

<https://doi.org/10.1038/s41746-025-02332-4>

Economic evaluation of a digital symptom checker for endometriosis using a Markov decision process model



Yihan Xu¹ ✉, Carley Prentice¹, Sergio Torres-Rueda², András Meczner^{1,3}, Jan Multmeier⁴, Aidan Wickham¹, Laura Kelly⁵, Anna Klepchukova¹, Heorhi Stsefanovich¹ & Liudmila Zhaunova¹

Digital symptom checkers (SCs) are increasingly used to support early symptom recognition and care-seeking, yet evidence on their cost-effectiveness remains limited. We conducted an economic evaluation of a digital SC for endometriosis, a prevalent but underdiagnosed condition, as a case study. We developed a Markov decision process model to compare the digital SC with the standard of care from a societal perspective. Over a 40-year horizon, the digital SC reduced diagnostic delay by 4.36 years, generated 0.049 quality-adjusted life years (QALYs) per person, saved \$5196.22 in costs, and produced an incremental net monetary benefit (INMB) of \$10,089.00 at a \$100,000/QALY threshold. Probabilistic sensitivity analysis confirmed the robustness of these findings, with an INMB of \$12,398.92 (95% CI: \$11,893.11–\$12,904.72). Scenario analyses showed that the SC remained cost-effective under a wide range of assumptions, with the greatest value realized when sensitivity and specificity were ≥ 0.7 , compliance exceeded 45%, and a time horizon of at least 10 years. This study provides the first economic evaluation of a digital SC for endometriosis and illustrates when and how digital SCs can deliver value to patients and health systems.

Digital symptom checkers (SCs), also referred to as symptom assessment applications, are tools that help individuals self-assess their symptoms and decide on whether and how to seek medical care¹. Digital SCs come in various forms, from simple web-based questionnaires to chatbots powered by large language models (LLMs)². These tools are not only increasingly used by individuals but are also valued by health professionals, who report that receiving patient-entered symptom information in advance can support more efficient consultations³. Some national healthcare systems have also adopted digital SCs, particularly during periods of strain, to create standardized remote assessments and guide patients to appropriate care settings^{4,5}.

Early evaluations of digital SCs primarily examined accuracy and safety, particularly in triage, with mixed results^{6–8}. Some studies have raised concerns about digital SCs variability in accuracy and potential disruption to the patient-physician relationship, underscoring the need for more validation and oversight^{8–11}. However, other studies suggest that digital SCs match or even outperform medical professionals in directing users to the appropriate care and exhibit high triage safety^{12–15}. When implemented effectively, digital SCs can support faster diagnosis and reduce healthcare system burden^{5,16,17}. This has led to growing interest in their use for conditions that

are underdiagnosed, slow to detect, or poorly understood by patients and even clinicians.

Endometriosis, a chronic gynecological condition affecting 6–10% of women of reproductive age, marked by inflammation, pelvic pain, and infertility, presents a compelling case for examining the potential value of digital SCs^{18,19}. Diagnosis of endometriosis is often delayed (by 7 years or longer), due to non-specific symptoms, limited availability of non-invasive yet accurate tests²⁰, and low disease awareness among both patients and healthcare providers^{19,21}. These delays can lead to reduced quality of life and higher societal costs through unnecessary healthcare use and loss of productivity^{22–25}. In the U.S., the total annual cost of endometriosis, including healthcare costs and productivity losses, is estimated at \$78–119 billion^{24,26}. In Europe, the annual costs per woman are around €9579, with productivity losses comprising up to 75% of total costs^{25,27}. These figures underscore the urgent need for early identification of high-risk individuals to mitigate these socio-economic impacts. Digital SCs may help shorten the path to endometriosis diagnosis and eventually treatment by prompting earlier symptom recognition and clinical engagement²⁸. Recently developed digital SCs embedded within women health apps have exemplified this approach, offering users medical guideline-based self-assessments of their

¹Flo Health, London, UK. ²London School of Hygiene & Tropical Medicine, London, UK. ³Institute for Clinical Data Management, Semmelweis University, Budapest, Hungary. ⁴TotalEffects GmbH, Berlin, Germany. ⁵York Health Economics Consortium, York, UK. ✉e-mail: yihan_xu@flo.health

Table 1 | Base case analysis results

Outcome	Intervention: Flo Symptom Checker	Comparator: Standard care	Incremental results
Costs per person (\$)	\$7394.00	\$12,590.22	−\$5196.22
- Direct medical costs	\$5113.00	\$9155.00	−\$4042.00
- Intervention costs	\$391.00	\$0.00	\$391.00
- Indirect costs	\$1890.00	\$3435.00	−\$1545.00
QALYs per person	18.80	18.75	0.049
Diagnostic delay (years)	2.90	7.27	−4.36
ICER (£/QALY)	–	Dominant	–
INMB (WTP = \$100,000/QALY)	–	\$10,089.00	–
INMB (WTP = \$50,000/QALY)	–	\$7642.68	–

Results are presented over a 40-years horizon, with both costs and QALYs discounted at 3% annually.

risk for gynecological conditions²⁹. Preliminary evidence suggests that some of them can help users identify symptom patterns indicative of endometriosis, providing a scalable, non-invasive screening tool for earlier symptom recognition^{30,31}.

