Adaptive Health Questionnaires: Methods, Implementation and User Impact

Vladyslav Malanin

[1]M.D., PhD Student, Department of sensor devices systems and technology contactless diagnostic, V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine

Abstract

The need for personalized, streamlined, and scalable ways to evaluate health is driving the advancement of adaptive health questionnaires that rely on computerized adaptive testing (CAT) to ensure questions are the most relevant, reduce the effort needed from patients, and guarantee accurate results. This article looks closely at the ideas behind, logic of, software details for, and ethics of using adaptive questionnaires in today's health systems. Using IRT and advanced machine learning techniques, adaptive questionnaires automatically adapt to each user's responses as they take the test, resulting in a more accurate measurement with a much shorter exam. This article covers essential aspects of adaptive assessment, for example, items under IRT (1PL, 2PL, 3PL), item bank creation and calibration, real-time scoring, and criteria for stopping answers.

Both the obstacles and new solutions related to digital health platforms, such as working with wearables and mobile health systems such as Apple HealthKit and Google Fit, are discussed. Important points about UX/UI design are explained, mainly highlighting that emotionally safe, accessible, and open solutions should be provided for adaptive interactions. Results are provided from both simulations and actual use by many people, which show that their suggestions are accurate, take less time, and are appreciated by all users within the study group. Special efforts are made to ensure that adaptive assessments do not cause unfairness, invade privacy, result in real-time changes, or lack clear instructions about consent.

It concludes by noting several difficulties yet to be overcome: using adaptive technologies for many health conditions, upgrading item banks as needed, interpreting AI-powered decisions, linking adaptive systems to national health systems, and applying them during emergencies and crises. It demonstrates that using adaptive health questionnaires helps achieve progress in precision medicine, collecting helpful data for public health, and adapting care to the needs of each user. People are asking for global standards, combined government efforts, and long-term studies to guarantee ethical and fair use of these tools across the world's health care.

Keywords: Computerized Adaptive Testing, Digital Health Assessment, Item Response Theory, Personalized Health Monitoring, Ethical AI in Healthcare

1. INTRODUCTION

1.1. Background of Health Questionnaires in Clinical and Digital Health

For many years, using health questionnaires has played a vital role in medicine by guiding doctors to collect PROs, check for health problems, assess wellness, and measure patients' quality of life.

Historically, surveys such as the General Health Questionnaire (GHQ), the Patient Health Questionnaire (PHQ-9), and the SF-36 have been chosen, as they are confirmed to measure people's psychological and physiological status. Thanks to tokens, doctors and nurses can discover symptoms, assess patient responses to treatment, and group individuals by their risks. They are involved in monitoring entire populations and developing healthcare policies for the public. Since digital health is on the rise with more mHealth apps, EHRs, and wearables, more people can now use health questionnaires. Using digital technology, experts can now quickly and inexpensively distribute mobile data, capture real-time information, and link the information to that gathered by fitness trackers and smart watches. As a result, data can be gathered over extended periods and monitored remotely thanks to modern technology. Still, adopting new methods in medicine creates challenges for the traditional questionnaire approach.

1.2. Limitations of Traditional Fixed-Form Questionnaires

Even though fixed-format questionnaires are reliable and frequently used, their weaknesses make them less effective, accurate, and user-friendly. The main problem is that these questionnaires hold the same questions for every participant, even if the inquiries are not relevant. Therefore, some users have to see the same information twice, while others have to deal with only a basic explanation, making the balance between being concise and having many details less than perfect. Since fixed questionnaires are not very flexible, they may lead to increased survey fatigue in people with certain health conditions or low computer skills, making it more difficult for them to complete the questionnaires or answer properly. Besides, all items in fixed-form assessments are valued the same, so their information value is not taken into account. Because of this, assessments are not as individualized to the needs of a respondent, and the results lack detail and become more difficult to understand. When asking questions on the internet, a static survey is rarely able to create the immediate and highly personal user experience clients desire. Hence, it is necessary to develop data-driven methods that optimize both the questions asked and their usefulness, all the while remaining reliable.

1.3. Emergence of Computerized Adaptive Testing (CAT)

Using IRT and adaptive algorithms, CAT was implemented as a better approach to fixed-form tests. Begun in testing, CAT is now being used in healthcare to vary the approach for each patient and ease the pressures on those being tested. To sum up, CAT uses each respondent's answers to decide on and present the subsequent item, which helps to estimate their trait by asking fewer relevant questions. With this technique, both the accuracy of measurement and the length of the test can be increased without lowering its validity or reliability. For example, CAT is being used in NIH PROMIS (Patient-Reported Outcomes Measurement Information System) in health settings with improved speed and satisfaction among patients. The growth of CAT in digital health goes hand in hand with current improvements in the field. CAT, when placed in health apps, EHRs, or devices for smart watches, makes it possible to continuously keep an eye on a user's health with little effort on the part of the user. Due to machine learning, current CAT algorithms can easily keep pace, which makes them the core of the next-level technology in digital health assessment.

1.4. Objectives of the Article

This article explores in detail how health-related questionnaires are designed, implemented, and used to impact people. At the beginning, the paper discusses the factors involved in health-related adaptive testing, such as Item Response Theory, selection of suitable items, and the way item banks are formed. Second, it studies the technological tools needed for deployment and shows how they can be used with EHRs, mHealth offerings, and devices for wearables. Experts need to focus on user experience design, data security, and interoperability to ensure a product is used successfully. Moreover, the paper examines what has been learned from deploying CAT in healthcare, focusing on measures including how quickly CAT responds, the accuracy of its findings, and users' satisfaction. Lastly, it talks about the problems related to deploying adaptive questionnaires, such as fairness, clarity, and diversity. Many researchers, clinicians, and developers may find this article useful for understanding how to use adaptive questions to personalize, scale, and improve the efficiency of digital health evaluations.

2. Theoretical Foundations of Adaptive Health Questionnaires

2.1. Concept of Computerized Adaptive Testing (CAT)

CAT adjusts the test topic and order of questions during the assessment depending on a person's previous responses. The theory behind CAT relies on Item Response Theory (IRT), a group of models that predicts the likelihood of a correct (or affirmative) answer to a question based on a person's traits and the specific parameters of the question. While classical test theory views all tests as having the same influence on the composite score, IRT provides different levels of discrimination, difficulty, and guessing to each item. Consequently, CAT systems can provide the most relevant item to each respondent at every stage, making the CAT test more efficient and accurate. When someone answers mental health questions showing a high level of resilience, the algorithm can jump to tougher items that are most revealing for those at the top of a mental health scale. Because of this, the test is shorter, requires less time from respondents, and maintains or improves its rigor.

Because CAT is adaptive, it can quickly help estimate someone's underlying characteristics. This type of questionnaire makes every person answer the same questions, regardless of whether they are needed, leading to both extra thinking for the users and tiredness for the respondents. Furthermore, it is difficult to adjust for different responses within individuals or for the functioning of items in subpopulations, so their effects cannot be generalized. CAT's algorithms, along with Bayesian methods or maximum information criteria, help personalize the assessment process for every user and allow for testing with a smaller number of questions. Healthcare professionals must be efficient, as time is short and both mental and physical effort need to be kept low, mainly for older or frequent patients. There are no special requirements for using CAT; digital health apps are easily updated to include it. Consequently, CAT allows doctors to have interactive conversations instead of just using questionnaires, making health tools better suited to personalized medicine.

