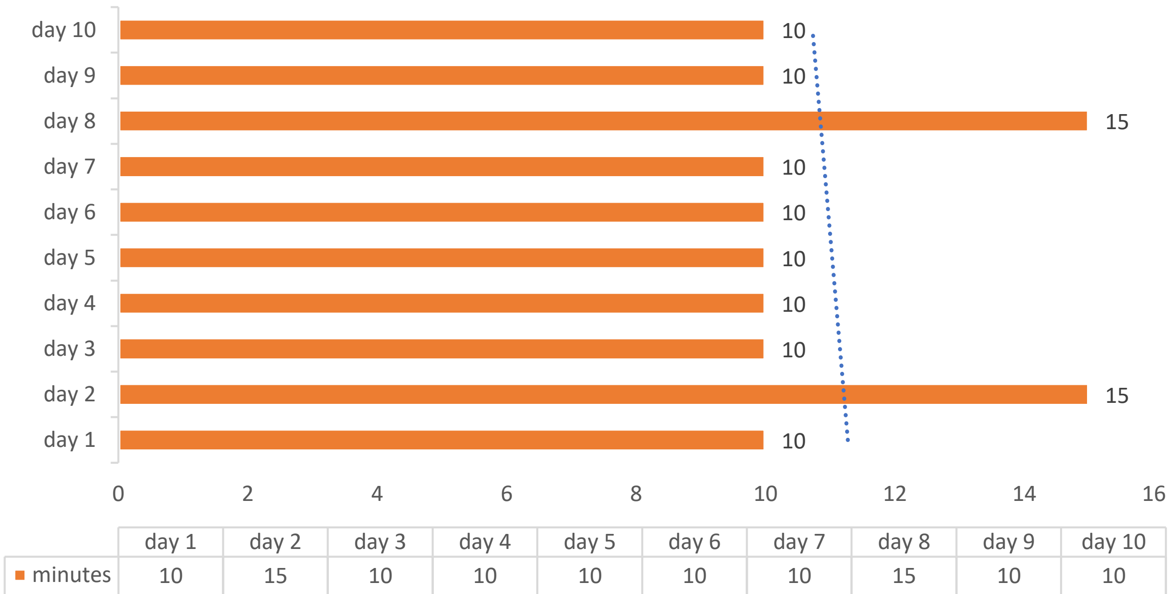
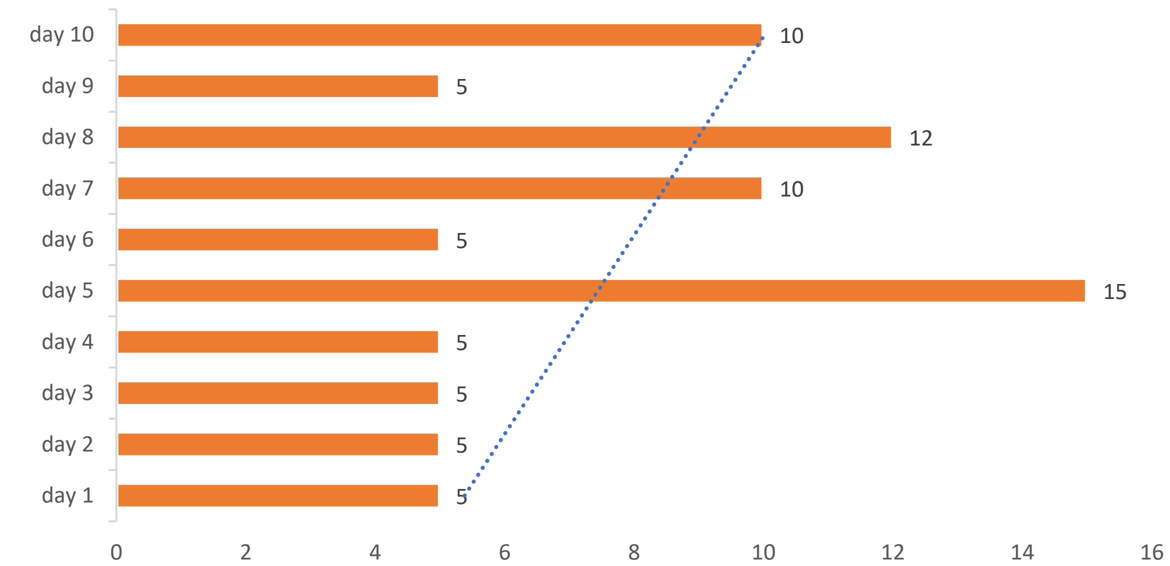