While digital SCs hold promise for facilitating earlier symptom identification of conditions like endometriosis, where delayed diagnosis and high societal costs make early identification important, there is currently no published evidence on their cost-effectiveness. Existing economic evaluations focus on diagnostic testing or treatment strategies for endometriosis, overlooking the role of early digital self-assessment tools on when and how women engage with the healthcare system^{32–36}.

This study aims to inform that gap in the evidence. We evaluated the cost-effectiveness of ‘Flo SC’, an early symptom self-assessment tool for endometriosis embedded within a female health digital platform Flo Health, compared to the standard of care (symptom-driven presentation to primary care and routine diagnostic and treatment pathways). We developed and validated a decision-analytic model to estimate Flo SC’s potential impact on time to diagnosis, healthcare costs, and quality-adjusted life years. In doing so, we aim to contribute to the small but growing evidence base on when and how digital SCs might offer value to patients and health systems alike.

Results

Base case analysis

The base case analysis, conducted over a 40-year time horizon and discounted accordingly, suggests that the Flo SC was likely to be a dominant strategy compared to standard care, potentially providing both cost savings and improved health outcomes.

Table 1 presents the average costs, quality-adjusted life years (QALYs), and diagnostic delays associated with Flo SC and the standard of care. The introduction of Flo SC improves diagnostic delay by 4.36 years and reduces average costs per patient by \$5196.22 compared to the standard of care. The incremental QALY gain per patient is modest at 0.049. The incremental net monetary benefit (INMB) per patient is estimated to be \$10,089.00 at a willingness-to-pay (WTP) threshold of \$100,000 per QALY, and \$7642.68 at a WTP of \$50,000 per QALY. Breaking costs down further, the majority of savings came from direct medical costs savings (\$4042 per patient), complemented by reductions in indirect costs (\$1545 per patient). In comparison, intervention costs (i.e., Flo SC subscription) were modest at \$391 per patient.

Deterministic sensitivity analyses

To assess the robustness of the base case findings, we conducted one-way deterministic sensitivity analyses (DSA), evaluating how uncertainty in

individual model parameters impacts the INMB of Flo SC versus the standard of care.

The one-way DSA results, presented in Fig. 1, indicate that health state utilities, Flo SC specificity and sensitivity, uptake probability, and compliance with the self-assessment results are the most influential drivers of cost-effectiveness. Higher specificity and sensitivity are associated with improved INMB, as are higher uptake and compliance with screening recommendations. In contrast, healthcare costs and productivity losses due to endometriosis symptoms have a relatively smaller impact. Utility values ranked as the top drivers for two reasons: (1) their impacts on NMB are direct and get amplified by the WTP threshold (which is high in our study); (2) with utilities bounded between 0 and 1, the same proportional uncertainty has a larger relative effect than costs, which vary across wider ranges.

We conducted a two-way DSA to examine the joint impact of two key parameters—Flo SC sensitivity and specificity—on its effectiveness in reducing diagnostic delay and its cost-effectiveness, as measured by INMB. Given that digital SCs often undergo iterative improvements in their algorithms, and that their accuracy may vary in real-world settings, a two-way DSA allows us to account for their inherent uncertainty and identify thresholds at which digital SCs remain cost-effective or reduce diagnostic delay.

As illustrated in Fig. 2, Flo SC generally leads to a positive INMB and a reduction in diagnostic delay across a range of sensitivity and specificity values, provided both remain above certain thresholds.

The left panel shows that higher Flo SC accuracy levels lead to positive INMB, indicating cost-effectiveness. The green zone, corresponding to higher sensitivity and specificity, represents combinations of sensitivity and specificity levels where Flo SC remains cost-effective. In contrast, the blue zone represents scenarios where the Flo SC could be cost-ineffective or even result in longer diagnostic delay. The contours within the green zone highlight that INMB increases with improved accuracy. At a WTP threshold of \$100,000 per QALY, the Flo SC achieves strong cost-effectiveness when both sensitivity and specificity are above 0.7.

The right panel depicts a similar pattern—high sensitivity and specificity are also crucial for reducing diagnostic delays. This result arises from the underlying logic of the care pathway, where any user flagged by Flo SC, whether a true or false positive, is likely to seek medical attention sooner. Higher sensitivity ensures more true cases are identified early, while higher specificity minimizes unnecessary referrals and reduces noise in the diagnostic process, indirectly facilitating faster diagnosis for true positives. As a result, the greatest reductions in diagnostic delay (up to 5 years) occur when both metrics exceed 0.8.