2.2. Historical Evolution and Adoption in Healthcare

CAT began in educational testing research of the 1970s and 80s, as a way to administer common tests like the GRE and ASVAB more smoothly. With the IRT framework, the first CAT models made it clear that a trait could be accurately measured using a smaller number of items than in traditional exams. When computers got stronger and user interfaces became more advanced, CAT techniques were applied in psychology and behavioral health. One of the main reasons for the shift to healthcare was the Patient-Reported Outcomes Measurement Information System (PROMIS), which the U.S. National Institutes of Health promoted. PROMIS used item banks calibrated with IRT to create adaptive tests for pain, fatigue, functioning, and mental health, proving that CAT can be used routinely in health settings. Next, CAT tools were made available in electronic health records, telemedicine, and systems for monitoring public health. Today, adaptive screeners for depression can change in real time based on what patients report, giving feedback that aids in diagnosis and deciding on a treatment plan.

The use of CAT in healthcare has been encouraged by the growing emphasis on value-based care and precision health, since these methods need simple, meaningful data that is easy for patients and staff to use. CAT systems perform these roles by speeding up the way assessments are done and allowing careful monitoring of how patients are doing. Various research studies have confirmed how valid and reliable this approach is, and its flexibility means it can benefit people in primary care, rehab, cancer, and mental healthcare settings. Also, CAT works well in long-term healthcare, where doctors need repeated evaluations that notice important differences. Because CAT changes the selection of items depending on earlier answers and trends, it allows for continual tracking of how a condition changes and responds to therapy. As a result, CAT has progressed from being used in academics to becoming a central part of health practice today.

2.3. Role in Precision and Personalized Medicine

With precision medicine focused on unique genetic, environmental, and lifestyle patterns, highly accurate measurement tools are necessary. CAT meets this need by providing exams that are psychometrically reliable and also match the needs of each person and culture. Using frail data, precision medicine helps identify groups of patients, guides actions taken by doctors, and tracks the results; adaptive questionnaires play a crucial role in efficiently and effectively collecting this data. By adjusting the test to each patient, CAT provides clearer and more acceptable results than fixed assessments that have the same structure for everyone. Because of personalization, people are more likely to follow the study rules and provide all their data, helping to ensure good data quality in both clinical trials and other care.

Also, CAT is capable of processing different types of health data, including what wearable devices and mobile apps record. When items in adaptive questionnaires are chosen based on a person's health at the time, the data helps create a more comprehensive picture of patient health. In the case that the wearable data records a rise in heart rate variability and poor sleep quality, the mental health screener may start with questions about stress and anxiety. As a result, we can look forward to improved health monitoring systems that adapt in real time and can personalize appropriate actions. These policy tools help to reduce the influence of language, literacy, and culture in assessing health, using adaptive ways and many languages. In general, CAT plays a significant role in transforming patient involvement and clinical data gathering, keeping the main ideas of personalized, predictive, and participatory healthcare.

Table 1. Comparison of Fixed vs Adaptive Questionnaires

Feature	Fixed-Form Questionnaires	Computerized Adaptive Testing (CAT)
Assessment Length	Typically long and standardized	Variable and optimized per respondent
Measurement Precision	Uniform across population	High precision tailored to respondent level
User Burden	Higher due to redundant questions	Lower due to relevant, concise questions
Relevance of Items	Same for all respondents	Personalized based on prior responses
Scalability in Digital Health	Limited interactivity	Seamless integration with mHealth and EHRs

Adaptability to Change	Static and inflexible	Dynamic, can evolve with user data

3. Methods and Algorithms in Adaptive Health Assessment

3.1. Item Response Theory (1PL, 2PL, 3PL Models)

The main mathematical basis for CAT in assessing health is Item Response Theory, which estimates the probability of an action based on someone's hidden skill level. Unlike classical test theory, IRT models examine the details of each item by its properties and how they work with the respondent's underlying trait. The one-parameter logistic (1PL) model, or Rasch model, believes the likelihood of a response depends on item difficulty and that all items can detect differences just as well. Using math, it says the chance of getting the answer right equals a logistic function of the difference between what the person knows and how hard the question is. In the 2PL model, an extra parameter called item discrimination scores can measure how well an item tells the difference between people with different levels of the trait. Because respondents might answer questions out of pure chance, the 3PL model introduces a guessing parameter, useful in educational settings and also when social bias is a factor in health-related surveys.

All of these IRT models are used to identify the most helpful items in adaptive testing. The degree to which an item reduces a respondent's uncertainty about their trait is influenced by its information function. Because health settings often look for subtle symptoms and changes, using these models ensures the algorithm gives importance to items with the right level of difficulty and the best separation for each individual. In clinical care, having exact estimates at the item level is vital, as mistakenly identifying a trait as greater or lesser can cause the wrong response. Also, advanced applications usually rely on MIRT models to study more than one trait at a time, making it possible to gain a richer and more complete picture of the patient's health problems.

3.2. Item Bank Design and Calibration

A high-quality adaptive health questionnaire is built around a quality and vast item bank that provides questions that have already been calibrated by experts. The item bank ought to span all the degrees of the trait and feature items that are well-suited to the needs of the group you want to test. When making an item bank, items are first developed, tested on a sample, and verified through the application of IRT modeling. Difficulty, discrimination, and guessing parameters are estimated by providing the items to a large and varied group of test-takers. The pattern of each item's response is now examined using ICCs and item information functions to assess its suitability for various parts of the trait spectrum.

Item banks for health applications should take into account different languages, levels of reading ability, and cultural differences. Frequently, researchers in international or multicultural studies need to calibrate and validate their tools in multiple languages. Additionally, differences in baseline conditions, such as limited mobility in the elderly versus acute injuries in the young, might require items to be calibrated separately for each group. It is important to consider the frequency of exposure to items during design, as often selecting the same item over and over can cause biases or issues related to test security. Therefore, item banks tend to have controls in place to ensure usage is uniform through the item pool while keeping the accuracy of the test. Test adaptivity and efficiency are greater when the item bank is larger, because it gives finer control over target items and trait assessment.

3.3. Algorithmic Flow of Adaptive Questionnaires

An adaptive questionnaire's algorithm includes three main steps: selecting items, evaluating answers, and deciding when to end the questionnaire. At the core of CAT is a loop that iteratively improves the estimate of the trait level after every item is answered, using IRT-based maximum likelihood or

Bayesian techniques. In the beginning, data analysis assumes the trait has a certain value and picks the item that reveals the most information at this presumed point. Based on the respondent's response, the algorithm adjusts the trait estimate and chooses the following item that helps the most in reducing the error, so every new item adds value to the estimation.

The score from each answer is processed directly, allowing the individual's trait estimate to update instantaneously. This type of estimation makes the accuracy of the solution better and allows the circuit to react immediately and respond to changed inputs. The stopping rule is commonly dictated by following a predetermined measurement error limit, a fixed quantity of items, or both. As a result, the algorithm could stop once the trait is estimated reliably or after examining a maximum of 10 items, whichever takes place first. Because some patients might get tired or experience discomfort, reducing the length of tests in health applications is crucial. Thus, the system is designed with a strong focus on keeping instructions brief and accurate. In more advanced setups, stopping criteria may use what the user says, how they respond, as well as factors like how long the session has lasted or detected stress.

3.4. Use of Machine Learning in Adaptive Logic

In classical CATs, IRT is central, but using machine learning has made modern health questionnaires quicker and easier to adapt, personalize, and predict. Training a machine learning model on lots of user responses can predict the next best thing to present to a user or help decide which items should come after one another to maximize engagement. Using decision trees or gradient boosting, supervised learning finds out how health outcomes are affected by items and can better forecast using these results. With clustering, unsupervised approaches help label people into subgroups and make it possible to develop customized items and programs that match their abilities.

Due to its approach to handling uncertainty, reinforcement learning, as a type of ML, is ideal for quickly updating health questionnaires. The CAT system's purpose is to help you find the right item (action) by increasing information gain (reward) and decreasing your workload (penalty). As time goes on, the system develops perfect strategies for things you purchase, making the service more efficient and customized. In addition, neural networks can help model complicated responses, providing better results in multimodal cases that use both speech, behavior, and physiological data. With these models, systems can notice when a user's state changes and modify the questionnaire to fit the new situation. The application of machine learning also allows the questionnaire system to adjust itself over time as it collects more information, which is very useful in digital health areas that are always adapting.

