A Randomized clinical trial using artificial intelligence to shorten wait times and boost customer satisfaction in pediatric outpatient services

Rakesh Kumar Mishra and Geetanjali Sharma*

1.Research scholar,Mathematics,BanasthaliVidyapith,Rajasthan,India-304022.

2. Faculty of Mathematics ,Banasthali Vidyapith,Rajasthan,India-304022

E-mail:mishrarbu303@gmail.com ; geetanjali.bu@gmail.com*Corresponding author

Abstract: Outpatient operations that are complicated are accompanied by a lot of paperwork and drawn-out wait times. We want to reduce wait times and raise visitor satisfaction. We created a programmed called Smart-doctor that uses artificial intelligence (AI) to aid users. At MGH Medical Centre, a randomized controlled trial was carried out. Participants were split into two groups—one with AI assistance and the other without. In the AI-assisted group, As a medical assistant, Smart-Doctor was employed An electronic medical satisfaction survey was requested to be completed following the visit. The queuing time was the main outcome, and the consultation and testing periods, time spent total and satisfied rating were the secondary objectives. Additionally employed were the Wilcox on multiple linear regression, the rank sum test, and ordinal regression

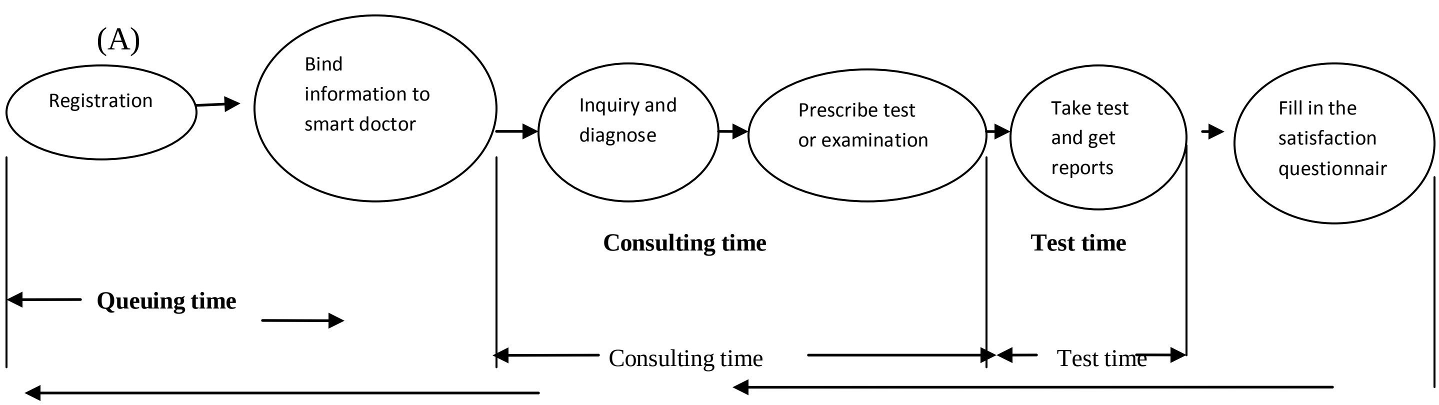
Introduction: The government's implementation of the two-child policy, which allows a couple to have up to three children, and India's population exceeding 1.4 billion people, have caused an increase of non-emergency visits to children's hospitals (1). The public has focused a great deal of attention on issues like long wait times, congested waiting rooms, and mistrust between patients and doctors (3, 4). These issues not only have a negative impact on patients' satisfaction with hospitals and physicians, but they also restrict the growth of the health and medical industry (5) The lengthy hospital wait times in India are a result of particular outpatient service protocols (6, 7), as the majority of patients can register themselves when they arrive at the hospital without an appointment (8). As a result, individuals must wait in queue multiple times to sign in, register and visit a doctor. The same is true for waiting to get the report, pay the bill, have an exam, and then waiting again to see their doctor. Particularly in major hospitals, there seem to be a lot of queuing segments. Additionally, as far as we are aware, the doctor's office doors have the longest lines (9). There is a large queue of patients waiting there, anxiously anticipating the announcement of their number. In our prior retrospective cohort study, we exposed the internal medicine department Introducing Smart-doctor, a medical assistant powered by AI. (19). Smart-doctor can treat multiple patients at once, model itself after how doctors reason and make decisions, and recommend the right tests and examinations for each patient. Patients can see a doctor without waiting in queue thanks to its assistance. All patients must use their smart phones to scan a special two-dimensional code before Smart-doctor can conduct an investigation. Patients still have to wait until clinical doctor reviews the test or examination results and makes a decision before moving on (Figure 1). The hospital information system was troubleshooted. due to the low quality of retrospective research in terms of data loss, logical mistake, and selection bias. Then, we conducted a randomized controlled experiment (RCT) to determine how Smart-doctor affected patients with respiratory and gastrointestinal disorders' wait times and satisfaction ratings.

