Optimizing the post-cataract patient journey using artificial intelligence-driven teleconsultation: a prospective case study.

Joukje C Wanten, Noël JC Bauer, Mohita Chowdhury, Aisling Higham, Nick de Pennington, Frank JHM van den Biggelaar, Rudy MMA Nuijts

Submitted to: JMIR Formative Research
on: February 24, 2025

Table of Contents

Optimizing the post-cataract patient journey using artificial intelligence-driven teleconsultation: a prospective case study.

Joukje C Wanten[1] MD; Noël JC Bauer[1] MD, PhD; Mohita Chowdhury[2]; Aisling Higham[2] MBBS, FRCOphth, MSc; Nick de Pennington[2] BM, BCh, MA, MBA, FRCS; Frank JHM van den Biggelaar[1] PhD; Rudy MMA Nuijts[1] MD, PhD

[1] University Eye Clinic Maastricht, Maastricht University Medical Centre+ Maastricht NL
[2] Ufonia Limited Oxford GB

Corresponding Author:
Joukje C Wanten MD

University Eye Clinic Maastricht, Maastricht University Medical Centre+
P. Debyelaan 25
Maastricht
NL

Abstract

Background: Given the increasing global demand for ophthalmologic care and the anticipated shortage of ophthalmology professionals, innovative solutions are essential for optimizing healthcare delivery. Digital health technologies offer promising opportunities to efficiently manage high patient volumes. Cataract surgery, with its established safety profile and routine postoperative care, is provides an ideal setting for these innovations. Structured clinical questions can effectively identify patients requiring further assessment, enabling clinicians to screen for complications through telephone consultations. Moreover, an artificial intelligence-based follow-up system could take this a step further by automating the process, reducing the need for clinician involvement while enabling more efficient postoperative screening for complications.

Objective: To assess the clinical safety and effectiveness of an artificial intelligence-based follow-up call system (Dora-NL1) in identifying patients who require further assessment after cataract surgery in the Netherlands.

Methods: Patients who underwent uncomplicated cataract surgery were eligible to participate. All patients received a Dora-NL1 follow-up call at 1 and 4 weeks postoperatively, in addition to routine postoperative cataract care. Dora-NL1 evaluated postoperative symptoms and made suggestions for care management decisions. The Dora-NL1 outcomes were compared with clinician assessments of recorded Dora-NL1 calls and regular care. User-acceptability was evaluated using the Telehealth Usability Questionnaire (TUQ).

Results: A total of 105 patients were included in the analysis. Dora-NL1 demonstrated high accuracy compared to clinician-supervised calls, with symptom evaluation accuracy of 89% to 99% and care management decision accuracy of 83% to 88%. At week 1, the sensitivity and specificity compared to standard telephone consultations were 100% and 42%, with no clinical concerns missed. However, at post-operative week 4, compared to an in-person hospital visit, Dora-NL1 did not identify unexpected management changes in 4 patients (4.1%). All 4 of these patients had issues that were only detected after slit lamp examination, reflecting the system's reliance on patient reported symptoms. Patients rated Dora-NL1 positively, with a mean TUQ score of 3/5, highlighting its simplicity, ease of use, and audibility.

Conclusions: Dora-NL1 is a safe and effective screening tool for postoperative cataract surgery, offering a safe alternative to telephone consultations, but this cannot fully replace in-person examinations. Clinical Trial: Not applicable.

(JMIR Preprints 24/02/2025:72574)
DOI: https://doi.org/10.2196/preprints.72574

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http
No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Optimizing the post-cataract patient journey using artificial intelligence-driven teleconsultation: a prospective case study.

Abstract

Background: Given the increasing global demand for ophthalmologic care and the anticipated shortage of ophthalmology professionals, innovative solutions are essential for optimizing healthcare delivery. Digital health technologies offer promising opportunities to efficiently manage high patient volumes. Cataract surgery, with its established safety profile and routine postoperative care, provides an ideal setting for these innovations. Structured clinical questions can effectively identify patients requiring further assessment, enabling clinicians to screen for complications through telephone consultations. Moreover, an artificial intelligence-based follow-up system could take this a step further by automating the process, reducing the need for clinician involvement while enabling more efficient postoperative screening for complications.

Objective: To assess the clinical safety and effectiveness of an artificial intelligence-based follow-up call system (Dora-NL1) in identifying patients who require further assessment after cataract surgery in the Netherlands.

Methods: Patients who underwent uncomplicated cataract surgery were eligible to participate. All patients received a Dora-NL1 follow-up call at 1 and 4 weeks postoperatively, in addition to routine postoperative cataract care. Dora-NL1 evaluated postoperative symptoms and made suggestions for care management decisions. The Dora-NL1 outcomes were compared with clinician assessments of recorded Dora-NL1 calls and regular care. User-acceptability was evaluated using the Telehealth Usability Questionnaire (TUQ).

Results: A total of 105 patients were included in the analysis. Dora-NL1 demonstrated high accuracy compared to clinician-supervised calls, with symptom evaluation accuracy of 89% to 99% and care management decision accuracy of 83% to 88%. At week 1, the sensitivity and specificity compared to standard telephone consultations were 100% and 42%, with no clinical concerns missed. However, at post-operative week 4, compared to an in-person hospital visit, Dora-NL1 did not identify unexpected management changes in 4 patients (4.1%). All 4 of these patients had issues that were only detected after slit lamp examination, reflecting the system's reliance on patient reported symptoms. Patients rated Dora-NL1 positively, with a mean TUQ score of 3/5, highlighting its simplicity, ease of use, and audibility.

Conclusions: Dora-NL1 is a safe and effective screening tool for postoperative cataract surgery, offering a safe alternative to telephone consultations, but this cannot fully replace in-person examinations.

Keywords: Artificial Intelligence; Cataract; Telemedicine; Digital Health; Clinical Study.