Probabilistic sensitivity analysis (PSA)

We performed a probabilistic sensitivity analysis (PSA) using 1000 Monte Carlo simulations to assess the joint uncertainty of key model input parameters, including costs, utilities, and transition probabilities. The PSA results show that Flo SC was cost-effective in 93.7% of scenarios at the \$100,000/QALY threshold. The mean cost per patient with Flo SC was \$7797.55, compared to \$13,076.09 under standard care, saving \$5278.54 per patient (95% CI: \$5212.85–\$5344.23). The intervention also resulted in a modest increase in QALYs (18.79 vs. 18.72), with an incremental gain of 0.071 QALYs (95% CI: 0.066–0.076). The INMB per patient was \$12,398.92 (95% CI: \$11,893.11–\$12,904.72). The incremental cost-effectiveness plane (Fig. 3) illustrates the majority of simulations (scatter points) falling in the lower right quadrant, indicating that Flo SC was highly likely to be cost-saving and more effective than standard care.

Additionally, Flo SC reduced the time to diagnosis by an average of 4.37 years (95% CI: 4.31–4.43) by prompting earlier self-recognition and clinical engagement. The probability that Flo SC reduces diagnostic delay by at least 3 years was 93.2%. Figure 4 shows a clear shift in diagnostic timing, with Flo SC users clustering around 3 years of delay (pink), compared to 7 years under standard care (blue).

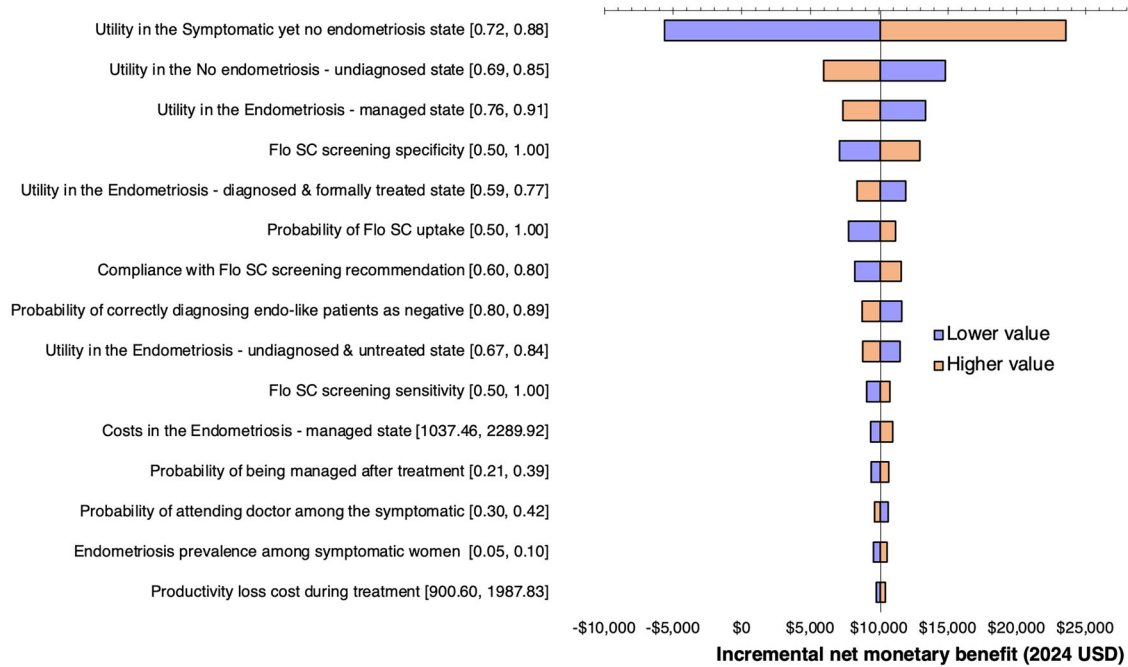


Fig. 1 | Deterministic sensitivity analysis results showing the impact of key parameter variations on the cost-effectiveness of the Flo Symptom checker versus standard of care at a willingness-to-pay threshold of \$100,000 per QALY.

Parameters at the top of the chart have the greatest influence on model outcomes, while those at the bottom have a smaller effect. Only the top 15 parameters are displayed for clarity.