Methods:

Study design and participants: The study was conducted at the 1,000-bed MGH Children's Medical Centre (JCMC), a tertiary hospital that is connected with the school of medicine at Mahatma gandhi hospital. This hospital introduced Smart-doctor in the internal medicine, respiratory medicine, and gastrointestinal departments in 2022. The number of annual outpatient visits was 90,000 in 2022 and 125,699 in 2023, respectively. 90.32% of these visits were handled by the traditional track, and 9.68% were handled by the AI-assisted track. An AI assistant called Smart-doctor is built using a machine learning approach that was discussed in both Liang's and our prior studies.

A model for natural language processing (NLP) that is driven by deep learning -based intelligent system called Smart-doctor suggests the tests and examinations that should be performed prior to consulting a physician, simulating the clinical reasoning process of a real doctor. From May 17 to July 30, 2023, participants were sought out in the department of internal medicine. Children with parental permission between the ages of 2 months and 18 were enrolled by uniformly trained investigators. During the visit, parents were present with every child. Children have to have a cough, diarrhoea, urinary pain, or vomiting as their main complaint to be eligible. Participants were excluded if their guardians refused to participate in the study or if their guardians found it impossible to support the study's continuation. The specifications were established based on the findings the initial test experiment: = 0.05, statistical power = 0.90, average queuing time of the AI-assisted group: 18.84 min, average queuing time of the conventional group: 25.38 min, standard deviation of the queuing times for the two groups: 19.44 min and 26.74 min.

(A)



(B)

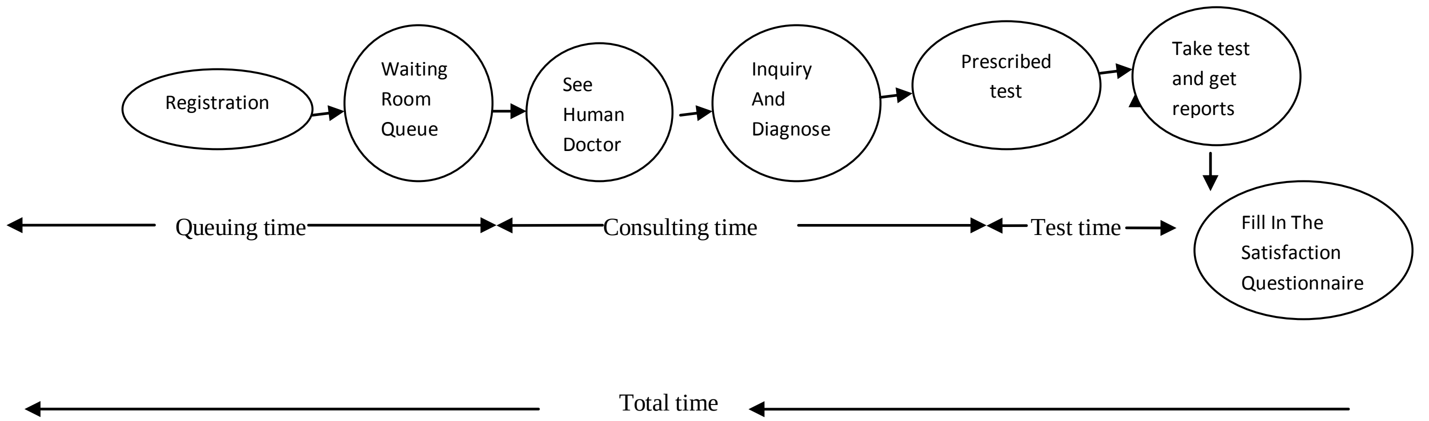


FIGURE 1

(Process of traditional group and AI-assisted group patient visits. (A) An outpatient visiting method powered by AI. Smart-doctor makes a diagnosis and also recommends tests and examinations for the patient. (B) The standard protocol for outpatient visits. A real human doctor makes a diagnosis and recommends tests or examinations for the patient.)

Intervention and grouping: Participants in the study were randomly assigned to one of two groups: those questioned by Smart-doctor (the AI-assisted group, the intervention), and those interviewed by real doctors (the traditional group, the control). According to the queue management system's guidelines (22), if a patient missed a round, they often had to wait for two further initial patients and one additional returning patient before receiving another round. After getting the results of the exams or tests, Both groups of patients were required to wait outside the doctor's office. It was necessary to report the outpatient number as a proposal violation (PV) during this time if the patient showed a strong desire to switch groups or withdraw. The investigator would alert the medical team if the patient's condition worsened so that the patient might receive prompt treatment. Investigators would provide the child's guardian instructions on how to complete the electronic satisfaction survey while they were waiting in queue.

Randomization and blinding setting: The allocation ratio between the two groups of eligible participants was 1:1. was used to create 800 random numbers, ranging from 0 to 100, with a fixed value . Patients were divided into two groups: those with odd numbers were placed in the traditional group, while those with even numbers were placed in the AI-assisted group. For this investigation, the blinding approach was inappropriate Despite the fact that interventions were chosen at random, the patients (or their guardians) had to carry out certain therapies. The majority of guardians had experience visiting a public hospital, and the outpatient procedure, which was identical to our standard practice, was typically similar in practically all public hospitals in India. Therefore, hiding the patients or their guardians based on the group they were assigned to was impractical.