Figure 22. Seeing a patient for a follow-up visit during treatment

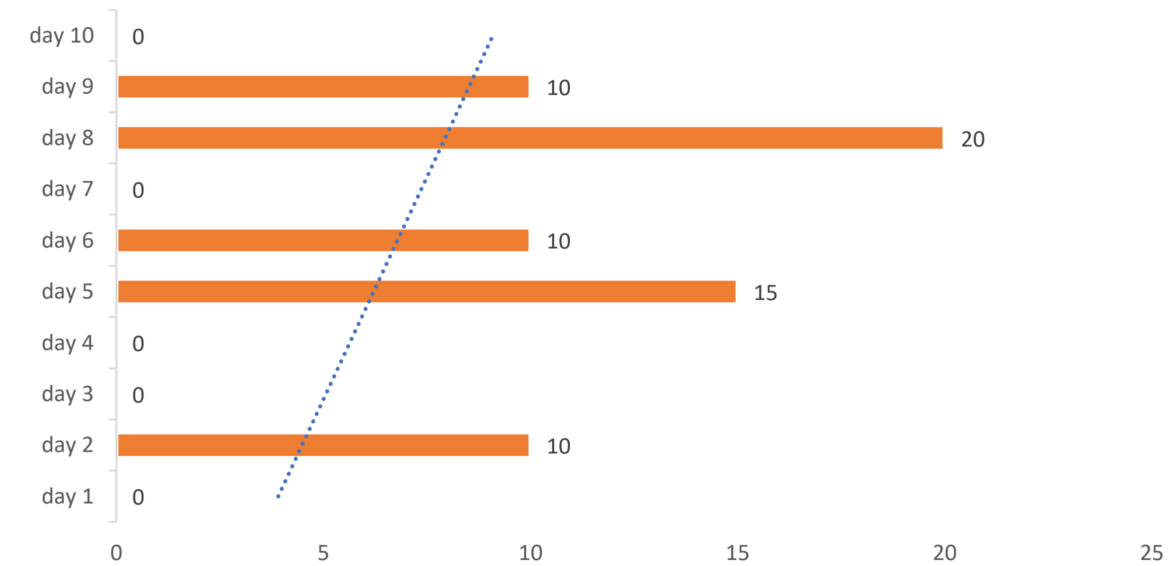
Visiting a patient for a second visit during treatment – on average, a nurse spends 10 minutes per patient.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
minutes	5	5	5	5	15	5	10	12	5	10

Figure 23. Seeing a patient for directly observed treatment

According to the trend line, examining a patient for treatment under direct supervision and involving a nurse in this process takes up to 8 minutes per patient on average.