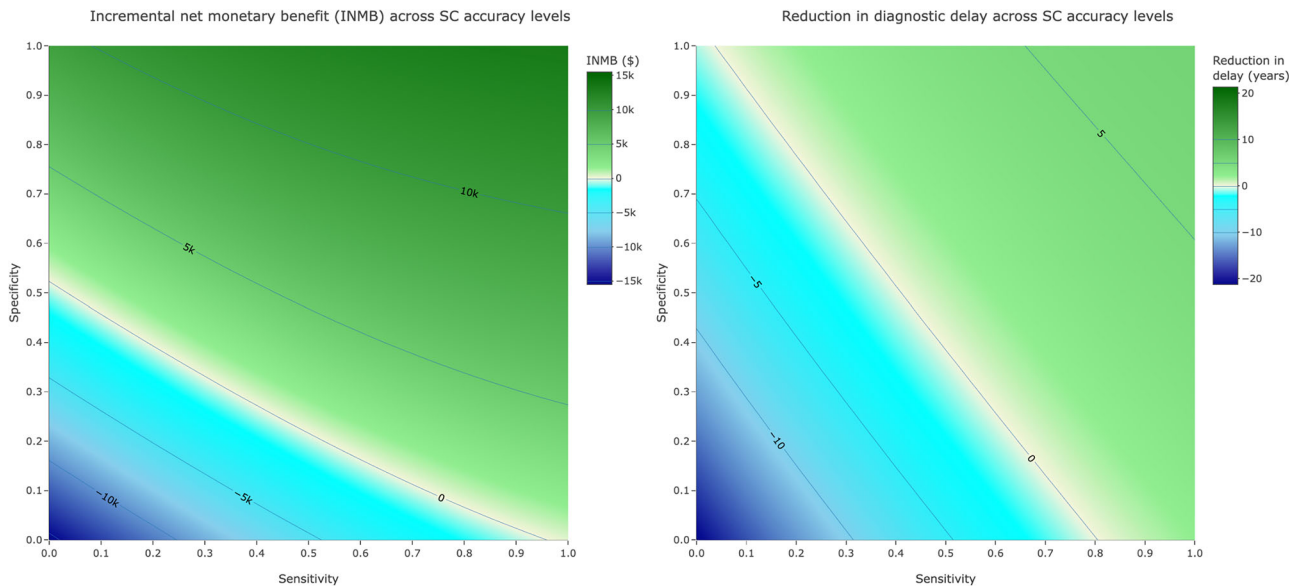


Fig. 2 | Impact of Flo Symptom Checker accuracy on incremental net monetary benefit (INMB) and diagnostic delay reduction: a two-way sensitivity analysis. The left panel illustrates the effect on INMB at a willingness-to-pay threshold of \$100,000 per QALY. The green zone represents combinations of accuracy thresholds where Flo SC remains cost-effective or reduces time to diagnosis, while the blue zone

indicates scenarios where cost-effectiveness is not achieved or leads to further diagnostic delay. The right panel shows the effects on diagnostic delay reduction. Higher specificity is particularly critical for reducing diagnostic delay, as indicated by the green zone, while lower specificity results in minimal or even longer diagnostic delay.

Scenario analyses

We conducted scenario analyses to further explore how uncertainty in key parameters could affect the health and economic outcomes of Flo SC. We first examined influential behavioral drivers identified from the DSA—compliance, uptake, probability of visiting a doctor while symptomatic. We then examined uncertainty associated with factors particularly relevant to

real-world implementation and payer decision-making, namely the time horizon and the subscription price of Flo SC.

Across most scenarios, Flo SC was a dominant strategy, providing both cost savings and greater health benefits (Table 2). However, when compliance was low (30%), the intervention became cost-increasing and was no longer considered cost-effective, with an ICER of \$265,177.35 per QALY. In

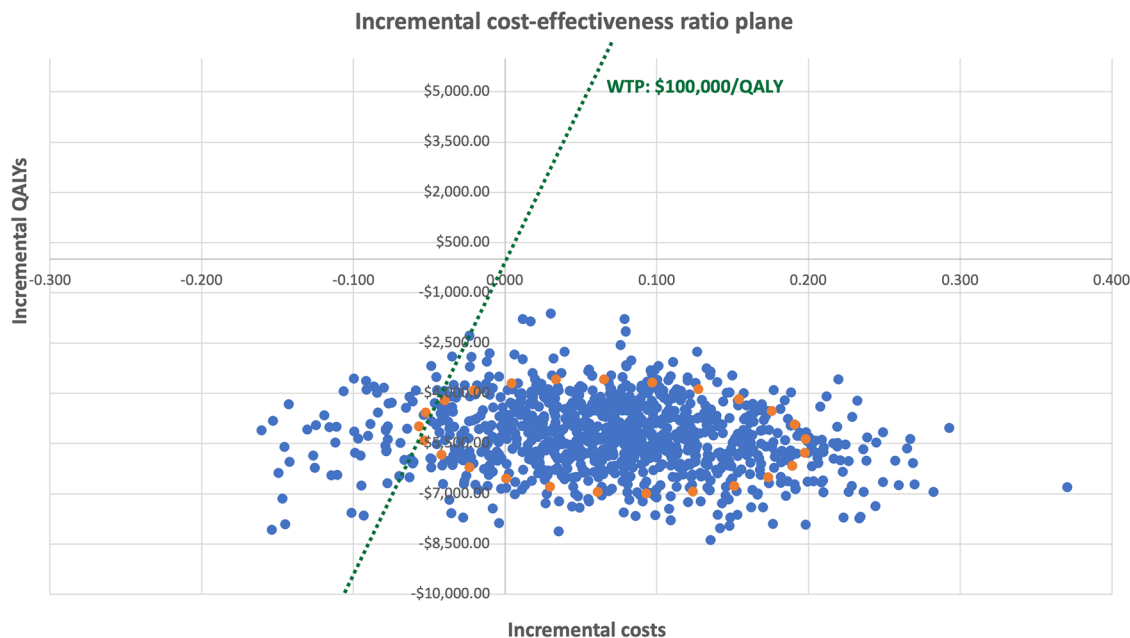


Fig. 3 | Probabilistic sensitivity analysis results, visualized on an incremental cost-effectiveness ratio (ICER) plane for Flo SC versus standard care. This scatter plot visualizes the results of 1000 Monte Carlo simulations, showing the incremental QALYs and incremental costs associated with Flo SC compared to standard care.