Satisfaction investigating and data collecting: An electronic questionnaire was used to research satisfaction. This survey related to a longer survey from the study conducted in secondary and tertiary hospitals in India (23). The eight areas—waiting time, autonomy, continuity, efficiency, effectiveness, knowledge, information, and empathy—that have been determined to be most significant in terms of the patient's assessment of the quality of treatment were included in the comprehensive satisfaction questionnaire This questionnaire's validity and reliability were confirmed among Indians. The questionnaire was completed by the patients' guardians because all of the patients were minors. There were two areas in the satisfaction survey. The first part of the survey asked for general information, such as the outpatient identification number, the method of registration (online or in person), whether an appointment had been set, whether a pre-diagnosis had been made, whether a pre-inquiry had been made, and the principal complaint. The second segment examined how parents felt about visiting, including their registration, pre-inquiry, pre-test, waiting time, interactions with the doctor and other medical personnel, their experiences while visiting, and the necessity of implementing Smart doctor in the hospital. With five possible responses for each entry on the Likert scale—extremely satisfied (5 points), somewhat satisfied (4 points), average (3 points), unhappy (2 points), and very dissatisfied (1 point)—the questionnaire was graded. We gave guardians the opportunity to select the average (3 points) response when a patient did not receive one of these services. The Hospital Information System (HIS) was used to record information on outpatient procedures, including general information about children (such as age, gender, method of registration and treatment, and registration and treatment timeframes and chargeable time. In order to correlate patient recruitment data with the information in the HIS, scientists made note of the patient's visit time and visit number.

Evaluating performance of smart-doctor:

We did an evaluation study prior to this trial in order to demonstrate the efficacy of Smart doctor in prescribing tests and examinations and to increase the confidence of doctors and guardians in AI-assisted inquisition. Electronic Health Records (EHRs) provided us with information on patients who visited the internal medicine section between May 2023 and July 2023. Finally, patients with gastrointestinal and respiratory disorders were chosen. The department of general internal medicine saw the most patients with these two types of disorders, which is why they were chosen. Furthermore, these patients' conditions were generally moderate, making them better candidates for processing with AI assistance. The guidelines for diagnosis were the Guideline for the Diagnosis of Respiratory Diseases in Children and the Guideline for the Diagnosis of Gastrointestinal Diseases in Children. Three highly qualified doctors were invited to review the Smart-doctor-recommended things in accordance with the industry's best practices. The first and second specialists independently assessed the listed cases. The two outcomes would be adopted if they were comparable. In the event that they weren't consistent, the outcomes will be decided after consulting a third expert missed or unnecessary exams or examinations were included in the unqualified recommendation of tests or examinations.

Outcomes and statistical analysis: The main result was the queuing time, which was the amount of time it took from patient registration to their initial entrance into the doctor's office. Secondary outcomes included consulting time (doctor's time spent ,asking questions, palpating, and prescribing), testing time (patient's time spent receiving auxiliary exams or tests), and total time (patient's time from registration to completing satisfaction survey). Other results included comparing the levels of satisfaction between the two groups and the variables influencing satisfaction and queue wait times. The use of Smart-doctor during the pre-test could result in negative outcomes, such as a nimiety, error, or absence of tests or examinations. Additionally, by comparing the percentage of doctor or Smart-doctor prescriptions that were repeated as well as the cost between the AI-assisted and conventional groups, Indirect evaluation of detrimental impacts was done in this study. Incomplete test or assessment materials are referred to be "twice prescription" when they are given to the patient during the initial appointment. That is to say, there might be some items that are missing. This occurred because the doctor or AI either forgot to prescribe the tests or examinations items or because they did so based on the patient's original test results or examinations' report. The Shapiro-Wilk test was used to determine whether the data were normal. First, we evaluated the patient characteristics using the t-test or Chi-square test after describing participant characteristics using Mean SD (standard deviation) and proportions. Second, using the Wilcoxon rank sum test, we compared the waiting time, consulting time, testing time, and overall time The Wilcoxon Rank Sum Test was used to compare the data after we computed the average scores for each entry in the satisfaction questionnaire. Fourth, we used multiple linear regression (MLR) to examine the relationship between the factors (group, gender, arriving on time or missing the turn, etc.) and the queuing time. Fifth, we used ordinal regression to examine the variables (group, wait time, turning up early or late, etc.) that affect the satisfaction rating. Sixth, the percentage of repeat prescriptions and the cost were compared between the two groups using the Chi-square test and the Wilcoxon rank sum test. Significant statistical values were defined as p 0.05 and 1- > 0.80.

Results:

Evaluation of smart-doctor’s performance:

For the evaluation, the EHRs of 2,118 patients with gastrointestinal disorders and 7,725 patients with respiratory diseases were collected. The prescription would be deemed to be accurate in the opinion of human doctors if the testing/examination of a patient performed by Smart-doctor was in perfect accordance with standards. The gold standard for pediatric respiratory and gastrointestinal disorders diagnosis was a set of guidelines. In terms of respiratory and gastrointestinal disorders, respectively, the accuracy was 0.92 and 0.85.