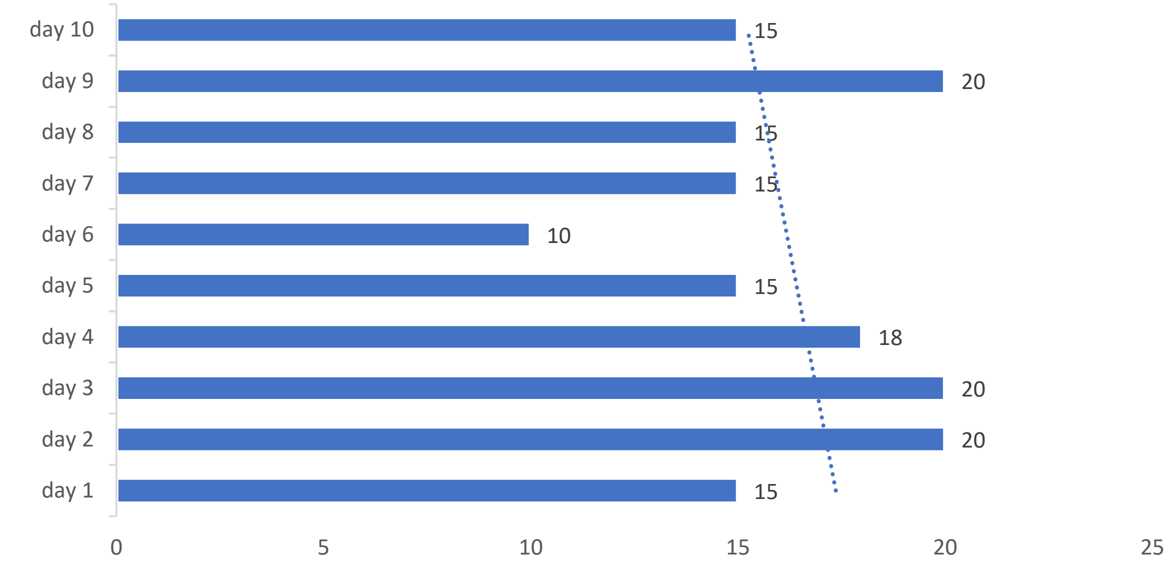


	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
minutes	0	10	0	0	15	10	0	20	10	0

Figure 24. Seeing a patient for their final visit once they complete treatment

After the patient's treatment is completed, the nurse meets with the patient for a final visit. The nurse takes approximately 10 to 20 minutes to provide the relevant services.

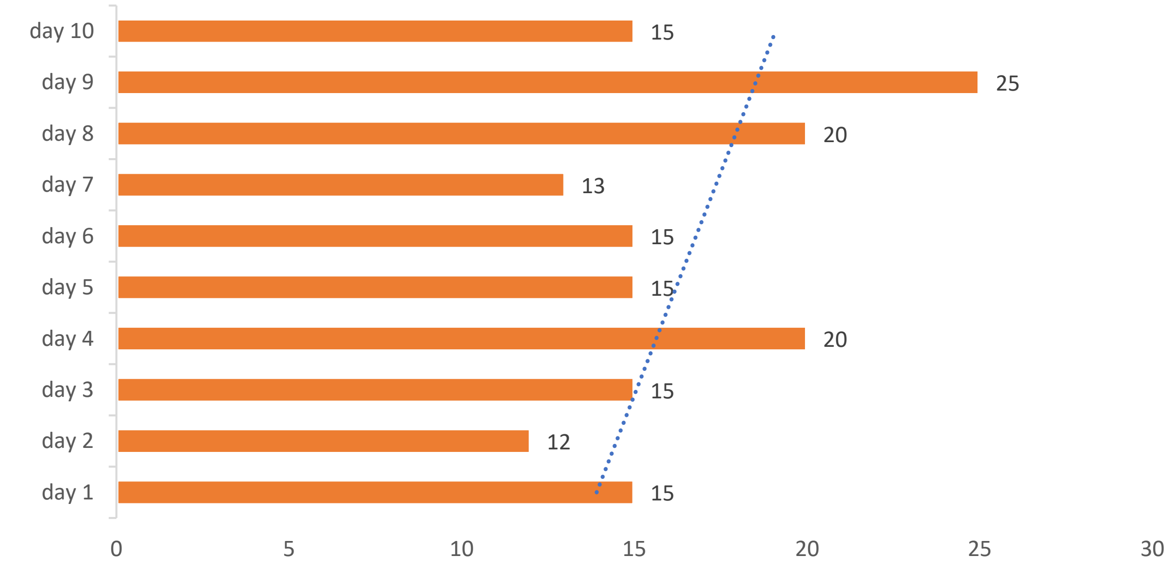
Here is a breakdown of the time required to perform a non-CAD X-ray interpretation provided by a radiologist.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
minutes	15	20	20	18	15	10	15	15	20	15

Figure 25. Interpreting a CXR without CAD

According to above Figure 25, we can analyze that the average time for a radiologist to perform an X-ray interpretation without CAD is about 16 minutes.