Introduction

Our healthcare system faces many challenges worldwide, including demographic shifts that contribute to an increasing mismatch between demand and supply. It is expected that, by 2035, the ophthalmology workforce will be insufficient to meet the growing demand for care.[1-3] Efficient management of high-volume workloads in ophthalmology is essential for maintaining healthcare quality and manageability, with eHealth offering promising solutions to improve accessibility in this field.[4-6] Cataract surgery, with its high safety profile and usually uneventful routine postoperative care, has the potential to be a target of application for these eHealth solutions.[7, 8] It is already proven that, in routine cataract cases, a structured set of clinical questions effectively identifies patients who require further assessments one week after surgery, and that replacing the first-week postoperative clinic visit with a nurse-led telephone call is safe and feasible.[9, 10] The voice-based automated telephone conversational assistant, Dora (Ufonia Limited, Oxford, United Kingdom), uses a machine learning conversational model to conduct follow-up after cataract surgery. This system applies standardized clinical questions to screen for complications in postoperative cataract surgery patients, potentially reducing the need for in-person follow-ups, and has already been validated for use in the United Kingdom (UK).[11] In this study, our aim is to validate and evaluate the clinical safety and effectiveness of the first Dutch version (Dora-NL1).

Methods

Recruitment

Adult patients (≥21 years) who underwent uncomplicated cataract surgery on either one or both eyes, with implantation of a monofocal (including toric) intraocular lens (IOL) targeted for emmetropia, were recruited for this study at the University Eye Clinic of the Maastricht University Medical Center. Patients were excluded if they had serious ocular comorbidities requiring face-to-face postoperative visits, underwent combined procedures (e.g. glaucoma or retinal surgeries), needed follow-up visits within one week due to complications, were not able to speak Dutch, or had significant hearing or cognitive impairments. All patients were informed about the study and signed an informed consent before enrolment. This prospective validation study was approved by the local medical ethics committee (Maastricht, the Netherlands) and executed in accordance with the Declaration of Helsinki.

Dutch model development and validation

The model for the autonomous cataract follow-up call system, Dora-NL1, was originally developed for English-speaking patients. In this study, the model was adapted for use in Dutch, with the initial Dutch system developed from the pre-trained English version. To facilitate this adaption, an intermediary translation function was integrated into the Dora-NL1 software, translating captured Dutch answers into English, formulating a response, then converting it back to Dutch for speech output. While this is a crucial step, the translation process introduces complexity and the risk of altering the original message. The Dora model has a finite number of potential utterances, and thus it was possible to confirm these were correctly translated into Dutch. To ensure the reliability of the Dora-NL1 software we ran a preliminary investigation before starting the study, by looking at the agreement between Dutch (Dora-NL1) and English (Dora-R2) calls. Twelve different postoperative scenarios were created, reflecting both the common and most severe outcomes of cataract surgery patients. These scenarios were simulated by a bilingual clinician researcher in both languages using a predefined strict script that underwent forward and back-translation and independent expert review to ensure accuracy and consistency. The responses from both models were compared to assess their consistency, defined as acceptable when no statistically significant differences were found in symptom assessment and care management decisions. The questions and responses used for model validation were similar to those used in the trial version of the Dora-NL1.

Postoperative follow-up

Patients undergoing bilateral cataract surgery underwent either immediate sequential bilateral

(ISBCS) or delayed sequential bilateral cataract surgery (DSBCS). For those undergoing unilateral surgery, the postoperative follow-up schedule aligned with that of ISBCS. All patients received regular postoperative care in addition to the Dora-NL1 calls, as shown in (Figure 1). One week after the first eye surgery, patients received both a Dora-NL1 call and a regular postoperative telephone consultation with a clinician. During the telephone consultations, a standardized set of questions was used to evaluate the patient's symptoms and identify any necessary management changes. Management changes included deviations from the standard eye drop tapering schedule, the addition of any drops (excluding artificial tears), the need for an additional hospital visit, or the implementation of further interventions (excluding suture removal). Clinicians conducting the telephone consultations were masked to the results of the Dora-NL1 calls. For patients receiving a toric IOL, a one-week follow-up hospital visit was scheduled instead of a telephone consultation. At 4 weeks postoperatively, another Dora-NL1 call was conducted within 24 hours before the regular outpatient clinic visit. The final postoperative visit at the outpatient clinic consisted of a routine ophthalmological assessment, including evaluation of the refractive status and visual acuity as well as slit lamp and retinal examination.

Dora-NL1 call

The postoperative Dora-NL1 call had the following conversational flow: greeting and introduction, patient identification, symptom evaluation, patient's queries, care management decision-making, user-acceptability question, and call closure. The key symptoms which were assessed, included redness, pain, reduced vision, flashing lights, and floaters. Dora-NL1 indicated for each symptom whether it was absent, present with or without clinical significance, or whether there was insufficient information. Dora is programmed to ask the patient to repeat once if the system did not have high confidence in understanding the first response. Based on the symptom evaluation, Dora-NL1 provided an overall outcome: 'no clinical concerns' or 'potential clinical concerns'. If there was 'incomplete information' about a symptom question, or if the patient asked a question that Dora-NL1 could not address, the outcome defaulted to 'potential clinical concerns'. The decision-making flowchart can be found in (Appendix 1). Depending on the overall call outcome, Dora-NL1 made care management decisions including 'no review recommended' or 'clinician review recommended'. All Dora-NL1 calls were recorded and subsequently reviewed by a clinician, who independently assessed symptoms and made clinical decisions based on the recordings, while masked to Dora-NL1's decisions and the regular care outcomes.

User-acceptability

At the end of each call, Dora asked patients to rate their likelihood of recommending the automated system to others on a scale of 1 to 10. The Net Promoter Score (NPS) was calculated by subtracting the '%-detractors' (scores of 1-6) from the '%-promoters' (scores of 9 or 10).[12] Additionally, after the final postoperative visit, patients were invited to complete the Telehealth Usability Questionnaire (TUQ), translated to Dutch. The TUQ consists of 20 questions that can be answered on a 5-point scale ranging from 'fully agree' to 'fully disagree'.[13]

Sample size and statistical Analysis

For this pilot-study a sample size of 100 participants who completed the 1-week call, was used. Data analysis was conducted using SPSS (IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY) and Excel (Microsoft Corp., Redmond, WA). Qualitative variables were summarized as frequencies and percentages, while descriptive statistics, including mean and SD, were calculated for quantitative variables. The agreement between the English and Dutch model in the developmental phase was tested using Chi-squared or Fisher's exact tests. The inter-observer reliability was calculated using the Cohen's Kappa coefficient (κ). The accuracy of the tool was defined in accuracy, specificity, and sensitivity, including 95% Confidence Intervals (CI) of these proportions calculated using the Wilson score interval. A significance level of ≤0.05 was applied.