The orange ellipse represents the 95% confidence interval of the ICER. The dotted green line represents the \$100,000 per QALY willingness-to-pay (WTP) threshold. Most points fall in the fourth quadrant, indicating that Flo SC is generally cost-saving and more effective.

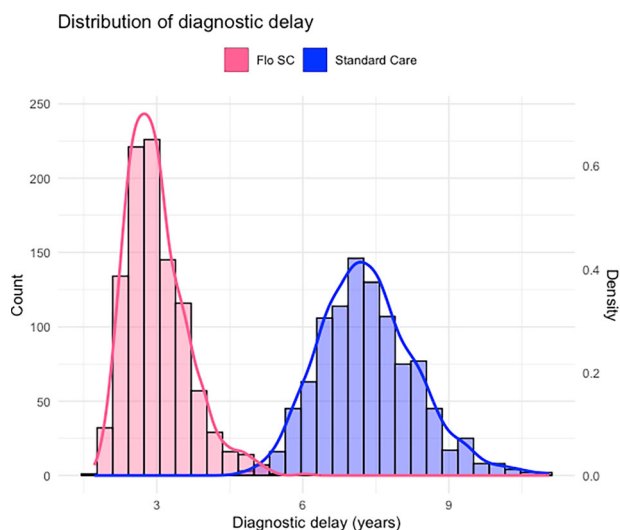


Fig. 4 | Distribution of diagnostic delay: standard of care versus Flo SC. This histogram compares the distribution of diagnostic delays between standard care (blue) and Flo SC (pink), based on 1000 Monte Carlo simulations. The density curve (scaled) on the right indicates the relative frequency of each delay. The Flo SC distribution is centered around a shorter diagnostic delay (3.04 years), whereas the standard care distribution shows a longer delay (7.41 years).

contrast, high compliance (90%) resulted in the greatest cost savings and QALY gains. While lower uptake (30%) still maintained cost-effectiveness and reduced diagnostic delay, the benefits were much more modest. To explore potential correlation between uptake and compliance, we conducted a two-way sensitivity analysis varying both parameters from 0.3 to 0.9. The results (see Supplementary Table 7a, b) reinforce the findings that cost-effectiveness is primarily driven by compliance, and when it is very low (0.3), increasing uptake actually worsened the outcomes, as more users engaged with the SC but did not act on its recommendations.

The spontaneous doctor visit scenarios showed that Flo SC was particularly cost-saving and beneficial when symptomatic women, regardless of their endometriosis status, were less proactive in seeking medical advice on their own. The time horizon analysis showed that shorter time horizons (e.g., 5 years) limited the scope of intervention effects, as some individuals may not develop symptoms of endometriosis within that period. Longer time horizons allowed for a more complete realization of the benefits, as Flo SC’s impact on healthcare utilization and symptom progression became more evident over time.

To test the robustness of our conclusions to pricing uncertainty, we varied the annual subscription cost from \$15 to \$240 per user (equivalent to 25%–400% of the current price). Across this entire range, Flo SC remained a dominant strategy, with INMB estimates ranging from \$8917 to \$10,382 per patient at a WTP threshold of \$100,000 per QALY.

Discussion

We found that a digital SC, when used as a self-assessment tool for endometriosis, was likely to be a cost-saving and health-improving intervention compared to standard of care. Over a 40-year horizon, on average, the digital SC shortened the diagnostic delay by 4.36 years, delivered modest QALY gains of 0.049 per person, reduced costs of \$5196.22 per person, and resulted in an INMB of \$10,089.00 per person at a WTP threshold of \$100,000 per QALY, primarily by improving early symptom recognition that supported reduction of diagnostic delays. The PSA showed that the digital SC remained cost-effective in more than 93.7% of simulations and reduced diagnostic delays by more than 3 years in 93.2% simulations, supporting the robustness of the base case findings.

Scenario analyses highlighted that digital SC accuracy, compliance, and uptake were the most critical drivers of cost-effectiveness. The digital SC delivered the greatest value when it provided accurate recommendations (both sensitivity and specificity were ≥ 0.7) and when users acted upon them (compliance exceeded 45%). Uptake alone had a smaller influence compared to compliance, but sufficient uptake remains a prerequisite for scaling the intervention in real-world settings. The model also showed that longer time horizons are essential to capture the full benefits of earlier diagnosis. Shorter horizons (e.g., 5 years) underestimated both cost savings and QALY