	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day 8	day 9	day 10
minutes	15	12	15	20	15	15	13	20	25	15

Figure 26. Preparing a person for a CXR and taking the CXR

The distribution of time for services provided by an X-ray technician to prepare a person for an X-ray and perform it is as follows, as shown in the graph above. We can see that the average time is 16 minutes.

Cost Estimation of CAD implementation

The main question is what does it cost to implement CAD, considering costs of (1) training staff on its use; (2) the software licensing and equipment; (3) the annual costs of software maintenance and support.

Therefore, it can be concluded that the introduction of CAD for TB diagnosis can significantly facilitate and speed up the diagnosis and treatment process, which can lead to time savings for healthcare professionals and improved outcomes.

After analyzing all the above data, which take into account the costs associated with the implementation of CAD, it should be noted that cost estimates consider the number of CAD licenses purchased (assume 1 CAD box is purchased per license, and boxes have a 5-year useful life), and the need for annual maintenance subscriptions.

It also considers costs of CAD installation and training: adjustable is the number of trainings provided for CAD along with initial installation, the frequency of CAD re-training, how long each training lasts, who provides re-training (assumes the CAD company gives the initial training), and the cadre and number of healthcare workers being trained.

Finally, it considers additional personnel costs associated with CAD, including time for a worker to upload images (the type of worker and time spent can be adjusted in the form above) and additional time, on average, a radiologist would spend interpreting each chest x-ray given they have additional information from CAD.

Having made the necessary calculations under the received data on financial costs, as well as the time costs of healthcare professionals, we can provide calculations that will demonstrate an analysis that considers a 10-year time horizon. We have annuitized all economic costs on this time horizon at a rate of 3% per annum.

Table 19. Financial and economic costs. Date populated for Lviv Region, price in USD

ANNUAL COSTS	Year 1	Year 2/5	Year 6	Year 7/10	Total
Financial Costs	$ 44 219	$ 11 494	$ 16 994	$ 11 494	$ 153 161
Economic Costs	$ 17 409	$ 15 684	$ 15 684	$ 15 684	$ 158 564

Table 20. Cost Estimation. Date populated for Lviv Region, price in USD

Input Costs	Year 1	Year 2/10	Total
CAD License Cost (Financial)	$ 25 500	-	$ 25 500
CAD License Cost (Economic)	$ 2 989	$ 2 989	$ 29 894
CAD Training and Installation Costs	$ 2 300	$ 575	$ 7 475
Staff Training Costs	$ 499	$ 499	$ 4 991
CAD Maintenance Costs	$ 10 200	$ 10 200	$ 102 000
CAD Equipment Costs (Financial)	$ 5 500	-	$ 11 000
CAD Equipment Costs (Economic)	$ 1 201	$ 1 201	$ 12 010
CAD Uploading and Interpretation Costs	$ 220	$ 220	$ 2 195

Tables 19 and 20 break down the annual and 10-year total costs for CAD implementation and use based on the values provided on the Cost Input and Data Entry Sheets. It also provides the various input costs considered. Financial costs represent the costs incurred that year for budget planning purposes, while economic costs take into account opportunity costs of these larger up-front costs (eg, CAD licenses and equipment) and annuitize these costs over the useful life or time horizon (whichever is shorter). Over the entire time horizon, economic costs will be more expensive than financial costs, because they consider opportunity costs; however, years when capital purchases (CAD licenses, equipment) are made, financial costs (budget impact) will be greater.