Results

Preliminary study results

The agreement between the Dutch (Dora-NL1) and English (Dora-R2) versions used twelve scenarios. The accuracy in terms of detecting symptoms, including redness, pain, floaters, and flashing lights, was 100%. With respect to reduced vision, the accuracy was 83.3%, affected by two missing data points where Dora was unable to collect or understand the patient's response (one from the English model and one from the Dutch model. The accuracy of the care management decisions was 91.7%, with the decrease attributed to these missing data points. There was no statistically significant difference found between the two language models (P=.318). Detailed responses for each scenario are provided in (Appendix 2).

Study results

Between July 2023 and April 2024, 182 patients who were scheduled for cataract surgery were screened for eligibility. Of these, 168 were suitable for participation, and 111 patients provided informed consent. In total, 105 patients (190 eyes) with a mean age of 72±7 years (range 50-84) were included in the analysis; 100 patients completed the 1-week follow-up with Dora-NL1, and 98 patients completed the 4-weeks follow-up. The patient flowchart can be found in (Figure 2). The baseline characteristics are summarized in (Table 1).

Postoperative week 1

The agreement between the assessments made by Dora-NL1 at week 1 and those made by the clinician reviewing the Dora-NL1 call is shown in (Table 2). The sensitivity ranged from 66.7% to 100% and specificity ranged from 78.6% to 97.4% in identifying key symptoms and making care management decisions, with kappa values ranging from 0.390 to 0.947. Only the detection of new floaters at week 1 showed a sensitivity of 66.67%, and a specificity of 96.81%. Sensitivity was reduced as Dora misunderstood 2 patients that had reported floaters – the patients initially said they had floaters which Dora misunderstood, as per Dora's algorithm to re-ask the question if it did not understand, when asked again the patient answered that they did not have floaters. The agreement between the automated care management decisions made by Dora-NL1 and the regular care is shown in (Table 3). At week 1, only patients who received regular telephone consultations were analyzed, excluding those with toric IOLs (n=13) who were scheduled for routine postoperative hospital visits. At this time, Dora-NL1 demonstrated 100% sensitivity with no false negatives and a specificity of 41.7%, yielding a false positive rate of 57.5%.

Postoperative week 4

(Table 4) summarizes the agreement between the Dora-NL1 assessments at week 4 and those by the clinician reviewing the Dora-NL1 call, with a 100% sensitivity and specificity ranging from 88.3% to 99.0% for symptom evaluation. Despite this agreement, three asymptomatic patients who were identified as having 'no clinical concerns' required an unexpected management change at the 4-week postoperative clinic visit, and one patient who had been asymptomatic of visual concerns at the week 1 Dora and clinician calls reported binocular diplopia. The three asymptomatic patients all had inflammatory cells detected during slit lamp examination; with one of these patients also having a residual lens fragment in the anterior chamber. They received topical corticosteroids and required additional follow-up visits. The fourth patient who reported binocular diplopia was evaluated by an orthoptist, who identified monocular diplopia in the contralateral eye due to postoperative changes after macular pucker peeling.

Acceptability and patient experiences

Patients rated their likelihood of recommending the Dora-NL1 calls, with mean scores of 7.59±1.77 at week 1 and 7.38±2.35 at week 4, shown in (Figure 3). The corresponding NPS for these calls were +13.5 and +12.6, respectively. The TUQ (n=98) had an overall mean score of 3.19±1.13. High scores were given for aspects such as time savings due to reduced travel (3.80±1.11), simplicity (3.89±0.96), ease of use (3.94±0.88), and audibility (3.97±0.91). However, the user experience of Dora-NL1 was not deemed equivalent to an in-person consultation with a clinician, with a mean score of 2.46±1.07. Detailed questionnaire results are available in (Appendix 3).

Tables

Table 1. Baseline characteristics. [a]

Demographics (n=111)	
Gender (F/M), n	61/44
Surgical risk factors, n	
- High myopia	3
- High hyperopia	3
- Glaucoma	2
- Corneal pathology	7
- Diabetic retinopathy	1
- Previous vitrectomy	3
Type of surgery, n	
- Unilateral	20
- DSBCS	29
- ISBCS	56
IOL characteristics (n=190)	
Type of IOL, n	
- Monofocal non-toric	169
- Monofocal toric	21
IOL Power (D), Mean±SD	20.0 ± 3.8
Refractive target (D), Mean±SD	-0.03 ± 0.12

[a]Abbreviations: D (Diopter), IOL (Intraocular Lens), SD (Standard Deviation).

Table 2. Symptom and outcome accuracy automated Dora vs. supervised Dora call week 1. [a-c]

Decision	Accuracy (%)	Sensitivity (%)	Specificity (%)	Kappa [95%CI]	P-Value

(n=100)	[95%CI]	[95%CI]	[95%CI]		
Redness[1]	96.0 [90.12; 98.4]	100 [56.6; 100.0]	95.79 [89.7; 98.4]	0.695 [0.416; 0.973]	<.001
Pain[2]	97.0 [91.6; 99.0]	100 [20.7; 100.0]	96.97 [91.5; 99.0]	0.390 [-0.149; 0.929]	<.001
Vision issue[3]	94.0 [87.5; 97.2]	100 [70.1; 100.0]	93.41 [86.4; 96.9]	0.718 [0.508; 0.928]	<.001
New floaters	95.0 [88.8; 97.9]	66.7 [30.0; 90.3]	96.81 [91.0; 98.9]	0.589 [0.262; 0.916]	<.001
Flashing lights	98.0 [93.0;99.5]	100 [86.2; 100.0]	97.37 [90.9; 99.3]	0.947 [0.874; 1.000]	<.001
Outcome	88.0 [80.2; 93.0]	100 [92.0; 100.0]	78.57 [66.2; 87.3]	0.763 [0.641; 0.885]	<.001