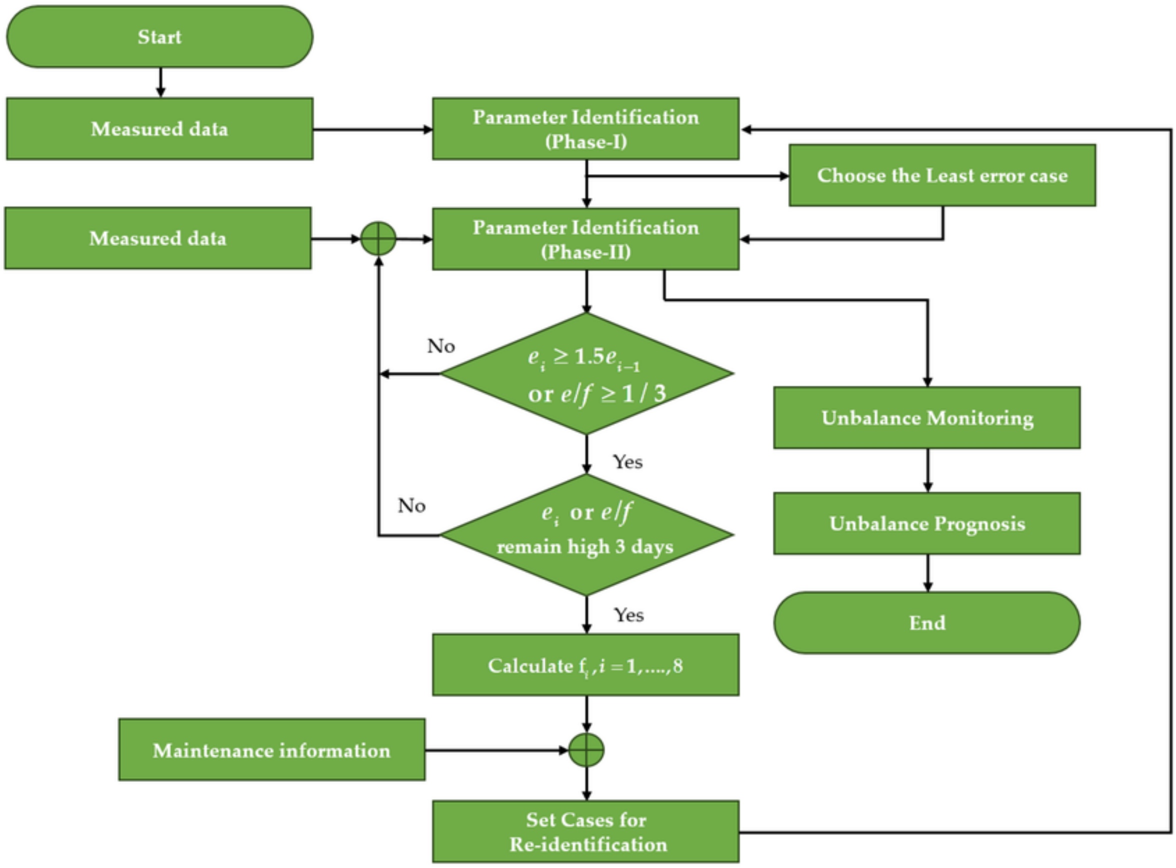
3.5. Adaptive Psychometrics in Health vs Education Domains

Whereas adaptive testing was first developed in education, it has unique requirements when applied in health that call for the use of health-specific psychometric rules. Within education, keywords like cognitive abilities are not easily influenced, so it is mostly done as a statistical, objective process. In comparison, characteristics related to health can be shaped by multiple things, often change over time, and are affected by people's opinions, other health issues, and the situation at hand. For this reason, health-based systems should allow for flexible structure, notice rapid patterns, and consider feelings and mental state better.

Also, when it comes to health assessments, there is more pressure, and the outcomes usually matter more to individuals than in testing at school. Misclassified results in depression screening instruments can lead to missing a case or performing ineffective treatment, which negatively affects patient health. So, CAT systems created to measure health should place importance on psychometric accuracy, as well as validity, fairness, and understanding in the field. Health assessments also differ in that they have to be suitable for a wide range of individuals, whereas educational tests are usually designed for similar people, such as students in one school. Because people from different cultures exist, it is important to design items, translate them, and calibrate tests so they are inclusive.

Also, education often sees adaptive testing being used as a summative tool to assess whether individuals have the required knowledge or should move on. Usually, the main goal in health is to help plan care, track improvement, or individualize treatment with imaging. For this reason, CAT tools focused on health should be able to detect soft differences in patient health over time and connect with different types of health data. Increased use of CAT, wearable gadgets, and telehealth raises the need for psychometrics that can reflect the ever-changing and complicated nature of modern healthcare.

Figure 1: Adaptive Questionnaire Flowchart with Real-Time Response Routing



4. Implementation in Digital Health Platforms

4.1. Integration with Wearable and Mobile Health Systems (e.g., Apple HealthKit, Google Fit)

Health monitoring and self-assessment are being reinvented as adaptive health questionnaires are added to digital platforms. When adaptive questionnaires are added to mobile health systems and wearables such as Apple HealthKit, Google Fit, Samsung Health, Fitbit, and Garmin Connect, they can work in a fuller and more active data environment. These services collect physiological and behavioral measurements obtained by many connected devices and present them all together. Adaptive health questionnaires can choose which questions to ask by matching with common APIs and permission protocols. When the device notes both a fall in regular exercise and a quickened heart rate while resting, the depression and chronic fatigue screener could first pay attention to items about energy, motivation, or psychosomatic issues. As a result, these polls can send out more useful and personalized insights at a faster speed. Furthermore, combining with mobile health apps makes it possible for the questionnaire to process user sensor data and move data collected in other ways to the user's health record or others taking care of the patient.

4.2. UX/UI Design Considerations for Adaptive Interaction

When making adaptive health questionnaires user-friendly, it remains important to blend simplicity, the ability to adjust to users, and overall engagement for the sensitive and possibly distressing nature of questions about health. Fixed-form surveys are less complex than adaptive questionnaires because the former do not need to vary their sequence of questions, type of interactions, ways responses are provided, or visual appearance based on the respondent's prior answers or ongoing input. For this reason, the interface has to be quick, aware of where the person is in the experience, and adjust to how the user's journey is unfolding. It covers giving users with special needs more accessible options, for example, larger text for blind or low-vision users, voice responses for those who can't use a keyboard, and multilingual features to support different cultures. Showing users only the information they need at each point is the basis of adaptive UX design, helping keep mental effort to a minimum and making the experience clearer. Using progress bars, confidence meters, or adaptive feedback prompts adds color to the system's performance and proves that adjustments are being made for the user's benefit. To maintain the same user path with variable questions, the interface needs modular parts that are updated as the backend sends new configuration data. Using helpful microcopy, caring tone, and positive feedback in design also helps users remain interested, especially during situations where the platform discusses trauma, suicidal feelings, or lingering illness. For that reason, using UX/UI well in adaptive health platforms is an obligation relating to ethical aspects and how complete and trustworthy the records become.