Table 2 | Scenario analyses results

Scenarios	Value	Incremental cost (\$)	Incremental QALYs	Incremental diagnostic delay (years)	ICER (\$/QALY)	INMB (WTP = \$100,000/QALY)
Base case		−\$5196.36	0.049	−4.36	Dominant	\$10,089.00
Compliance						
Low	0.3	\$2124.89	0.008	5.68	265,177.35	−\$1323.58
High	0.9	−\$6988.52	0.059	−5.69	Dominant	\$12,868.90
Uptake						
Low	0.3	−\$2677.41	0.028	−2.48	Dominant	\$5436.65
High	0.9	−\$5513.27	0.052	−4.59	Dominant	\$10,671.12
Spontaneous doctor visit						
Low	20%	−\$4772.94	0.073	−7.21	Dominant	\$12,037.29
High	50%	−\$5247.78	0.040	−3.05	Dominant	\$9241.93
Time horizon						
5 years	5	−\$184.88	0.008	−0.23	Dominant	\$984.60
10 years	10	−\$1180.71	0.021	−1.15	Dominant	\$3299.69
20 years	20	−\$3531.70	0.041	−3.34	Dominant	\$7621.87
Flo subscription fee						
Lower: 25%	\$7.50	−\$5489.32	0.049	−4.36	Dominant	\$10,381.96
Lower: 50%	\$15.00	−\$5391.67	0.049	−4.36	Dominant	\$10,284.31
Lower: 75%	\$22.50	−\$5294.02	0.049	−4.36	Dominant	\$10,186.65
Higher: 150%	\$44.99	−\$5001.06	0.049	−4.36	Dominant	\$9893.70
Higher: 200%	\$59.99	−\$4805.75	0.049	−4.36	Dominant	\$9698.39
Higher: 400%	\$119.98	−\$4024.52	0.049	−4.36	Dominant	\$8917.16

Results are presented over a 40-years horizon, with both costs and QALYs discounted at 3% annually.

gains, as many women in the simulated cohort had not yet developed symptoms.

Previous studies on digital SCs have largely focused on accuracy and triage safety^{6–8,12,37}, with economic outcomes underexplored¹⁷. This study helps fill that gap and illustrates that digital SCs can offer economic value, particularly in conditions such as endometriosis, where diagnostic delays are costly and prolonged³⁸. Our model also showed that savings were driven primarily by direct medical costs than by reductions in indirect costs related to productivity loss. This pattern is consistent with prior real-world evidence from the U.S.²⁶ and implies that our findings are potentially relevant from the perspective of healthcare payers, for whom direct medical costs are the primary concern, even though the societal burden of productivity loss remains considerable.

Our sensitivity analyses suggest that digital SC accuracy, especially specificity, was a key driver of cost-effectiveness. Our findings are aligned with previous analyses that digital SCs tend to be too risk-averse, often over-triaging patients for medical attention^{11,37}. While reduced sensitivity may delay diagnosis due to an increase in missed cases/false-negatives, low specificity can lead to unnecessary referrals/false positives that increase downstream costs. This highlights the importance of balancing the trade-offs between sensitivity and specificity in product iteration.

User behaviors—specifically, compliance and uptake—also played a major role in determining value for money. Even a well-performing digital SC will yield limited benefit if users do not engage with it or fail to act on its recommendations. These findings echo earlier work emphasizing that the success of digital health tools depends not only on technical performance but also on user adoption and adherence^{39,40}.

Moreover, the cost-effectiveness of early endometriosis diagnosis aligns with findings from studies on menstrual health interventions and digital tracking applications^{41,42}. These studies emphasize that empowering women with accurate, accessible information can lead to more proactive healthcare engagement. Consistent with these findings, our scenario analyses showed that Flo SC showed greater value when women were less likely

to seek care on their own—highlighting its potential to reduce inequities in healthcare access. Indeed, in settings where symptom awareness is low or barriers to care are high, digital SCs may play a particularly valuable role in encouraging timely clinical evaluation and treatment.

The model also demonstrated that a longer time horizon is necessary to fully capture the benefits of earlier diagnosis. Because the simulated cohort began at age 20, shorter horizons excluded individuals who had not yet developed symptoms. As a result, both health gains and cost savings appeared smaller over 5 years, but became more pronounced over 10- and 20-year horizons. This supports the use of long-term modeling for evaluating digital interventions that affect disease trajectories over time.

This study is the first known published economic evaluation of a digital SC for endometriosis and among the few assessing the cost-effectiveness of digital SCs more broadly.

The findings offer broader insights into the conditions where digital SCs may be most valuable. Digital SCs are likely to deliver the greatest economic and health benefits in conditions characterized by prolonged diagnostic delays, stigma, high societal costs, and low care-seeking rates. For example, perimenopausal symptoms, which often remain under-recognized despite significant impacts on quality of life and workforce participation^{43–45}, may benefit from digital SCs tailored to improve awareness and engagement.

The modest QALY gains observed in our model reflect both conservative assumptions on utility of diagnosis (we assumed none) and the nature of endometriosis, which has low mortality risk despite the morbidity burden⁴⁶, which naturally limits the magnitude of QALY improvements. Moreover, treatment effects for endometriosis are constrained by relapse risk and imperfect effectiveness: available therapies such as hormonal treatments or laparoscopy can relieve symptoms but are not curative and may involve side effects or peri-operative discomfort^{47,48}. Many patients also experience recurrent symptoms despite treatment^{49,50}. Consequently, even when Flo SC accelerates treatment, the resulting QALY gains remain modest. For conditions with more responsive treatments and higher

mortality risk, digital SCs may deliver larger health gains and, in turn, greater cost-effectiveness.

Our scenario analyses also reinforce the need for long-term evaluations of digital SCs, as shorter time horizons substantially underestimated both health and economic gains.