[a]Amount of cases with ‘incomplete information’: [1] n=4; [2] n=1; [3] n=1.
[b]Accuracy = (TP+TN)/(TP+TN+FP+FN)
[c]Abbreviations: FN (False Negatives), FP (False Positives), TN (True Negatives), TP (True Positives)

Table 3. Accuracy clinical decision-making automated Dora vs regular telephone consultation week 1. [a]

	Decision regular telephone consultation		
Dora decision	Potential concerns n, (%)	No clinical concerns n, (%)	Total n, (%)
Potential concerns n, (%)	1 (1.1)	50 (57.5)	51 58.6)
No clinical concerns n, (%)	0 (0.0)	36 (41.4)	36 (41.4)
Total n, (%)	1 (1.1)	86 (98.8)	87 (100)

[a]Sensitivity: 100% [20.7; 100.0], Specificity: 41.9% [32.0;52.4], Kappa: 0.016 [0.00;0.398] P=.398.

Table 4. Symptom and outcome accuracy automated Dora vs. supervised Dora call week 4.

Decision (n=98)	Accuracy (%) [95%CI]	Sensitivity (%) [95%CI]	Specificity (%) [95%CI]	Kappa [95%CI]	P-Value
Redness[1]	99.0 [94.4;99.8]	100 [34.2; 100.0]	99.0 [94.3; 99.8]	0.795 [0.403; 1.000]	<.001
Pain[2]	98.0 [92.9; 99.4]	100 [34.2; 100.0]	97.9 [92.7;99.4]	0.657 [0.216; 1.000]	<.001
Vision issue	96.9 [91.4; 99.0]	100 [56.6; 100.0]	96.8 [90.9; 98.9]	0.754 [0.487; 1.000]	<.001
New floaters[3]	88.8 [81.0; 93.6]	100 [51.0; 100.0]	88.3 [80.3; 93.3]	0.823 [0.656; 0.990]	<.001
Flashing lights	94.9 [88.6; 97.8]	100 [77.2; 100.0]	94.1 [87.0; 97.5]	0.809 [0.648; 0.970]	<.001
Outcome	82.7 [74.0; 88.9]	93.6 [79.3;100.0]	77.6 [66.3; 85.9]	0.640 [0.491; 0.789]	<.001

[a]Amount of cases with ‘incomplete information’: [1] n=1; [2] n=3; [3] n=2.
[b]Accuracy = (TP+TN)/(TP+TN+FP+FN)
[c]Abbreviations: FN (False Negatives), FP (False Positives), TN (True Negatives), TP (True Positives)

Discussion

Principal Results

The purpose of this study was to evaluate the clinical safety and effectiveness of the Dutch Dora-NL1 model in identifying patients who require further assessment following uncomplicated cataract surgery. Our findings demonstrated high overall accuracy of Dora-NL1’s symptom evaluation and care management decision-making at both 1 and 4 weeks after surgery, when compared to the clinician-supervised Dora-NL1 calls. At week 1, Dora-NL1 exhibited a perfect sensitivity of 100% for care management decision-making when compared to a regular telephone consultation. However, at 4 weeks, clinical issues in four asymptomatic patients were identified during face-to-face consultations, which were not detected by Dora or by the clinician reviewing the calls. In terms of

user-acceptability, most patients expressed a positive attitude towards the tool, highlighting its simplicity, ease of use, and audibility.
The Dora-NL1 model was designed to err on the side of caution, thereby ensuring that patients classified as having ‘no concerns’ by Dora are unlikely to have underlying clinical issues. Additionally, in cases of ‘incomplete information’, patients were assumed to have potential concerns, which may have increased the rate of false positives and with that, lowered the specificity. The sensitivity and specificity of the automated Dora-NL1 outcomes compared to the supervised Dora-NL1 calls were high both at 1 and 4 weeks. Only the sensitivity for floaters at week 1 was reduced as Dora misunderstood two patients. In addition, the kappa agreement for the symptom ‘pain’ at week 1 was low due to the low prevalence of pain symptoms in this cohort (n=1). When comparing Dora-NL1 to regular care at week 1, no clinical issues were missed by Dora-NL1, resulting in 100% sensitivity, while the specificity was 41.86% due to the high rate of false positives.

Comparison with Prior Work

Previous studies have demonstrated that telephone follow-ups are a safe method for identifying significant issues after cataract surgery.[10, 14, 15] In this study, three patients had inflammatory cells, which required a management change during their 4-week hospital visit. These complications were detected during slit lamp examination, highlighting a limitation of telephone consultations, which may fail to identify such issues in asymptomatic patients. Postoperative inflammatory cells usually peak at day one after surgery and may take up to 3 months to resolve.[16] Studies have demonstrated a correlation between prolonged intraocular inflammation and an increased likelihood of developing cystoid macula edema (CME), suggesting a threshold beyond which extended corticosteroid treatment is advisable.[17] However, randomized controlled trials comparing corticosteroid treatments vs. placebo for inflammation after cataract surgery have shown that only 2.7-3.8% of patients in the placebo groups developed CME within three months post-surgery vs. 1.4% to 3.8% in the corticosteroid-treated groups.[18, 19] These findings pose the question of clinical relevance of the presence of inflammatory cells in asymptomatic patients who have no symptoms of red eye or ocular pain or decreased visual acuity due to clinically significant CME. Additionally, another patient in our cohort who required a management change at 4 weeks erroneously reported experiencing binocular diplopia while there appeared to be monocular diplopia in the non-operated eye. Notably, this symptom was also missed by the clinician-led call at week 1.