Flow chart of the randomized controlled trial: A 540-person sample was thought to be sufficient. We adjusted the sample size by 10% to account for participant rejection, and the predicted final sample size was 594. 626 (84.59%) of the 740 patients enrolled during the trial period were eligible for randomization (Figure 2). However, 69 parents declined to take part due to temporary circumstances and unfamiliar cell phone functions. Due to the short application time and limited community reach of Smart-doctor, mistrust was another major factor in midway exits. 45 patients who had complex concurrent symptoms were also omitted because they did not match the inclusion criteria. 626 (100%) of those who were randomly assigned completed the patient satisfaction survey. Eleven patients who were supposed to be in the conventional group were moved to the AI group for breaking the proposal, and 15 more patients in the AI-assisted group were moved to the conventional group per their requests. In the end, 313 patients were in the AI-assisted group and 313 patients were in the control group, totaling 626 patients in the Full Analysis Set (FAS) and ITT analysis. Then, 610 participants were divided into the Per-Protocol Set (PPS) and an AT analysis was carried out (302 in the control group and 298 in the AI-assisted group).

Basic characteristics of participants: There is no significant difference is found in age between the AI assisted and conventional groups .

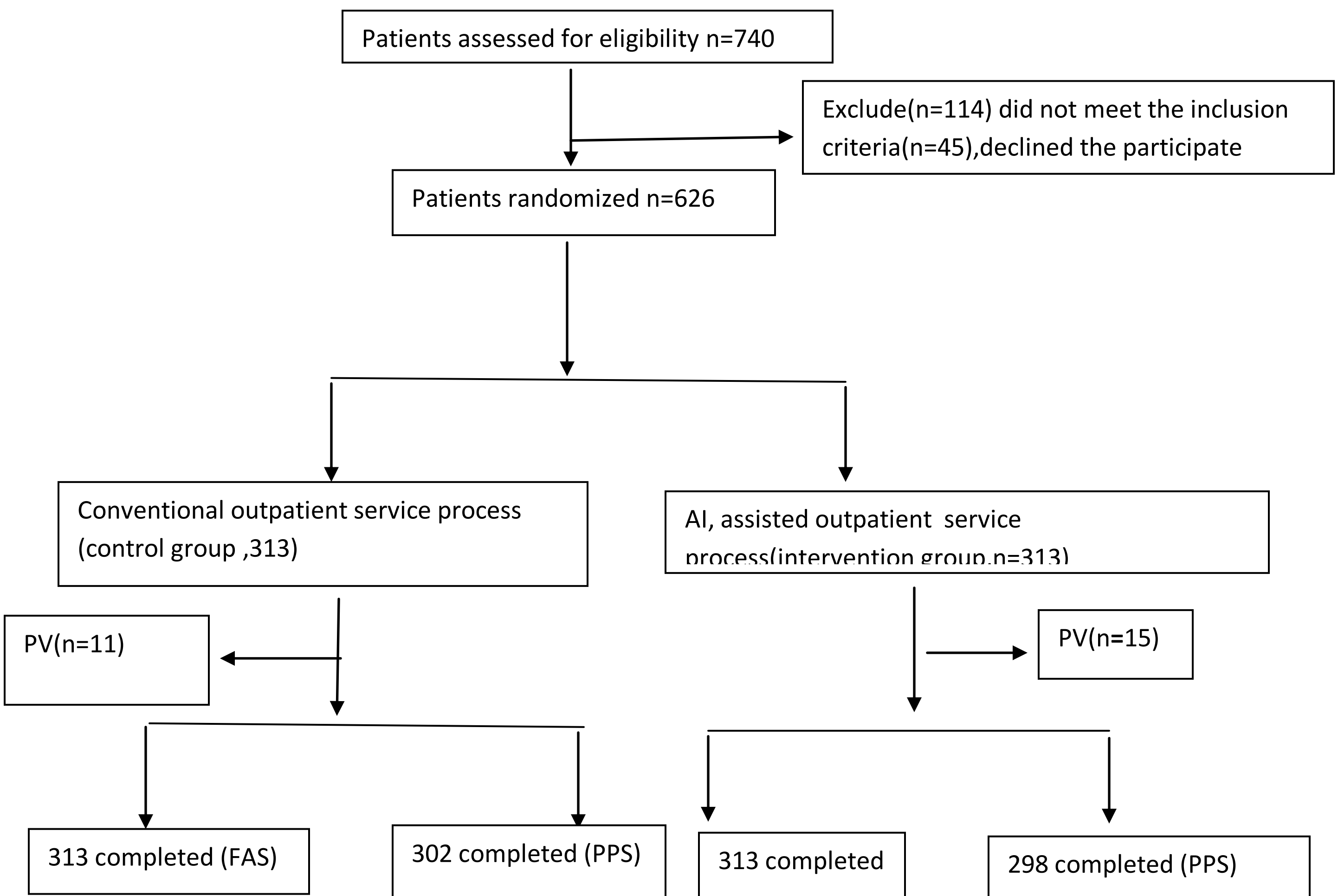


Table 1 shows the baseline characteristics of the participants. There were 159 (51.02%) boys and 168 (52.60%) boys in the two groups, respectively (p = 0.522). We firmly based our patient recruitment on their primary complaints, and we ultimately included 232 cases of diarrhea, 215 cases of cough, 146 cases of urine pain, and 33 cases of vomiting. Patients' primary complaints did not differ between groups, according to the Chi-square test ($p = 0.821$). The registration process, appointments, pre-inquiries, missing their turn, and weekdays were the primary areas where the control and intervention groups differed. More persons registered using mobile devices, scheduled appointments in advance, used the pre-inquiry service, missed their turn, and visited on weekdays during outpatient service in the AI-assisted group ($p <0.0$).