As digital SCs proliferate, particularly through consumer-facing health apps⁵¹, healthcare systems will increasingly confront decisions about how to integrate, regulate, and reimburse such tools. This study provides a methodological framework for evaluating digital SCs and offers early evidence that they can deliver clinical and economic benefits under realistic assumptions. Looking ahead, the rapid emergence of LLM-based health applications^{52,53} may further transform the digital SC landscape, offering opportunities for more personalized, conversational symptom triage while also introducing new challenges around specificity, safety, and trust. Expanding evaluation frameworks to address the opportunities and risks introduced by AI-powered tools will be crucial to ensure that the next generation of digital symptom checkers delivers meaningful clinical and societal value.

Despite these promising results, several limitations must be noted. The model assumes that women undergo digital SC check-in every 6 months while symptomatic. In practice, usage patterns are likely more variable, influenced by symptom severity, healthcare access, and user experience. Additionally, the model assumes that women receiving a false negative from the SC continue seeking care at the same rate as those who never used the tool. While not universally accurate, persistent or worsening symptoms often prompt continued care-seeking, meaning the model likely captures broader care-seeking dynamics despite these simplifications.

Another simplification concerns the modeled transition from “endometriosis-like symptoms” to “no endometriosis”. In the standard care pathway, symptom resolution only occurs after a clinical consultation, whereas spontaneous resolution without medical input is possible in reality. This may slightly underestimate quality of life in the standard care group; however, deterministic sensitivity analysis showed that even with optimistic utility values (0.85), the SC remained cost-effective (INMB: \$7078.09 per patient).

The model also uses average disease trajectories and does not distinguish between phenotypes of endometriosis. Fertility-related outcomes were also not modeled, in line with prior findings³⁴ that pain-related QALY gains tend to dominate due to their frequency and immediacy. This assumption likely underestimates the cost-effectiveness by excluding potential long-term benefits that may further favor the intervention. Treatment-specific responses and postmenopausal utility values were also averaged, and we assumed quality of life after menopause is similar for women with or without endometriosis—consistent with evidence suggesting symptom relief post-menopause¹⁹.

Real-world adoption remains a source of uncertainty. The Flo SC compliance was based on early data based on U.S. users and may shift with broader implementation. Last but not least, while the societal perspective captures a broad set of benefits, including productivity gains, it may not fully align with the incentives of all decision-makers in the U.S. healthcare system. Private payers, for instance, may be less likely to support tools where financial benefits primarily accrue to individuals or society rather than the insurer.

To address those limitations, future research should focus on the following priorities. First, real-world validation through longitudinal studies using electronic health records, claims data, or prospective cohorts to assess whether digital SCs meaningfully change diagnostic timelines, healthcare use, and outcomes⁵¹. Second, integration into care pathways is key; embedding digital SCs into telemedicine platforms, payer portals, or digital triage systems could improve early intervention, especially in complex healthcare systems like the U.S.^{54,55}. Third, boosting user engagement and equity requires addressing barriers such as digital literacy, trust, and access. Tailoring digital SCs for the diverse needs of users will be essential to maximize uptake and impact^{39,40,56}. Fourth, future research could also explore payer-specific perspectives and quantify how benefits are

distributed across stakeholders, particularly in fragmented or non-universal healthcare systems.

By addressing these gaps, digital SCs can be further optimized not only for accuracy and usability, but also for their real-world impact, cost-effectiveness, and integration into equitable care.

In conclusion, this study provides the first economic evaluation of a digital SC for endometriosis screening. Using a validated decision model, we found that the tool can reduce diagnostic delays, improve health outcomes, and be cost-saving under a wide range of assumptions. These findings offer timely evidence for healthcare systems exploring the integration of digital triage tools—particularly for underdiagnosed, high-burden conditions. When implemented effectively, digital SCs may enhance early detection and support more efficient, equitable care.

Methods

Modeling overview

We developed a Markov decision process model to evaluate the cost-effectiveness of a digital SC (‘Flo SC’) for endometriosis compared to standard care. The analysis adheres to best practices in health economic modeling, following the CHEERS 2022 guidelines⁵⁷ (see Supplementary Table 1). Following a pre-specified health economics analysis plan, the model simulates two parallel cohorts of 10,000 women of 20 years of age, based in the United States, and presenting with symptoms suggestive of endometriosis. We chose this starting age to reflect the typical onset of endometriosis during reproductive years^{18,58}.

Each cohort starts in the healthy state and enters a decision tree capturing initial symptom assessment and care-seeking behaviors, followed by a Markov state-transition process representing diagnosis, treatment and management for endometriosis. Transitions occur at 6-month intervals, reflecting typical timelines for symptom checking and follow-up consultations. The intervention cohort is assumed to use Flo SC, a chatbot-based tool embedded within the Flo Health app, which guides users through a structured, guideline-based questionnaire to self-assess symptoms. Importantly, the model assumed that Flo SC augments, rather than replaces, existing care pathways, supporting users in recognizing symptoms earlier and navigating the healthcare system more effectively. For the comparator cohort, women follow standard of care pathways, with symptom recognition and healthcare engagement based on routine care-seeking. Model outcomes include time to diagnosis, healthcare costs, productivity losses, and QALYs.