This scenario of cost forecasting, separating financial and economic costs, is an important tool for analyzing the cost of implementing CAD in Ukraine. This approach will allow management and financial analysts to estimate the total cost of implementation, taking into account both actual costs in a particular period of time and opportunity costs.

In addition, this approach allows us to understand the overall economic benefits of investing in technology. Taking into account opportunity costs helps to evaluate cost-effectiveness and rationalize the decision to implement CAD.

However, the effectiveness of this approach may depend on the accuracy of opportunity cost forecasting and the correctness of accounting for depreciation costs over the useful life of the technology.

For a particular context, it is important to have access to reliable data on opportunity costs and possible depreciation costs, and to analyze their impact on the cost-effectiveness of implementation. To this end, we have performed cost projections and taken into account all the necessary factors. Accordingly, it can be concluded that this approach will be useful for the effective management of CAD implementation in Ukraine and for maximizing its benefits.

Conclusions

Chest X-ray is an important method for detecting pulmonary TB. Including X-rays in screening and selection programs can help to interpret X-rays, to detect TB earlier by providing a preliminary screening and identifying suspicious cases that require further clinical evaluation. The overwhelming majority of respondents have heard of computerized/automated diagnostic systems and know that CAD systems are not currently used in Ukraine. After conducting a thorough analysis through in-depth interviews with healthcare professionals involved in TB screening and national and regional TB experts, it is clear that the introduction of CAD has advantages and disadvantages.

Speaking about the potential advantages and disadvantages of implementing CAD systems in the medical industry, participants expressed the opinion that using CAD can speed up a radiologist's work, increase the efficiency and speed of examinations and obtain results. Accordingly, it can increase the number of patients examined per day.

At the same time, the participants expressed their concerns that there may be a lack of options for interpreting the results, and that there will be a need to refer to a radiologist in case of pathologies (communication problems). There are also concerns about possible overdiagnosis.

Speaking about the WHO recommendation to replace the X-ray interpreter, participants in the in-depth interviews expressed general opinions that it is better to implement CAD in Ukraine without replacing a human reader - it can be used for screening and triaging TB patients without a human reader. However, under national program guidelines and protocols, a healthcare professional may be involved in the preliminary diagnosis process and review the DST report for confirmation – this may be a radiologist, clinician or phthisiologist who is part of the screening program.

In general, the interviewees emphasized that the implementation of CAD systems in Ukraine could potentially face difficulties due to insufficient fluency in computer systems and skills in working with office programs in general.

Therefore, the introduction of CAD requires thorough training for healthcare professionals who will be involved in the examination of patients for TB.

To effectively implement CAD products into routine practice, the proper threshold must be defined to indicate possible TB cases and trigger further diagnosis. Establishing the ideal threshold for each case requires customization of CAD products to suit the local context and how they are intended to be used. Decisions need to be made on the purpose of the screening and the allowable costs.

In other words, it is important to determine the right level of sensitivity or accuracy of the CAD system so that it does not miss possible TB cases and does not trigger too often without proper reason. This requires customizing CAD products for specific use cases, as well as carefully considering the goals and budget for the TB diagnostic process.

To summarize the features of CAD calibration, it should be recalled that according to a toolkit to support the effective use of CAD for TB screening guide provided by WHO [3], the gold standard to compare CAD system performance and calibrate CAD system for further application is GeneXpert MBT/Rif.

As the routine Ukrainian TB testing algorithm does not involve the process of performing GeneXpert testing for all individuals after an X-Ray, unfortunately, it was not possible to collect all the required data (in particular, the required number of GeneXpert results) to perform a calibration study according to the WHO guideline – Taking into account an indicator «sensitivity» = 90%, the number of confirmed TB cases should be 138 cases and number of confirmed non-TB cases required (assuming the same precision and similar specificity as above) should be 138 cases (based on +-5% precision), with the overall enrolment size of 276 cases. Considering the realized sample based on the available results for GeneXpert (N = 145, positive = 78, negative = 67), the number doesn't meet the minimum sample according to the WHO protocol.