4.3. Data Collection and Privacy Challenges

The use of adaptive questionnaires online introduces many issues concerning data security, privacy, and giving proper consent. Adaptive systems collect information from different contexts and, when linked to sensors or previous medical charts, can provide thorough and unique profiles of patients' health. The valuable information obtained from analyzing patients' medical records increases the chance that someone could abuse the data, identify patients, or access them unauthorized. Since adaptive systems are flexible, people may see different sets of questions and receive different sets of instructions, which makes traditional approaches to consent problematic. Because of this, digital solutions should always comply with strict data protection standards such as GDPR, HIPAA, or CCPA, depending on the location. This means, users must agree to the terms, data is not kept without reason, it is encrypted, and there are well-defined rules on how long data is kept. These platforms should use privacy-by-design ideas such as different access controls based on roles, records of data uses, anonymization when needed, and dashboards for controlling and viewing data sharing permissions. Additionally, it becomes crucial for users to see how their information is affecting the outcome of the questionnaire in situations where decisions are made about health care. The IEEE P7000 and ISO/TS 82304 are new standards that guide ethical AI and trustworthy digital health systems, insisting on a proper relationship between adaptability and privacy, fairness, and user control.

4.4. Backend Architecture: APIs, Data Pipelines, Feedback Loops

The system that underpins adaptive questionnaires must be flexible, able to scale as required, handle countless user requests, and manage errors and response time efficiently. There is generally a RESTful or GraphQL API in the architecture that helps talk between the frontend and core services, such as item selection, scoring logic, user profiles, and audit logging. Since latency has to be minimal, these APIs are required for support next-best item searches, recording the user's answers, and restoring trait values in real time. The adaptive engine is hidden behind the API layer and uses both calibrated IRT logic and machine learning models, running in containerized microservices to ensure the solution can scale and be easily maintained. The majority of response information, metadata, and device information, plus optional sensor input data, are ensured safe storage in secure cloud databases by a central data pipeline using services such as Google Cloud Healthcare API, AWS HealthLake, or Microsoft Azure Health Data Services. Then, the data is applied to adjust the service's models, offer users more personalized options, and help both end users and clinicians with dashboards for visual analysis. Feedback loops make it possible for this framework to learn from user activity, high dropout rates, or negative moods and then adjust its item selection process or detect potential users who may be at risk. Thanks to CI/CD, any changes to the item bank, scoring process, or privacy are applied without any breaks in the system. In addition, backend systems should be compatible with outside systems such as FHIR-based EHRs, telemedicine services, and research databases, relying on adherence to codes such as HL7, SMART on FHIR, and OAuth 2.0 authentication. These elements work together to form a technology foundation that makes personalized, up-to-date, and clinically important assessments possible.

4.5. Real-World Case Studies (e.g., Life Calculator)

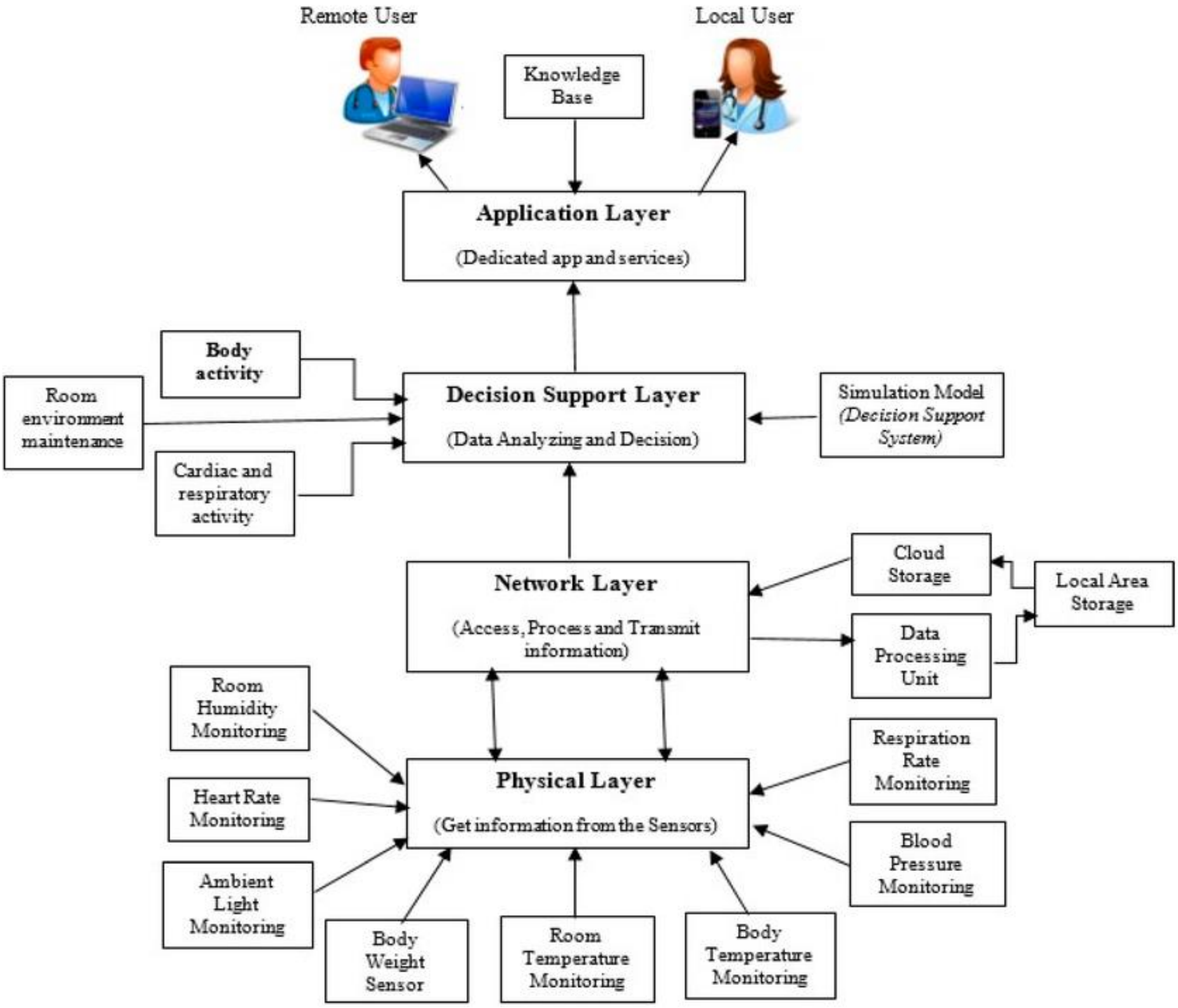
The National Academy of Sciences of Ukraine received a dissertation that contained the Life Calculator, an excellent real-world case of implementing adaptive health assessment. This resource effectively transforms adaptive algorithms into a digital health tool that can be used by many users. The data from over 260,000 users allowed Life Calculator to connect with Apple HealthKit and Google Fit, adjusting its questionnaire flow according to details like age, lifestyle choices, and test results. Users who showed signs of cardiovascular risk were asked specific questions about their sleep, exercise habits, and ECG information. When new data became available, Random Forest and XGBoost were used to instantly update the life expectancy predictions on the backend. Its modular approach made it possible for the system to work well with different national health registries, wearable gadgets, and external databases. The frontend interface looked clear and gave users the freedom to navigate on any device. It also modified the presented items, featured progress indicators, and personal suggestions on health. On average, it took 4 minutes and 15 seconds to complete the course, and around 77% of users completed it, showing that adaptive learning supports both engagement and less tiredness. The use of the system in the Ukrainian Ministry of Health's digital initiative emphasizes that it is both important and flexible for society. The success of its portrayal in the media and its real effect on health services stresses that adaptive questionnaires are practical for public health, monitoring chronic diseases, and preventive care. The Life Calculator is an example of future systems that help people and that rely on predictive models, consider users' needs, and make systems from different countries and continents able to interact.