The English Dora-R1 model has been investigated in the UK, showing it to be a safe alternative for cataract surgery follow-up, with automated calls conducted 3-4 weeks postoperatively and reviewed by an ophthalmologist.[11, 20] The sensitivity of care management decision-making is 93.75% (95%CI 84.76;98.27) and the specificity is 86.26% (95%CI 79.16;91.56), with an accuracy of 88.72 (95%CI 93.42;92.79).[11] These results are consistent with our findings in Tables 2 and 4. In the UK cohort, 10 patients who had been discharged following a Dora-R1 call – based on both Dora’s and the supervising clinician’s decision – experienced unexpected management changes at their 4-week routine hospital visit. These patients were either asymptomatic or exhibited mild symptoms that were not reported during the Dora-R1 call. Among them, three were diagnosed with CME, five presented with inflammatory cells, one experienced a refractive surprise, and one developed early posterior capsular opacification.[11] These findings are consistent with our results, as the majority of cases not detected via the telephone call were in asymptomatic patients with inflammatory cells in the anterior chamber.

The NPS for Dora-R1 ranges from +45 to +51, with mean scores of 8.55±1.67 and 8.59±2.05. The TUQ had an overall mean score of 3.96±0.47 (95%CI 3.76;4.16).[11, 21] Our study reported lower patient acceptability scores which likely resulted from differences in study design; unlike the UK-studies, where patients received either only a Dora-R1 call, or a Dora-R1 call along with a hospital visit at 4 weeks, all our participants received both automated Dora-NL1 calls and regular follow-ups including telephone consultations and hospital visits. Consequently, patient expectations may have differed as they directly compared Dora to the regular care they received. Additionally, the Dora-NL1 model, based on the pre-trained English version, is constrained by the more limited availability of Dutch language models at the time of Dora-NL1's development, affecting the text-to-speech and speech-to-text functionalities.[22] This limitation may have contributed to a less smooth user experience. Among patients who reported lower acceptability scores, the primary reported reasons were the limited understanding by Dora-NL1, which did not facilitate genuine conversation, and lacked personal engagement.

Limitations

Limitations of this study mainly included the potential selection bias. There were eligible patients who decided to not participate due to the expected difficulties of the Dora-NL1 call or who strongly opposed eHealth solutions in general. This population may not represent the broader cataract population, limiting generalizability. In addition, this is a mono-center study, which reduces the applicability to other healthcare facilities in the Netherlands as there is a high variability in practice patterns of the postoperative cataract care pathway.

Future implementation

Our findings indicated that at the 1-week follow-up, the Dora-NL1 model accurately identified 41.4% of patients as having 'no clinical concerns'. This suggests that approximately four out of ten patients could avoid clinician callbacks, potentially enhancing healthcare efficiency and reducing costs. Telephone reviews have been shown to be a safe time- and cost-saving alternative to short-term hospital visits after routine surgeries, with high patient satisfaction.[14, 15, 23] Moreover, automated calls allow clinicians to save time and focus on other care tasks, while ensuring a standardized approach to symptom evaluation and decision-making. This consistency reduces variability, improves the quality of patient care, and facilitates the collection and comparison of patient outcomes. In the UK cohort, the average staff cost savings amount to £35.18 per patient compared to regular care.[11] Further research is necessary to evaluate the potential cost-effectiveness in the Netherlands.

To implement automated calls in routine postoperative cataract care, several practical issues must be considered, particularly the significant variability in postoperative care across clinics and countries.[24] In some clinics, patients may not receive follow-up visits, as research indicates that for those undergoing uncomplicated cataract surgery without preexisting ocular comorbidities, follow-up can be safely deferred to 2 to 4 weeks post-surgery without significant differences in visual outcomes.[25, 26] However, automated phone calls for interim check-ups could enhance patients' confidence and reassurance during their recovery without straining healthcare resources. Despite this potential benefit, the current Dutch Cataract Guidelines do not yet recommend remote care for the final 4-week postoperative follow-up due to the limited evidence regarding its safety, and the necessity of visual acuity and refraction assessments for the mandatory quality registry of cataract surgeries in the Netherlands.[27] Further development of the Dora-NL1 model is expected to improve its accuracy and expand its applicability.

Conclusions

Based on the results of this study, we can conclude that the Dutch artificial intelligence-based cataract follow-up call system, Dora-NL1, is a safe alternative for telephone consultations in the postoperative cataract care pathway, and is suitable as a screening tool for postoperative complications during the follow-up of cataract surgery. However, as this tool only captures subjective information, it is not able to fully replace final 4-week in-person hospital visits after cataract surgery in the Netherlands. Future research is necessary to investigate the practical implications and (cost)effectiveness of this AI-driven teleconsultation system.

Funding

This study was funded by a research grant from UitZicht and the Bayer Ophthalmic Care Award (BOCA). UitZicht is a cooperation of several Dutch funds for people with visual impairment to fund a.o. scientific research into eye diseases. The BOCA is an award for healthcare professionals and researchers in ophthalmology for healthcare innovations.

Conflicts of Interest

J.C. Wanten has no financial disclosures.
N.J.C. Bauer is a consultant for Alcon, has received lecture fees from Teleon Surgical B.V., and has received research support of Teleon Surgical B.V. and Alcon.
A. Higham is an employee at Ufonia Limited.
M. Chowdhury is an employee at Ufonia Limited.
N. de Pennington is the Director and Shareholder of Ufonia Limited.
F.J.H.M. van den Biggelaar has no financial disclosures.
R.M.M.A. Nuijts is a consultant for Alcon, Johnson&Johnson and Carl Zeiss Meditec, has received lecture fees from Alcon and Ophtec, and has received research support from Alcon, Teleon Surgical B.V. and Carl Zeiss Meditec.

Abbreviations

BOCA: Bayer Ophthalmic Care Award
CI: Confidence Interval
CME: Cystoid Macular Edema
D: Diopter
DSBCS: Delayed Sequential Bilateral Cataract Surgery
FN: False Negatives
FP: False Positives
IOL: Intraocular Lens
ISBCS: Immediate Sequential Bilateral Cataract Surgery
NPS: Net Promoter Score
SD: Standard Deviation
TN: True Negatives
TP: True Positives
TUQ: Telehealth Usability Questionnaire
UK: United Kingdom

Data availability

Data is available upon reasonable request.