However, even with 145 available GeneXpert results, the analysis results show overall relatively high CAD system performance – the area under the curve (AUC) is 0.79 (95% CI: 0.72 to 0.87). With CAD system threshold based on Youden's index – 0.668 – the Sensitivity and Specificity of CAD system are 69.2% and 82.1% respectively.

When comparing results with the GeneXpert as a gold standard, the CAD system's Sensitivity is lower compared to Radiologist (62.9% vs. 94.9%), however, the CAD system's Specificity is significantly higher (82.1% and 29.9%).

In general, even on the relatively small sample size with available GeneXpert results, CAD system with a threshold of 0.668, based on Youden's index from the ROC Curve, showed worse performance in the context of predicting GeneXpert results; thus, according to the obtained data and the results of the analysis, implementation of CAD system in Ukraine could help decrease the number of unnecessary diagnostic procedures, what potentially could save both time of healthcare professionals and money of healthcare facility that provides TB testing. However, to fully understand both the financial and health impact of CAD system implementation in Ukraine, further prospective calibration research is needed to achieve required by WHO guidelines sample size and, thus, CAD system threshold determination.

Data analysis was performed using R version 4.3.1 (2023-06-16 ucrt) [4].

Estimating the costs of implementing CAD systems for TB helps to maximize results and ensure more efficient and targeted use of these technologies in medical practice. Here are some ways that CAD can help save time for healthcare professionals:

- Automated image processing: CAD systems can automatically process lung X-rays, identify pathological signs of TB, and highlight them for further analysis. This can significantly speed up image analysis, which is usually time-consuming.

- Reduced time for interpretation: CAD systems can quickly analyze large amounts of data and detect TB signs, allowing doctors to speed up the process of interpreting results and making decisions.
- Automated recognition: CAD can automatically detect pathological changes in images, helping physicians focus on confirming the diagnosis and prescribing treatment.
- Disease monitoring: CAD can monitor changes in lung images in terms of identification and treatment effectiveness, which helps physicians respond more quickly to changes and choose the best treatment strategy.
- Collaboration with other systems: CAD systems can integrate with other medical systems to automate data exchange and facilitate collaboration between healthcare professionals.

All of these benefits allow healthcare professionals to save time analyzing and interpreting images, which in turn leads to faster diagnosis and treatment, which is important for TB patients. During the study, CAD was purchased, and the process of preparing for installation at the Pulmonary Health Center has already begun, which will facilitate more efficient and appropriate TB screening, early diagnosis and timely treatment.

Financial costs and economic costs are important for the cost estimation of CAD implementation. Financial costs include the money actually spent on the purchase of equipment, software, staff training, payment for services, etc. Economic costs help to provide broader insights into the impact of CAD on TB detection performance.

In this regard, the analysis of the cost estimates of CAD implementation suggests that the introduction of CAD systems will not lead to a reduction in financial costs. The highest costs will be incurred in the first year of CAD implementation, including budget planning costs.

Thus, in Ukraine, CAD is not a replacement tool, but rather a tool to speed up the interpretation of CXR, so individuals with abnormalities can be tested for TB faster. A trained radiologist will still read all images, but a screen/no screen decision for many individuals can be made quickly.

References

1. Wikimedia Foundation. (2023, September 9). F-score. Wikipedia. https://en.wikipedia.org/wiki/F-score
2. Google. (n.d.). Classification: Roc curve and AUC | machine learning | google for developers. Google. https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc
3. Determining the local calibration of computer-assisted detection (CAD thresholds and other parameters). A toolkit to support the effective use of CAD for TB screening. https://apps.who.int/iris/bitstream/handle/10665/345925/9789240028616-eng.pdf
4. R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.