Time and cost comparison between the conventional group and the AI-assisted group:

The initial allocation plans were rejected by certain guardians during the intervention. As a result, three sets of data from ITT, PP, and AT were utilized. ITT analysis was used to compare the time-related results between the two groups in Table 2. At time outcomes, it was discovered that time variables were much Queuing time was decreased more in the AI-assisted group than in the control group. (7.386[3.97,31.70] vs. 20.72 [6.50,72.20] min, $p < 0.01$), consulting time (0.35 [0.19,0.89] vs. 2.582[1.60, 3.20] min, $p < 0.01$), and total time (40.20 [25.40, 72.70] vs. 110.40 [67.40, 150.30] min, $p < 0.01$Additionally, there was no difference in test/examination times between the conventional group and the AI-assisted group. (17.92 [10.10, 31.16] vs. 17.83 [13.19, 26.87] min, $p = 0.774$). Supplementary Tables 1, presented the respective results of PP and AT analysis, which were consistent with the results of ITT analysis (Table 2). The rate of twice-prescribed medications was the same in both groups, at 3.83% in the AI group and 6.07% in the traditional group. ($p = 0.269$). Moreover, patients in the AI-assisted group spent less on test fees (120.72 ± 110.66 vs. 134.16 ± 149.41, $p < 0.01$), examination fees (9.65 ± 38.43 vs. 27.25 ± 72.97, $p < 0.01$), and drug fees (25.15 ± 84.54 vs. 116.30 ± 128.40 $p < 0.01$) than those in the conventional group

Score of satisfaction between the traditional group and the AI-assisted group:

Figure 3 compares the satisfaction ratings between the conventional group and the AI-assisted group. The average score for each of the seven satisfaction rating questions in both groups was higher than 3 points. The registration, pre-inquiry, waiting time, service from a doctor or other medical staff, and overall satisfaction levels in the AI-assisted group were all higher than those in the control group. (p 0.01). The waiting time satisfaction score showed the biggest difference the AI-assisted group improved by 1 point over the conventional group ($Z = –8.020$, $p < 0.01$). The AI-assisted pre-test was not administered to conventional group patients before to consultation; hence the item's default score was set to 3. Consequently, we did not compare the two groups' pre-test results.

TABLE 1 Characteristics of participants.

Characteristics	AI-assisted group N = 313(%)	Conventional group N = 313(%)	Total N = 626	P
Gender				
Male	159 (51.02)	168 (52.60)	327	0.522[a]
Female	154 (48.98)	145 (47.40)	299	
Self-reported symptom				
Nausea and vomiting	19 (6.07)	14 (4.47)	33	0.821[a]
Cold and Cough	105 (33.55)	110 (35.14)	215	
Abdominal pain and diarrhea	115 (36.74)	117 (37.38)	232	
Frequent urination urgency	74 (23.64)	72 (23.00)	146	
Registration way				
Mobile phone	115 (36.74)	58 (18.53)	173	<0.01[a]
Machine & service window	198 (63.26)	255 (81.47)	453	
Appointment or not				
Yes	104 (33.23)	54 (17.25)	158	<0.01[a]
No	207 (66.13)	259 (82.45)	466	
Missing	2 (0.64)	0 (0.00)	2	
Pre-inquiry or not				
Yes	194 (61.98)	65 (20.77)	259	<0.01[a]
No	119 (38.02)	248 (79.23)	367	
Missing the turn or not				
Yes	30 (9.58)	90 (28.75)	120	<0.01[a]
No	283 (90.42)	223 (71.25)	506	
Weekdays or weekends				
Weekdays	267 (85.30)	210 (67.09)	477	<0.01[a]
Weekends	46 (14.70)	103 (32.91)	149	

AI, Artificial intelligence.
[a] Chi-square test.

Liner regression of queuing time: Results of the multiple linear regressions for factors affecting queuing time are presented in Supplementary Table 3. Group, gender, missing the turn or not, weekdays or weekends, appointment-free or not,

TABLE 2 ITT analysis of time between AI-assisted group and conventional group

Time Variables	AI-assisted group (N=313) Median (P_{25}, P75)	Conventional group (N=313) Median (P_{25}, P75)	P
Queuing time, min	8.78 (3.97, 33.88)	46.10 (17.53, 87.79)	<0.01[c]
Consulting time, min	0.35(0.18, 0.99)	2.68(1.82, 3.80)	<0.01[c]
Test time min	18.9 (11.10, 30.16)	17.93(13.19, 26.87)	0.874
Total time min	40.20(26.40, 73.80)	110.4(68.40, 164.40)	<0.01[c]

(Intention-to-treat or ITT. Artificial intelligence or AI. Interquartile range is IQR. A The time spent in queue between registering and visiting a doctor. B Time allotted for consultation: While the patient is in the doctor's office, the doctor takes the necessary time to speak with them, examine them, and write down their prescription. C Time spent performing lab tests or imaging examinations on a patient. D Total time: The period of time between the patient's admission and discharge from the hospital. Test of Wilcoxon rank sum.)