Table 2. Digital Platforms Supporting Adaptive Health Tools and Their Capabilities

Platform	Data Sources	Adaptive Capability	Integration Support	Use Case Examples
Apple HealthKit	iOS wearables, apps	High (via APIs)	FHIR, REST API	Cardiovascular monitoring, fitness tracking

Google Fit	Android wearables, apps	Medium-High	REST API, OAuth	Activity logging, sleep tracking
Fitbit	Proprietary sensors	Medium	SDK, Web API	Heart rate, step count, stress detection
Garmin Connect	Sports sensors, GPS	Medium	REST API	Athletic training, chronic condition care
Samsung Health	Mobile & wearable data	Medium-High	SDK, REST API	Stress, activity, heart rate trends

Figure 2: Architecture of an Adaptive Health Questionnaire System



5. Experimental Validation and Performance Evaluation

5.1. Metrics: Accuracy, Completion Time, Drop-out Rates

The effectiveness of adaptive health questionnaires should be measured using several different performance factors, both addressing psychometrics and the needs of users. The key performance metrics are accuracy, how long it takes to complete each activity, and the number of times users leave, showing different perspectives on how the system works. Accuracy showcases the extent to which the adaptive system's evaluation of traits agrees with clinical standards or with values determined by experts. In psychometric terms, these tests are evaluated by seeing how closely their estimated scores line up with results from interviews or standardized assessment tools. If the adaptive logic matches consistently across different regions, it supports both its internal structure and its wider applicability. From when an incident is opened until the user submits their final response provides a way to judge both efficiency and how much the user had to do. Adaptive systems should make tests shorter to complete, while still ensuring accurate measurement, and this balance can be measured objectively by reviewing the average and median time it takes diverse cohorts to finish the test. The dropout rate lets us know about engagement, the challenge of the questions, and how comfortable users are using the system. People leaving the test at a high rate might be due to UX issues, unsuitable items being used, or the way questions are presented. They serve as key elements of evaluation protocols and lend real support to claims about the performance of systems.

5.2. Benchmarking with Traditional Methods

Comparing adaptive questionnaires with standardized, unchanging forms is a main way performance validation in health assessment is done. When carrying out such studies, both systems serve similar participants at the same time, making it easy to check their efficiency, accuracy, and experiences. On average, adaptive questionnaires are more efficient than fixed-form versions in collecting data and reflect subjective information. Psychometric reliability equivalent to or even better than static PROMIS CAT forms can be reached with 40–70% fewer items in the adaptive versions. Thanks to these changes, adults with mental health issues, functional limitations, or chronic illnesses now have an improved chance to use these surveys. In addition, benchmarking helps detect important differences in bias and the range of test performance. Unlike fixed-form products, adaptive instruments can adjust assessment items and show less error because they select topics according to the individual's approximate level of the trait. In comparative studies, adaptive systems usually score higher for user satisfaction, mainly when the interface is made mobile. Such studies are key for validating adaptive systems, directing their use in clinical settings, acquiring regulatory clearances, and improving algorithm models with observed results.

5.3. Results from User Testing or Simulations

Checking adaptive questionnaires with real people and in simulations gives a good picture of how the system works in practice and in different scenarios. Adaptive questionnaires are used in user testing to study how well people use the tools, their engagement levels, and the concordance between the tools and clinical diagnosis. Live testing of the Life Calculator system during its national health deployment on 260,000 users showed an average completion time of 4.15 minutes, a 23% dropout rate, and an approved accuracy of more than 90% in important modules such as cardiovascular check and stress screening. Such achievements indicate that adaptive systems can be deployed widely in different kinds of environments. In contrast, simulations use either fake or previous response logs to test how the system performs under a range of situations, such as variations in characteristics, lengths of the item bank, and chosen algorithms. These results allow us to understand its ideal efficiency, how much data is needed to maintain a certain level of accuracy, and how the system functions when there are few data points or errors. Using the simulated screener, the researchers achieved a standard error of measurement below 0.3 using only six items, while requiring far fewer than the PHQ-9's nine items. The agreement between results from simulations and field data demonstrates that adaptive questionnaires are both correct in principle and helpful in practical use.

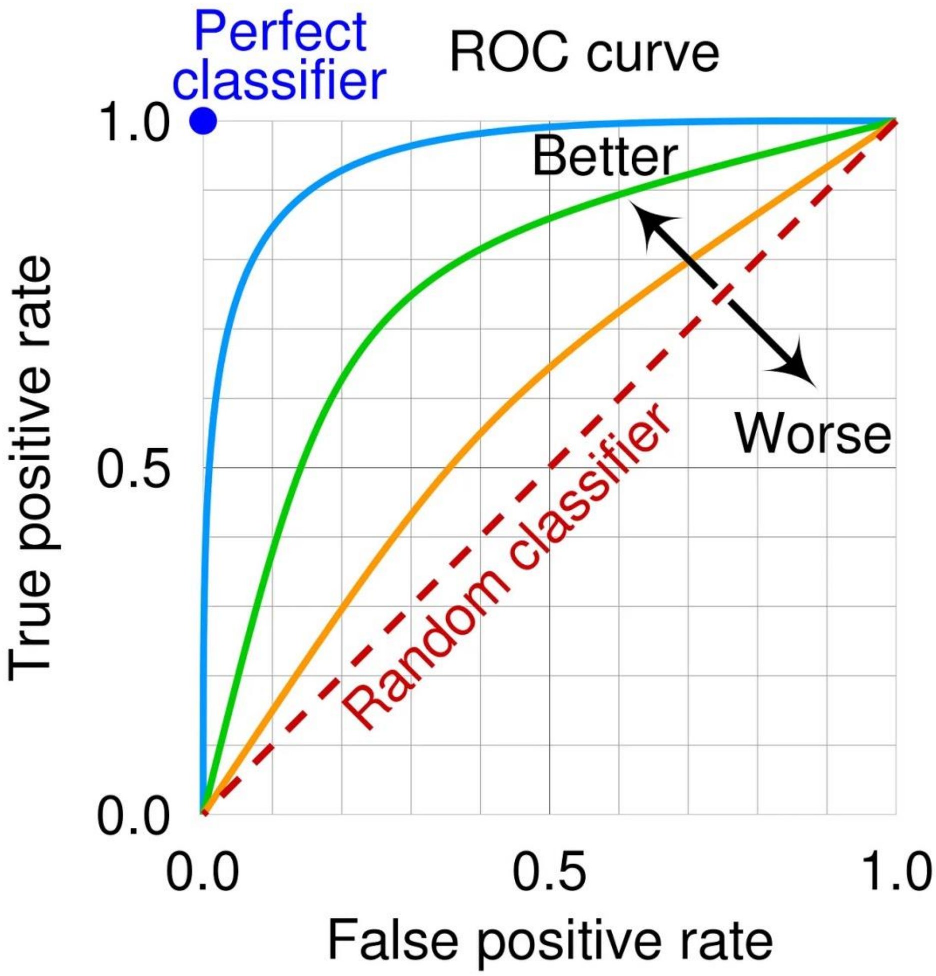
5.4. Interpretation of Key Findings and Implications

The tests and research demonstrate that using adaptive health questionnaires in online health assessment is a meaningful step forward, as it brings together strict statistics and easy use. As IRT-based and machine learning algorithms achieve high scores with many patients, it is clear that they support the delivery of clinical tests that are adapted to each person. The significant cuts in both the time taken to finish a test and the number of items, with no fall in reliable diagnosis findings, indicate that these tools are especially suitable for uses where time and attention are not easy to spare, including emergency rooms, rural clinics and self-assessment applications for the elderly and chronically ill. Because dropout rates are low when an application is used in real situations, we can assume users find the content helpful and are highly engaged with it, most probably due to smart UX design. Such findings will have significant impacts on future digital health care. In daily work, clinicians may use adaptive screeners to simplify intake procedures and repeat assessments, and health systems may implement them for large-scale population health care and support. Researchers can use adaptive systems to receive sharp, high-quality data that aids customized and precise measures in epidemiology and researching how best to help at the individual level. If evidence supports the findings, regulatory agencies might use them as a reason to add adaptive devices to regular care plans, as long as repeated validations maintain their safety, fairness, and interpretability. The conclusions demonstrate that technology should respond to people's requirements in healthcare, not the other way around.