References

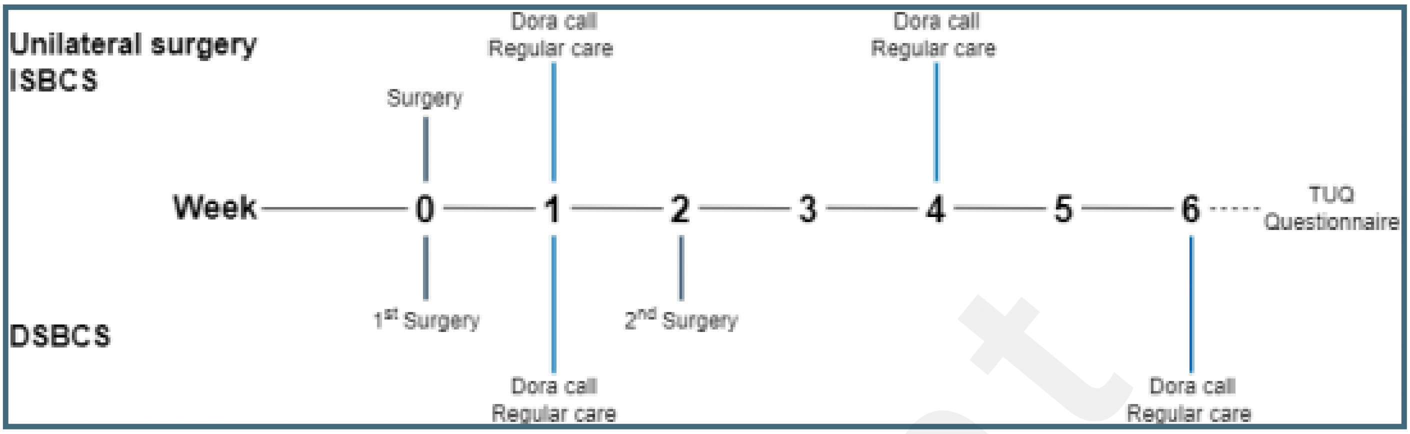
1. Berkowitz ST, Finn AP, Parikh R, Kuriyan AE, Patel S. Ophthalmology Workforce Projections in the United States, 2020 to 2035. Ophthalmology. 2024 Feb;131(2):133-9. PMID: 37739231. doi: 10.1016/j.ophtha.2023.09.018.
2. Dall TM, Gallo PD, Chakrabarti R, West T, Semilla AP, Storm MV. An aging population and growing disease burden will require a large and specialized health care workforce by 2025. Health Aff (Millwood). 2013 Nov;32(11):2013-20. PMID: 24191094. doi: 10.1377/hlthaff.2013.0714.
3. Murphy A, McElnea E, Byrne S. Health technology assessment: A primer for ophthalmology. Eur J Ophthalmol. 2018 Jul;28(4):358-64. PMID: 29973070. doi: 10.1177/1120672117747040.
4. Li JO, Liu H, Ting DSJ, Jeon S, Chan RVP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. Prog Retin Eye Res. 2021 May;82:100900. PMID: 32898686. doi: 10.1016/j.preteyeres.2020.100900.
5. Tognetto D, Brezin AP, Cummings AB, Malyugin BE, Evren Kemer O, Prieto I, et al. Rethinking Elective Cataract Surgery Diagnostics, Assessments, and Tools after the COVID-19 Pandemic Experience and Beyond: Insights from the EUROCOVCAT Group. Diagnostics (Basel). 2020 Dec 2;10(12). PMID: 33276612. doi: 10.3390/diagnostics10121035.
6. Tognetto D, Giglio R, Vinciguerra AL, Milan S, Rejdak R, Rejdak M, et al. Artificial intelligence applications and cataract management: A systematic review. Surv Ophthalmol. 2022 May-Jun;67(3):817-29. PMID: 34606818. doi: 10.1016/j.survophthal.2021.09.004.
7. Lindegger DJ, Wawrzynski J, Saleh GM. Evolution and Applications of Artificial Intelligence to Cataract Surgery. Ophthalmol Sci. 2022 Sep;2(3):100164. PMID: 36245750. doi: 10.1016/j.xops.2022.100164.
8. Stein JD. Serious adverse events after cataract surgery. Curr Opin Ophthalmol. 2012 May;23(3):219-25. PMID: 22450221. doi: 10.1097/ICU.0b013e3283524068.
9. Moustafa GA, Borkar DS, Borboli-Gerogiannis S, Greenstein SH, Lorch AC, Vasan RA, Kloek CE. Optimization of cataract surgery follow-up: A standard set of questions can predict unexpected management changes at postoperative week one. PLoS One. 2019;14(9):e0221243. PMID: 31536500. doi: 10.1371/journal.pone.0221243.
10. Lim JS, Tan SS, Yeo YL, Hong M, Teo AWJ, Lee YF, et al. Replacing the postoperative week 1 visit after routine phacoemulsification with a telephone consult. Can J Ophthalmol. 2024 Apr 8. PMID: 38604239. doi: 10.1016/j.jcjo.2024.03.015.
11. Meinert E, Milne-Ives M, Lim E, Higham A, Boege S, de Pennington N, et al. Accuracy and safety of an autonomous artificial intelligence clinical assistant conducting telemedicine follow-up assessment for cataract surgery. EClinicalMedicine. 2024 Jul;73:102692. PMID: 39050586. doi: 10.1016/j.eclinm.2024.102692.
12. Krol MW, de Boer D, Delnoij DM, Rademakers JJ. The Net Promoter Score--an asset to patient experience surveys? Health Expect. 2015 Dec;18(6):3099-109. PMID: 25345554. doi: 10.1111/hex.12297.
13. Parmanto B, Lewis AN, Jr., Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil. 2016 Spring;8(1):3-10. PMID: 27563386. doi: 10.5195/ijt.2016.6196.
14. Ruiss M, Pai V, Pilwachs C, Bayer N, Palkovits S, Findl O. Quality assurance via telephone interviews after cataract surgery: An explorative study. PLoS One. 2024;19(3):e0298149. doi: 10.1371/journal.pone.0298149.
15. Al-Ani HH, Li S, Niederer RL. Telephone follow-up one day post-cataract surgery. Clinical & experimental optometry. 2023;106(7):741-5. doi: https://dx.doi.org/10.1080/08164622.2022.2146482.