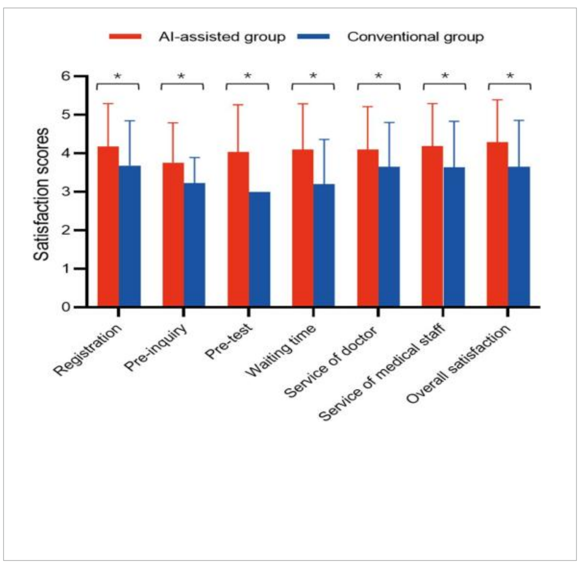


FIGURE 3 Scores of satisfaction for Guardian. Signifies a statistically significant difference. The pre-test/examination was not given to the conventional group, hence the default selection was "normal" (3 points). As a result, the scores of the traditional group and the AI-assisted group were not compared.

Discussion

This study was carried out at MGH Children's Medical Centre, a representative, specialized children's hospital that is connected to the university. This project is innovative in that it uses AI to arrange tests and examinations before doctors ask for them. There could be a number of determining elements when it comes to queuing time. Grouping and whether patients missed the turn had the biggest effects on multiple linear regression. The group of patients decided if there would be an AI assistant in the session and whether there would be a queue to see the doctor. The Smart-doctor helped reduce the wait time in the queue. Due to personal factors, the round missing caused the waiting time to extend. One-third of the patients experienced issues with missing their turn and having to wait longer in queue. Additionally, the wait time was quicker on the weekends than it was throughout the workweek. This might have happened as a result of fewer people going to the hospital on weekends. People in India are used to receiving outpatient care during the week without an appointment because few hospital employees in that country are required to work on weekends or holidays. Queuing times may have been affected by gender-related factors as well. Boys tended to wait in queue for a longer period of time, according to Supplementary Table 3's findings. However, this occurrence has no logical explanation. In Figure 2, we discovered that each item's satisfaction score in the AI-assisted group was much greater than that of the conventional group. According to our first hypotheses, the increase in satisfaction may have been caused by a reduction in patient wait times in the AI-assisted group as well as an improvement in guardians' satisfaction with medical professionals and other personnel, maybe as a result of more personal attention given to them in the AI-assisted group. They most likely requested staff assistance since they were using an AI-assisted process. In light of this, Supplementary Table 4, the model 1 of ordinal regression These findings might imply that cutting waiting times alone is unlikely to have an impact on patients' satisfaction with hospital services (24). Therefore, we believe there are a number of causes. First, it was impossible to blind patients to how they would be treated during the intervention's implementation, which could have resulted in bias (25, 26). Second, the AI-assisted group's patients lacked a clear understanding of how to use Smart-doctor. Here's where we needed the assistance of our investigators The traditional group's patients, though, were certainly familiar with the procedures, so it's possible that the researchers focused less on them. Therefore, similar to the Hawthorne Effect, a rise in patient satisfaction may have been the result of medical staff giving the patient more attention rather than simply taking less time (27, 28). Finally, the goal of our intervention was service completion; we did not tell patients of test results or reasons for treatment delays. Patient satisfaction has been linked to informational interventions that are more thorough (29, 30).

The accuracy of Smart-doctor's prior verification was 0.92 or so. We discovered that AI-assisted inquisition might behave effectively in outpatient service taking into account the fact that there was also a doctor to check the items to avoid omission and excessive. The verification served as the foundation for the widespread use of Smart-doctor to patients. This study used the rate of twice-prescribed medications and the cost of outpatient services to assess the negative consequences, which were defined as nimiety, mistake, and omission of tests and examinations. Actually, the findings indicated that both the AI-assisted group and the conventional group would experience a circumstance in which the doctor initially did not prescribe enough tests or examinations but would do so after patients received their reports and went back in to see the doctor. The two groups' differences did not, however, differ in a statistically significant way. The fact that the expenditures of the AI-assisted group were much lower than those of the conventional group shows that the patients did not undergo additional tests or examinations We really restricted the test items that Smart-doctor may recommend in our study to increase guardians' acceptance of it. This was demonstrated by the fact that Smart-doctor was only able to recommend straightforward, inexpensive, and minimally invasive procedures like basic blood tests and abdominal ultrasounds. Additionally, since all of the patients we included had similar illnesses, the necessities could be met with the minimal supplies. Based on these findings, Smart-doctor was nearly as effective as a licensed physician.