Table 3. Experimental Results: Average Length, Completion Time, and Accuracy Across User Groups

User Group	Fixed-Form Length (Items)	Adaptive Length (Items)	Completion Time (mins)	Accuracy vs Gold Standard (%)
General Adults	20	6	4.2	91.3
Older Adults (65+)	18	5	4.6	89.8
Adolescents (13–18)	22	7	3.9	92.1
Chronic Illness Group	25	8	4.8	93.5
Low Literacy Users	19	5	4.4	88.7

Figure 3: ROC Curve or Adaptive Efficiency Curve



6. User Impact and Ethical Considerations

6.1. User Experience: Personalization vs Predictability

Adaptive health questionnaires place people in a dilemma, where providing custom experiences leads to less predictability and yet remains important for trust. Real-time selection and adjustment of issues and replies boost how useful surveys seem to users and lower mental strain by offering only essential questions according to prior responses or information. Because of this adaptivity, users may feel more satisfied and could experience more accurate responses, especially when facing long tests. On the other hand, personalization leads to differences that may confuse users. Because adaptive designs may vary in structure and content, users might feel confused or unsure when the setup changes or when each section is not clearly explained. Because decisions can be unexpected or unexplained, this can confuse people, make them mistrust the system, or believe it weighs evidence with no clear explanation. For this issue to be addressed, adaptive systems should have user interface changes, helpful tips, and noticeable signs to indicate why the system asks each question. Finding the right point at which personalization and predictability meet is crucial; if there is too much effort to change, it could lead to trust issues, and if there is too little, the main idea behind the assessment could be lost. Hence, design methods such as participatory design and iterative testing of usability help calibrate this balance, considering the users' minds, feelings, and background.

6.2. Accessibility and Inclusiveness Across Demographics

Helping everyone participate easily, including those with special needs, should always be the main concern in developing health questionnaires. All users, regardless of their age, reading skills, comprehension of languages, disability, socioeconomic position, and cultural traditions, must be able to use these tools. To put it into action, this requires adaptive systems to feature multimodal methods, so people with all abilities and choices can use them more easily. The questions must be either culturally neutral or available in the local language to stop anyone from feeling left out. To ensure no unfair results occur for some demographics, item banks should be calibrated on different populations. When adaptive systems depend on digital tools, they must handle the digital divide, which often excludes those who cannot easily get to smartphones, the internet, or learn to use digital tools. Approaches to beat these inequalities could include having no internet need, easier interfaces, or connecting with health centers or community workers. It is also important to include inclusiveness in the algorithm, so that adaptive logic responds appropriately to different ways that symptoms may show up, depending on individual social situations. It's not enough to simply check accessibility; the process must include repeated user reviews, data about different groups of people, and real tests to ensure everyone benefits.

6.3. Bias and Fairness in Adaptive Question Logic

It is especially challenging to prevent algorithmic bias and to ensure fairness exists for all subgroups in an adaptive questionnaire system. Bias in adaptive logic can appear as items being selected differently, reduced predictive accuracy, or favoritism that routinely impacts users differently depending on race, gender, age, or similar protected attributes. The appearance of these biases may be caused by uneven item sets, data sets that do not fairly represent patients, or continuing biases in clinical standards. Because a calibration set uses white, middle-income individuals, an adaptive screener for anxiety might inaccurately diagnose distress in immigrant or minority groups, as their ways of showing distress are not always the same. Using an item selection algorithm that biases towards hard or stigmatizing questions can lead to violations of fairness and psychological safety for users identified by demographics. Therefore, fairness must be checked through comprehensive bias reviews of errors, item use, and different results. We can address these issues by using adversarial testing, fairness-aware machine learning, or post-hoc recalibration. If the logic behind picking items is clear and if the process is examined by outsiders, it adds strength and accountability to the program. Essentially, fairness in adaptive systems should be seen as a moral requirement to treat everyone justly, embrace inclusion, and respect users, making sure those historically left out of healthcare data are given equal treatment.

6.4. Psychological Impact of Real-Time Tailoring

Real-time tailoring in adaptive questionnaires strongly affects a person's psychology, mostly in sensitive fields such as mental health, trauma, chronic illness, and palliative care. Though adaptive systems aim to provide relevant content and raise engagement, some users might become uncomfortable, unsettled, or bothered if they believe the system is examining their weaknesses too closely. Sometimes, asking about someone's health suddenly shifts to asking about suicide may seem silly or discomforting if no background context is given. Sometimes, real-time changes can make users feel watched more closely, especially if the technology collects information from things like wearables. Users may be curious about the extent of the system's awareness of them, how their information is processed, and if their answers play any role in decisions made by machines. Adaptive systems need to use emotional intelligence in their design to manage and predict people's reactions. You need to match how difficult the questions are, give helpful feedback, and give users ways to skip or avoid stressful questions. Systems like providing immediate references, including crisis materials, and having healthcare monitoring in real time can help further shield users from damage to their mental health. In addition, systems should make it clear that changes are for support, not for judging or labelling students. By understanding and acting upon these dynamics, developers and clinicians help ensure adaptive assessments support user choice, safe feelings, and effective results rather than harm them.

6.5. Consent, Transparency, and Explainability

Making informed decisions and ensuring transparency in adaptive health questionnaires is important because many of the processes used are both intricate and difficult to explain. Because users are not given complete details about the questions they will see online, traditional consent is not always strong enough in digital environments. As a solution, adaptive systems should offer dynamic consent so that users know exactly what each answer means for the rest of the test, which data points are gathered, and why they are being used. It can be done by including reference to consent forms, organizing areas of the app according to needs, and remembering to alert users about their choices and rights. Understanding the process is necessary as well; users should be able to read clear information on the system, what algorithms are used, and the meaning of their results. It covers both sharing software source code or public studies on the project, and explaining the way governance, checks, and accountability are structured. Part of explainability in machine learning-based adaptivity is translating the steps of a complex algorithm so that a human can make sense of them. Through techniques such as LIME and SHAP, we can convert the output of a black-box algorithm into clear explanations for users. Governments and regulators in the EU, the U.S., and many other regions are making explainability a requirement for new AI governance laws. It supports users by allowing them to choose wisely, notice mistakes, and support their own interests. If consent, transparency, and explainability are not part of a system, even the best adaptation algorithm may breach ethics and lose the trust of users.

7. Challenges and Future Directions

7.1. Scalability Across Conditions and Populations

One main difficulty in creating and putting adaptive health questionnaires into practice relates to their ability to address a wide variety of health issues and populations. While current adaptive testing works well for depression, physical function, and chronic disease, it must be expanded with more effort and support focused on new studies, various cultures, and the expansion of its infrastructure. Symptoms and their effects vary from person to person because of differences in the way patients perceive and report health conditions. As an illustration, the symptoms of cognitive impairment are not the same in older adults as in young people with brain injuries, so the assessment needs special questions and a different approach. Symptom reporting is also affected by gender, culture, and economic background; therefore, health solutions need to be created and tested to include all demographics. In order to reach this breadth, we must train models on vast, well-labeled data, test them on comprehensive methods, and have domain experts take part in making the test questions. A system that can handle millions of users simultaneously, on many devices and networks, is crucial for adaptive technology. Thus, it is important to have cloud solutions, distributed systems, and smart caching in hospitals and high-traffic healthcare portals. Scalability also includes supporting different languages and different educational levels, since systems for low-resource places or the world must not depend solely on complicated text. This means scalability includes more challenges than just increasing computational capacity and needs care in areas such as specialists' participation, the technology's reliability, and offering inclusive features to all.