16. De Maria M, Iannetta D, Cimino L, Coassin M, Fontana L. Measuring Anterior Chamber Inflammation After Cataract Surgery: A Review of the Literature Focusing on the Correlation with Cystoid Macular Edema. Clin Ophthalmol. 2020;14:41-52. PMID: 32021067. doi: 10.2147/OPTH.S237405.
17. Aaronson A, Taipale C, Achiron A, Aaltonen V, Grzybowski A, Tuuminen R. Relationship Between Prolonged Intraocular Inflammation and Macular Edema After Cataract Surgery. Transl Vis Sci Technol. 2021 Jun 1;10(7):15. PMID: 34125145. doi: 10.1167/tvst.10.7.15.
18. Donnenfeld E, Holland E. Dexamethasone Intracameral Drug-Delivery Suspension for Inflammation Associated with Cataract Surgery: A Randomized, Placebo-Controlled, Phase III Trial. Ophthalmology. 2018 Jun;125(6):799-806. PMID: 29397189. doi: 10.1016/j.ophtha.2017.12.029.
19. Tyson SL, Bafna S, Gira JP, Goldberg DF, Jones JJ, Jones MP, et al. Multicenter randomized phase 3 study of a sustained-release intracanalicular dexamethasone insert for treatment of ocular inflammation and pain after cataract surgery. J Cataract Refract Surg. 2019 Feb;45(2):204-12. PMID: 30367938. doi: 10.1016/j.jcrs.2018.09.023.
20. Khavandi S, Lim E, Mole G. PATIENT ACCEPTABILITY OF TELEPHONE FOLLOW UP AFTER CATARACT SURGERY. BMJ Leader. 2020;4(Supplement 1):A41-A2. PMID: 638512468. doi: https://dx.doi.org/10.1136/leader-2020-FMLM.110.
21. Khavandi S, Lim E, Higham A, de Pennington N, Bindra M, Maling S, et al. User-acceptability of an automated telephone call for post-operative follow-up after uncomplicated cataract surgery. Eye (London, England). 2023;37(10):2069-76. doi: https://dx.doi.org/10.1038/s41433-022-02289-8.
22. Chowdhury M. LE HA, McKinnon R., Ventoura N., He Y.V., de Pennington N. Can Large Language Models Safely Address Patient Questions Following Cataract Surgery? Proceedings of the 5th Clinical Natural Language Processing Workshop 2023:131-7.
23. Soegaard Ballester JM, Scott MF, Owei L, Neylan C, Hanson CW, Morris JB. Patient preference for time-saving telehealth postoperative visits after routine surgery in an urban setting. Surgery. 2018 Apr;163(4):672-9. PMID: 29398042. doi: 10.1016/j.surg.2017.08.015.
24. Rossi T, Romano MR, Iannetta D, Romano V, Gualdi L, D'Agostino I, Ripandelli G. Cataract surgery practice patterns worldwide: a survey. BMJ Open Ophthalmol. 2021;6(1):e000464. PMID: 33501377. doi: 10.1136/bmjophth-2020-000464.
25. Deshpande R, Vora U, Mangiraj V, Dole K, Deshpande M. Can the postoperative follow-up visit be deferred up to four weeks after an uneventful cataract surgery? - A randomized controlled trial. Indian J Ophthalmol. 2021 Jun;69(6):1409-13. PMID: 34011710. doi: 10.4103/ijo.IJO_2390_20.
26. Kessel L, Andresen J, Erngaard D, Flesner P, Tendal B, Hjortdal J. Safety of deferring review after uneventful cataract surgery until 2 weeks postoperatively. J Cataract Refract Surg. 2015;41(12):2755-64.
27. Gezelschap NNO. Richtlijn Cataract - Zorg op afstand bij cataract. 2023.