Patients were unfamiliar with the pre-diagnosis exam because Smart-doctor was only used for a brief period of time and other domestic institutions lacked a comparable application. at order to conduct the pre-diagnosis exam at the hospital during the trial, we recruited patients. In actuality, the pretest was not necessary for patients to take inside the hospital. The AI-assisted inquisition was able to learn about patients everywhere and at any time as the system was online and the medical card was bound. As a result, the appropriate prescription would be created, and patients may travel right to the hospital to have their tests or examinations. By transferring some of the stages that might be performed outside the hospital, we think AI-assisted inquisition can lessen the strain on hospitals. The disease's nosocomial transmission has also emerged as a key factor in the severity of COVID-19. The danger of infection may be increased by the long queue outside the clinic. However, the adoption of Smart-doctor can lessen the number of patients who must wait in the hospital for care, which can lessen nosocomial COVID-19 transmission during the epidemic. This provides other children's hospitals with a good model. The paediatrics department was the setting for our research, which focused on the most prevalent illnesses. We think that in today's highly specialized hospital, repetitive, arduous labor is what AI technology needs to do. For instance, patients with common colds and coughs frequently simply require a standard blood test before their doctor can recommend further drugs. However, due to the misdiagnosis of AI, individuals with complex and rare disorders may even pursue a more detouring path. Given the situation, we should be more cautious while utilizing AI. The current trial offered a number of advantages; earlier retrospective cohort studies had been carried out, giving this RCT a foundation and experience. In order to minimize selection bias during implementation, we precisely established inclusion and exclusion criteria and randomly allocated patients to interventions or control groups. The primary findings also matched those of the retrospective cohort study and what we anticipated. We also carried out a clinical registry prior to the study to direct the RCT. The fact that this study was carried out at a children's hospital and that the outpatient procedure was nearly exclusively guided by parents posed a significant drawback. As a result, in addition to the children, other variables that affected queuing time and satisfaction included the parents' occupation, income, and level of education. Unfortunately, because they were concerned about their children's illness, parents were not sufficiently cooperative when completing out the questionnaire. To make the form's completion by parents quicker, we simplified it. The subsequent multi-factor study was unable to investigate the impact of parental factors on the length of time spent in queue and level of satisfaction. The inclusion criteria for the study were a further source of its limitations. Patients' main symptoms were moderate and frequently occurring illnesses including cough, diarrhea, urine pain, or vomiting In actuality, these patients' primary complaints indicated that they had respiratory, urinary, or intestinal conditions. This was due to the fact that we restricted the conditions of patients included since the prescriptions that Smart-Doctor could write had a predetermined cap on their contents. We didn't include patients with complex symptoms, which could have influenced the results in a negative way (31). Beyond them, only based on evidence can intelligent systems resolve issues. With non-evidence-based systems or with unusual cases, doctors struggle.

Conclusion

The use of Inquisition with AI support in other institutions, particularly at India, where patients with common illnesses can register and consult a doctor from home, is highly beneficial. As a result, the resources are more likely to encounter challenging and complex patients, making it easier to use medical resources judiciously. We intend to introduce Smart-doctor to additional MGHMC departments soon. We will also roll out Inquisition with AI support to additional children's hospitals in Jaipur at the same time as we assess Smart-doctor's impact from the standpoint of a multi-center study.

Acknowledgment

We appreciate the participation and cooperation of all the kids and their parents in this study. We acknowledge the efforts of Dr Mohit mangol (a gastroenterologist), Harshit bhora (a respiratory surgeon), and Rahul gupta (an infectious diseases specialist) to the creation of the Smart-doctor system. We acknowledge the contributions of the gastroenterologists Atul chobey, Neha bashith, to assessment of the diagnostic effectiveness of the Smart-doctor system. We also acknowledge the reviewers of the manuscript, Xiumin Wang and Xiaodan Yu.

References

1. Abdus-Salam, R. A., Adeniyi, A. A., & Bello, F. A. (2021). Antenatal Clinic Waiting Time, Patient Satisfaction, and Preference for Staggered Appointment–A Cross-Sectional Study. Journal of Patient Experience, *8,* 23743735211060802.

2. Al-Harajin, R. S., Al-Subaie, S. A., & Elzubair, A. G. (2019). The association between waiting time and patient satisfaction in outpatient clinics: Findings from a

tertiary care hospital in Saudi Arabia. Journal of family & community medicine, *26*(1), 17

3. Almomani, I., & Alsarheed, A. (2016). Enhancing outpatient clinics management software by reducing patients' waiting time. Journal of infection and public health, 9(6), 734-743.

4. Bleich, S. N., Özaltin, E., & Murray, C. J. (2009). How does satisfaction with the health-care system relate to patient experience?. Bulletin of the World Health Organization, 87(4), 271-278.

5. Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety*, *28*(3), 231-237.
6. Dinh, M. M., Enright, N., Walker, A., Parameswaran, A., & Chu, M. (2013). Determinants of patient satisfaction in an Australian emergency department fast-track setting. *Emergency Medicine Journal*, *30*(10), 824-827.

7. Ehrler, F., Wu, D. T., Ducloux, P., & Blondon, K. (2021). A mobile application to support bedside nurse documentation and care: a time and motion study. JAMIA open, *4*(3), ooab046.