7.2. Dynamic Updates and Item Bank Maintenance

Since adaptive health questionnaires depend on how accurate, relevant, and robust their item banks are, updating and maintaining them is a difficult and necessary duty. Most static systems keep the same information for a long time, but adaptive systems have to change constantly as medical science, language, society, and user habits develop. Diagnostic approaches, cultural changes, and new advances in monitoring symptoms can cause old items to become obsolete. Ongoing monitoring of the data in the item bank is necessary to find items whose usefulness drops, takes too much time to complete, or becomes unfair to certain subgroups. Because of this, there should be a formal system for handling versions, archiving data, and refreshing the item response models. The fact that Dynamic Data makes changes easy also requires companies to ensure all stakeholders and users are aware and that every update is validated and clearly communicated to avoid issues with transparency and compliance. In such systems, suggestions for new items could be generated through analyzing users' reactions without supervision, but these must be fully reviewed and tested before being integrated. Maintaining an item bank becomes more difficult in these deployments because any updates have to reach users and be applied without causing confusion or a disruption in their data. When updates are semantically consistent due to compatibility with SNOMED CT or LOINC, they work better with other medical software. At the end of the day, the success of adaptive systems relies on continuing the curation of their question banks and involving clinicians, data scientists, ethicists, and patient representatives together in governance.

7.3. AI Explainability in Adaptive Models

Since adaptive health questionnaires now rely more on machine learning and AI to select and score items, many health experts agree that better transparency is necessary. You can measure and interpret the contributions of all test items because traditional IRT makes the process mathematically obvious. Yet, many modern adaptive engines use deep learning models and ensemble-based methods, as well as reinforcement learning systems. These developments often end up generating predictions that humans have a hard time explaining. As a result, introducing this technology into healthcare can be difficult due to legal compliance issues and a lack of trust among users in situations where diagnosis, treatment, and access to services are determined by such tools. LIME, SHAP, and counterfactual reasoning are among the techniques that help interpret how a model decides, but not all users can understand them, and experts are often needed to interpret them. Also, adaptive systems to combine both explainability and adaptivity, as extremely clear or fixed choices may end up being abused and misused, but unknown artificial effects might threaten the rights of users. Regulators such as the European Union and the US FDA are now enforcing that adaptive technology in health meaningfully shows transparency and manages risks through planned strategies. Consequently, developers need to focus on understandable models, give explanations that work for both patients and clinicians, and record how algorithms function. The next step in research should involve developing models that join the smartness of AI with the trackability of manually coded systems, making adaptive evaluations both responsive and straightforward.

7.4. Integration with National Health Systems

Adapting adaptive health questionnaires for use in national health systems can open new doors while presenting a range of difficult logistical, technical, and policy problems. Using adaptive assessments in these systems makes it much easier to look after patients using risk-based monitoring, to divide groups properly, and to target screening at a wide population level. The use of adaptive tools lowers the time patients are in the system, normalizes their answers, and helps guide quick triage when staff are swamped. Even so, national integration requires medical information to work well together, be standardized, and follow healthcare data rules. Adaptive systems must comply with HL7 FHIR, ICD-11, SNOMED CT, and other health IT standards, in addition to linking with current medical term lists to keep data meaningful and movable. Integration between systems needs to comply with all relevant privacy and security frameworks, such as GDPR, HIPAA, and local cybersecurity rules that can change a lot from one country to another in their scope and application. Problems arise when there are significant differences between city and country hospitals and when there are complicated systems for buying new equipment, making changes, and instructing staff. People and society must also be addressed about surveillance, algorithmic decisions, and the problem of unequal access to healthcare. Success in Estonia, Israel, and the NHS digital upgrades in the UK highlights that adaptive systems become key features of national health infrastructure when paired with important objectives. The field should give priority to flexible architectures and strong cooperation among healthcare businesses and government to bring modern innovation on par with systemic trustworthiness.

7.5. Use During Crises (e.g., Pandemics, Wars)

In times of pandemic, conflict, disasters, or migration, the use of adaptive health questionnaires is very important because the usual health care system can get overloaded. This situation requires fast, flexible, and relevant tools to assess health needs. Adaptive systems installed on mobile devices can support triage for sick individuals, check mental wellbeing, measure how well a patient is doing, or release relevant health information, while causing little physical contact and conserving resources. During the outbreak of COVID-19, countries used adaptive symptom screeners to help make sure testing was prioritized for those at greatest risk, self-isolation plans were made for infected patients, and medical assistance was available to those who needed it most. The systems reacted by updating their questions in real-time according to COVID-19 trends, new information from different regions, and changing categories for cases. Mobile-based assessments have also been used in countries like Ukraine or Syria to check for PTSD, look for signs of malnutrition, or measure chronic illnesses in people who are away from proper healthcare. Often, ethical and logistical concerns are more advanced during crisis deployments. If digital infrastructure is not always working well, literacy is not constant, and trust in institutions is low, it is crucial to provide service offline, support multiple languages, and design for different cultures. Because improper use of data might lead to persecution, stigma, or exploitation for people in these contexts, any data gathered should receive extra attention and be kept safe. Notably, adaptive logic needs to be frequently rechecked and updated as conditions on the ground, for example, new symptoms, changes in migration, or new public health directions, develop. With climate issues, political instability, and global health emergencies worsening, health questionnaires should be able to respond to crises and work as key parts of humanitarian aid and emergency response.

8. Conclusion

This significant move to adaptive health questionnaires is changing the way health assessment devices are made and used, leading to fresh methods for both collecting and using health information online. We have discussed the key theories, main technology, and practical impact of CAT on digital health. Using Item Response Theory, machine learning, and real-time routing, adaptive questionnaires provide more exact and efficient results than fixed-form questionnaires. Studies suggest that adaptive tools effectively reduce how much people struggle to respond, make assessment results more accurate, and increase engagement among a wide range of users. Also, when worn and linked with mobile health platforms, such systems track important signs of health in real time. Being able to adapt assessment approaches to a person's reactions supports both personalized medicine and broader goals within healthcare, for example, prompt detection, prevention, and equal access.

Along with their technology, adaptive health questionnaires have important effects on digital health progress. They demonstrate how assessment is moving from fixed models to flexible, adaptable systems that understand a person's health. Healthcare applications, online portals, and tools used for emergencies all use them because of their versatility and ability to work at scale. As a result, they help shift healthcare from fixing medical issues to preventing them using data. In the same way, such tools may spread high-quality health screening to underserved communities by being multilingual, operating seamlessly with little bandwidth, and being available through voice commands. When implementing such systems, important ethical, legal, and technical problems have been spotted that need to be addressed for their transformative potential to be used wisely. They show that working together closely among technologists, clinicians, ethicists, policymakers, and users is crucial for designing and directing adaptive systems.

To maintain accuracy and usefulness in adaptive health questionnaires, we need standard frameworks for their development, testing, and implementation. These frameworks must account for how the tool works for everyone, fairness checks in the algorithms, connections with current health databases, and following recent rules on AI and consent. Organizations and authorities responsible for global health should put effort into designing setup paths and ethical principles that mix new solutions with responsibility. Additionally, future studies should investigate ways to include adaptive assessments in parts of healthcare where they are less common, connect different types of data to improve overall health modeling, and analyze enduring outcomes both inside clinics and in communities. Examinations over time will show if adaptive tools help correct diagnosis, improve care, make health care fairer, and please users. All in all, adaptive health questionnaires are more than a tool: they are changing the way digital health functions by helping ensure healthcare is more personalized, fair, and effective.