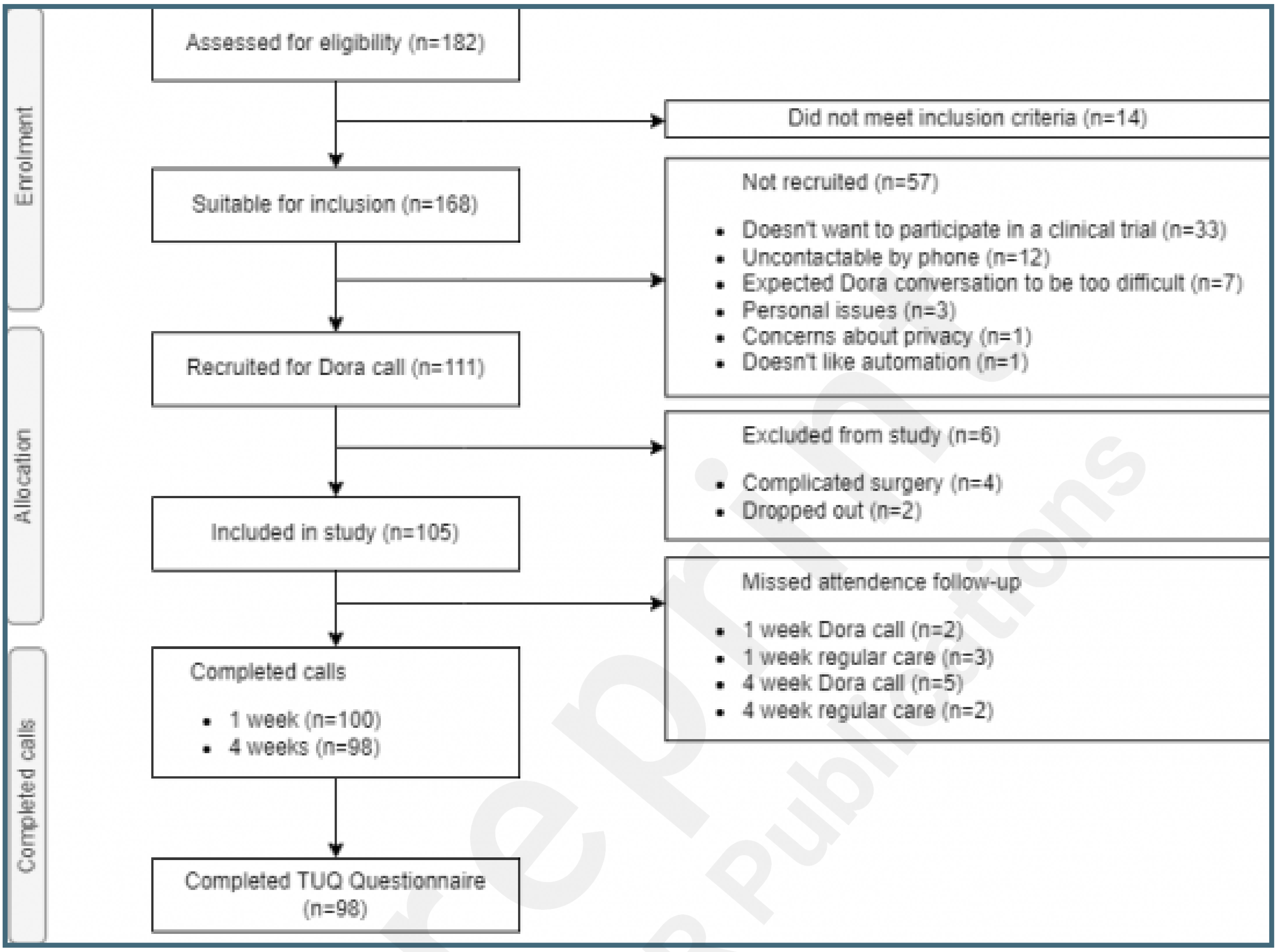
Supplementary Files

Figures

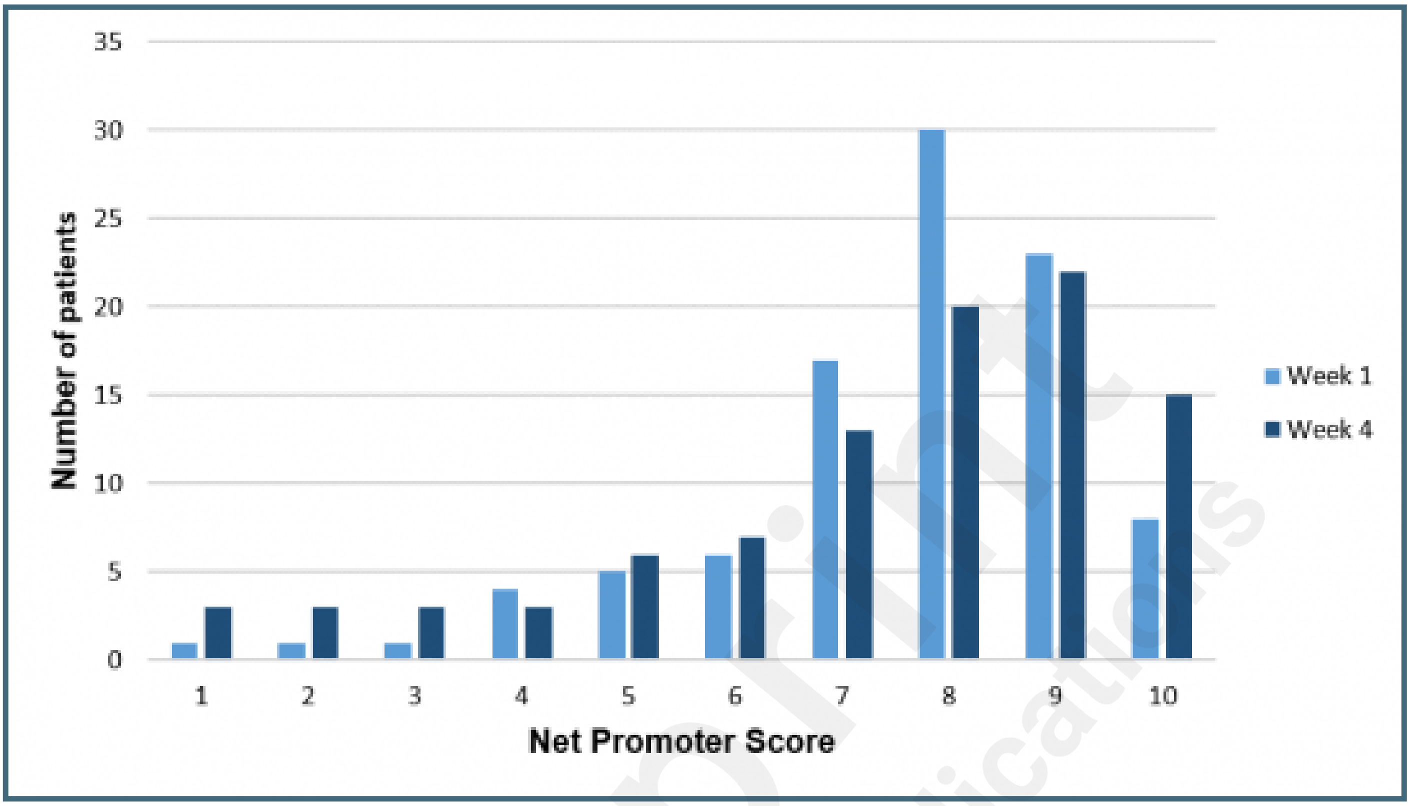
Patient pathway.



Flow chart study.



Net Promotor Score Dora-NL1 Calls at 1 and 4 weeks postoperatively.



Multimedia Appendixes

Dora decision algorithm and symptom classification.
URL: http://asset.jmir.pub/assets/e4b6473c24ad8b80c6121a3a2f0b9d36.pdf

Agreement between Dutch and English models across different patient scenarios.
URL: http://asset.jmir.pub/assets/2b3b7934aa991b3c011ddb679c846afc.docx

Outcomes Telehealth Usability Questionnaire (TUQ).
URL: http://asset.jmir.pub/assets/5ce4886e53ceef2529defd2f643fef61.docx