8. Harding, K. E., & Bottrell, J. (2016). Specific timely appointments for triage reduced waiting lists in an outpatient physiotherapy service. *Physiotherapy*, *102*(4), 345-350.

9. Harding, K. E., Snowdon, D. A., Prendergast, L., Lewis, A. K., Kent, B., Leggat, S. F., & Taylor, N. F. (2020). Sustainable waiting time reductions after introducing the STAT model for access and triage: 12-month follow up of a stepped wedge cluster randomized controlled trial. *BMC* Health Services Research, 20, 1-9.

10. Krier, M., Kaltenbach, T., McQuaid, K., & Soetikno, R. (2011). Potential use of telemedicine to provide outpatient care for inflammatory bowel disease. Official journal of the American College of Gastroenterology| ACG, 106(12), 2063-2067.

11. Li, X., Tian, D., Li, W., Dong, B., Wang, H., Yuan, J., ... & Liu, S. (2021). Artificial intelligence-assisted reduction in patients' waiting time for outpatient process: a retrospective cohort study. BMC health services research, *21*, 1-11.
12. Liang, H., Tsui, B. Y., Ni, H., Valentim, C. C., Baxter, S. L., Liu, G., ... & Xia, H. (2019). Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nature medicine, *25*(3), 433-438.

13. McCarthy, M. L., Ding, R., Zeger, S. L., Agada, N. O., Bessman, S. C., Chiang, W., ... & Bessman, E. S. (2011). A randomized controlled trial of the effect of service delivery information on patient satisfaction in an emergency department fast track. Academic Emergency Medicine, *18*(7), 674-685.

14. Melms, L., Schaefer, J. R., Jerrentrup, A., & Mueller, T. (2021). A pilot study of patient satisfaction with a self-completed tablet-based digital questionnaire for collecting the patient's medical history in an emergency department. *BMC* Health Services Research, 21(1), 1-13.

15. Merkow, R. P., Kaji, A. H., & Itani, K. M. (2021). The CONSORT framework. JAMA surgery, *156*(9), 877-878

16. Monaghan, T. F., Agudelo, C. W., Rahman, S. N., Wein, A. J., Lazar, J. M., Everaert, K., & Dmochowski, R. R. (2021). Blinding in clinical trials: seeing the big picture. *Medicina*, *57*(7), 647.

17. Pabinger, C., Lothaller, H., Lorenz, A., & Dammerer, D. (2021). Telemedicine versus on-site treatment at a surgical university clinic: Study of 225 consecutive patients. International Journal of Medical Informatics, 151, 104460.

18. Ren, W., Sun, L., Tarimo, C. S., Li, Q., & Wu, J. (2021). The situation and influencing factors of outpatient satisfaction in large hospitals: Evidence from Henan province, China. BMC health services research, *21*(1), 500.

19. Rezk, F., Stenmarker, M., Acosta, S., Johansson, K., Bengnér, M., Åstrand, H., & Andersson, A. C. (2021). Healthcare professionals' experiences of being observed regarding hygiene routines: the Hawthorne effect in vascular surgery. BMC Infectious Diseases, 21(1), 1-10.

20. Rohleder, T. R., Lewkonia, P., Bischak, D. P., Duffy, P., & Hendijani, R. (2011). Using simulation modeling to improve patient flow at an outpatient orthopedic clinic. Health care management science, 14, 135-145.

21. Sharma, K. A., Yadav, A., Sridhar, C., Malhotra, N., Biji, S., & Kumari, K. (2020). Reducing the waiting time to initiation of infertility treatment at a tertiary care centre in India. BMJ Open Quality, 9(4), e000975.

22. Sial, T. A., Liu, J., Zhao, Y., Khan, M. N., Lan, Z., Zhang, J., ... & Rajpar, I. (2019). Co-application of milk tea waste and NPK fertilizers to improve sandy soil biochemical properties and wheat growth. Molecules, *24*(3), 423.

23. Tatum, M. (2021). China's three-child policy. *The Lancet*, *397*(10291), 2238.

24. Wang, X., Jiang, R., Li, J., & Chen, J. (2018). Burströ m B, & Burström K. What do patients care most about in China's public hospitals? Interviews with patients in Jiangsu province. BMC Health Services Research, 18(1), 97.

25. Winsett, R. P. (2021). Editor View on Why Authors Need to Consider Blinding Within a Manuscript. Progress in Transplantation, *31*(3), 199-200.

26. Wu, D., Cui, W., Wang, X., Huo, Y., Yu, G., & Chen, J. (2021). Improvement in outpatient services using the WeChat calling system in the Shanghai Children's Hospital. Pakistan Journal of Medical Sciences, *37*(4), 993.

27. Xie, W., Yang, X., Cao, X., & Liu, P. (2019). Effects of a comprehensive reservation service for non-emergency registration on appointment registration rate, patient waiting time, patient satisfaction and outpatient volume in a tertiary hospital in China. BMC health services research, 19, 1-7.

28. Yan, J., Yao, J., & Zhao, D. (2021). Patient satisfaction with outpatient care in China: a comparison of public secondary and tertiary hospitals. International Journal for Quality in Health Care, 33(1), mzab003.