References

1. Meijer, R. R., & Nering, M. L. (1999). Computerized adaptive testing: Overview and introduction. Applied psychological measurement, 23(3), 187-194. https://doi.org/10.1177/01466219922031310
2. Triantafillou, E., Georgiadou, E., & Economides, A. A. (2008). The design and evaluation of a computerized adaptive test on mobile devices. Computers & Education, 50(4), 1319-1330. https://doi.org/10.1016/j.compedu.2006.12.005
3. Magis, D., & Barrada, J. R. (2017). Computerized adaptive testing with R: Recent updates of the package catR. Journal of Statistical Software, 76, 1-19. https://doi.org/10.18637/jss.v076.c01
4. Devine, J., Fliege, H., Kocalevent, R., Mierke, A., Klapp, B. F., & Rose, M. (2016). Evaluation of Computerized Adaptive Tests (CATs) for longitudinal monitoring of depression, anxiety, and stress reactions. Journal of Affective Disorders, 190, 846-853. https://doi.org/10.1016/j.jad.2014.10.063
5. Liu, Q., Zhuang, Y., Bi, H., Huang, Z., Huang, W., Li, J., ... & Chen, E. (2024). Survey of computerized adaptive testing: A machine learning perspective. arXiv preprint arXiv:2404.00712. https://doi.org/10.48550/arXiv.2404.00712
6. Yu, X., Cheng, Y., & Chang, H. H. (2019). Recent developments in cognitive diagnostic

computerized adaptive testing (CD-CAT): A comprehensive review. Handbook of diagnostic classification models: Models and model extensions, applications, software packages, 307-331. https://doi.org/10.1007/978-3-030-05584-4_15

7. Vokinger, K. N., Nittas, V., Witt, C. M., Fabrikant, S. I., & Von Wyl, V. (2020). Digital health and the COVID-19 epidemic: an assessment framework for apps from an epidemiological and legal perspective. Swiss medical weekly, 150, w20282. https://doi.org/10.3929/ethz-b-000465761
8. Haverinen, J., Keränen, N., Falkenbach, P., Maijala, A., Kolehmainen, T., & Reponen, J. (2019). Digi-HTA: Health technology assessment framework for digital healthcare services. Finnish Journal of eHealth and eWelfare, 11(4), 326-341. https://doi.org/10.23996/fjhw.82538
9. Ienca, M., Schneble, C., Kressig, R. W., & Wangmo, T. (2021). Digital health interventions for healthy ageing: a qualitative user evaluation and ethical assessment. BMC geriatrics, 21, 1-10. https://doi.org/10.1186/s12877-021-02338-z
10. Marsch, L. A., Campbell, A., Campbell, C., Chen, C. H., Ertin, E., Ghitza, U., ... & Young, S. (2020). The application of digital health to the assessment and treatment of substance use disorders: The past, current, and future role of the National Drug Abuse Treatment Clinical Trials Network. Journal of substance abuse treatment, 112, 4-11. https://doi.org/10.1016/j.jsat.2020.02.005
11. Kanzler, C. M., Lessard, I., Gassert, R., Brais, B., Gagnon, C., & Lambercy, O. (2022). Reliability and validity of digital health metrics for assessing arm and hand impairments in an ataxic disorder. Annals of Clinical and Translational Neurology, 9(4), 432-443. https://doi.org/10.1002/acn3.51493
12. Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. Information Fusion, 96, 156-191. https://doi.org/10.1016/j.inffus.2023.03.008
13. Frances, M. Y., & Solon, T. K. (2014). Item response theory for measurement validity. Shanghai archives of Psychiatry, 26(3), 171. https://doi.org/10.3969/j.issn.1002-0829.2014.03.010
14. Downing, S. M. (2003). Item response theory: applications of modern test theory in medical education. Medical education, 37(8), 739-745. https://doi.org/10.1046/j.1365-2923.2003.01587.x
15. Reise, S. P., & Haviland, M. G. (2005). Item response theory and the measurement of clinical change. Journal of personality assessment, 84(3), 228-238. https://doi.org/10.1207/s15327752jpa8403_02
16. Embretson, S. E., & Reise, S. P. (2013). Item response theory for psychologists. Psychology Press. https://doi.org/10.4324/9781410605269
17. Nguyen, T. H., Han, H. R., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. The Patient-Patient-Centered Outcomes Research, 7, 23-35. https://doi.org/10.1007/s40271-013-0041-0
18. Becker, J., Schwartz, C., Saris-Baglama, R. N., Kosinski, M., & Bjorner, J. B. (2007). Using item response theory (IRT) for developing and evaluating the pain impact questionnaire (PIQ-6™). Pain Medicine, 8(suppl_3), S129-S144. https://doi.org/10.1111/j.1526-4637.2007.00377.x
19. Damaševičius, R., Jagatheesaperumal, S. K., Kandala, R. N., Hussain, S., Alizadehsani, R., & Gorriz, J. M. (2024). Deep learning for personalized health monitoring and prediction: A review. Computational Intelligence, 40(3), e12682. https://doi.org/10.1111/coin.12682
20. Venčkauskas, A., Štuikys, V., Toldinas, J., & Jusas, N. (2016). A model-driven framework to

develop personalized health monitoring. Symmetry, 8(7), 65. https://doi.org/10.3390/sym8070065

21. Esposito, M., Minutolo, A., Megna, R., Forastiere, M., Magliulo, M., & De Pietro, G. (2018). A smart mobile, self-configuring, context-aware architecture for personal health monitoring. Engineering Applications of Artificial Intelligence, 67, 136-156. https://doi.org/10.1016/j.engappai.2017.09.019
22. Clifton, L., Clifton, D. A., Pimentel, M. A., Watkinson, P. J., & Tarassenko, L. (2012). Gaussian processes for personalized e-health monitoring with wearable sensors. IEEE Transactions on Biomedical Engineering, 60(1), 193-197. https://doi.org/10.1109/TBME.2012.2208459
23. Almtireen, N., Altaha, H., Alissa, A., Ryalat, M., & Elmoaqet, H. (2024, September). Ai-driven mobile app for personalized health monitoring. In 2024 22nd international conference on research and education in mechatronics (rem) (pp. 338-342). IEEE. https://doi.org/10.1109/REM63063.2024.10735595
24. NVPS, R. K., Damaševicˇius, R., Jagatheesaperumal, S. K., Hussain, S., Alizadehsani, R., & Gorriz, J. (2023). Deep Learning for Personalized Health Monitoring and Prediction: A Review. Authorea Preprints. https://doi.org/10.22541/au.168723724.45330026/v1
25. Gerke, S., Minssen, T., & Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In Artificial intelligence in healthcare (pp. 295-336). Academic Press. https://doi.org/10.1016/B978-0-12-818438-7.00012-5
26. Schönberger, D. (2019). Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. International Journal of Law and Information Technology, 27(2), 171-203. https://doi.org/10.1093/ijlit/eaz004
27. Bekbolatova, M., Mayer, J., Ong, C. W., & Toma, M. (2024, January). Transformative potential of AI in healthcare: definitions, applications, and navigating the ethical landscape and public perspectives. In Healthcare (Vol. 12, No. 2, p. 125). MDPI. https://doi.org/10.3390/healthcare12020125
28. Goktas, P., & Grzybowski, A. (2025). Shaping the future of healthcare: Ethical clinical challenges and pathways to trustworthy AI. Journal of Clinical Medicine, 14(5), 1605. https://doi.org/10.3390/jcm14051605
29. Goirand, M., Austin, E., & Clay-Williams, R. (2021). Implementing ethics in healthcare AI-based applications: a scoping review. Science and Engineering Ethics, 27(5), 61. https://doi.org/10.1007/s11948-021-00336-3
30. Smallman, M. (2022). Multi scale ethics—why we need to consider the ethics of AI in Healthcare at different scales. Science and Engineering Ethics, 28(6), 63. https://doi.org/10.1007/s11948-022-00396-z