This structure was adapted from the National Institute for Health and Care Excellence (NICE)’s economic model of endometriosis care pathway³⁴ and aligns with previous published work³³ and best practices modeling guidance^{59,60}. Model conceptualization and assumptions included were refined through consultation with independent health economists and U.S.-based gynecologists and primary care physicians to ensure that the model structure reflects key features of the diagnostic and treatment care pathway for endometriosis in the US as well as patients’ care-seeking behaviors in the US healthcare system.

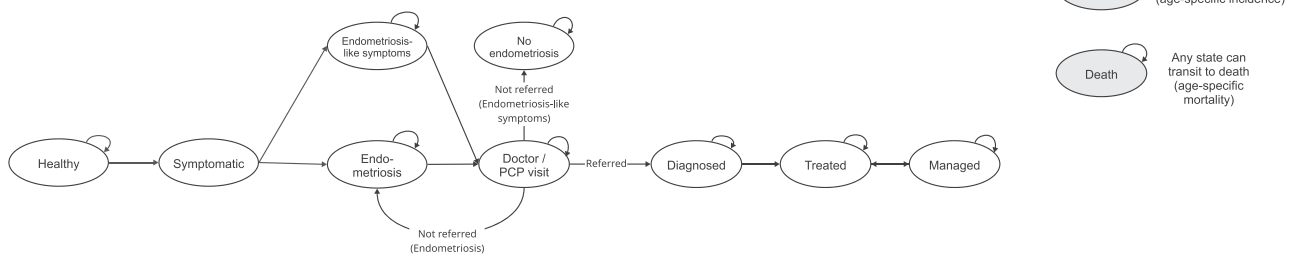
As recommended in recent reviews of economic evaluations in women’s health⁶¹, the model adopts a societal perspective, incorporating both direct healthcare costs and productivity losses. A 40-year time horizon was chosen to reflect the typical reproductive lifespan, with menopause marking symptom relief in most cases¹⁹. Costs and QALYs were discounted at an annual rate of 3%, consistent with standard economic evaluation practice in the U.S. setting⁶².

Model structure

The model reflects key stages of the endometriosis care pathway: symptom onset, initial care-seeking, diagnosis (or misdiagnosis), treatment, relapse, symptom resolution at menopause, and death. A schematic representation of both model arms is shown in Fig. 5. Women in each cohort either have endometriosis or experience similar symptoms due to other causes. To reflect real-world diagnostic uncertainty, we explicitly modeled both true-positive and false-positive pathways, including symptomatic women

Schematic diagram of Markov decision process

Standard care



Flo SC

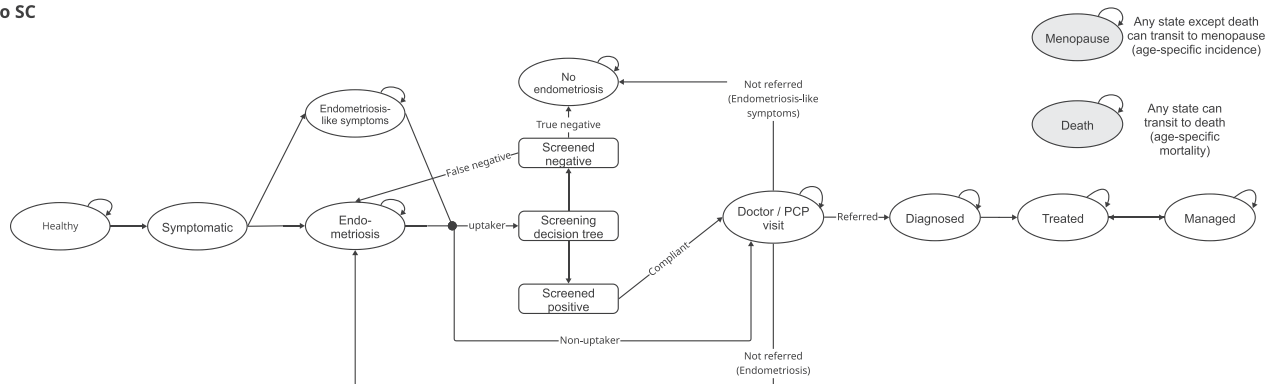


Fig. 5 | Schematic diagram of the endometriosis care pathway model.

without endometriosis who may be mislabeled to seek medical help or informed that their symptoms were not suggestive of endometriosis via Flo SC.

At model entry, healthy women may develop symptoms and transition into one of two underlying conditions: endometriosis or other causes. How a woman transitions to subsequent stages depends on the care pathway. In the intervention cohort, women first receive self-assessment via the tool, which influences the likelihood and time to seek care. In the standard of care cohort, care-seeking is based on women’s spontaneous behavior—they consult primary care physicians (PCPs) and, if needed, specialists, undergoing diagnosis and treatment based on cumulative diagnostic accuracy and empirical treatment response.

Diagnosis and treatment pathways were simplified by averaging treatment effects and omitting rare transitions such as delayed-onset endometriosis. Unlike the prior models³⁴, we excluded a separate “diagnosed but untreated” state, based on clinical expert input indicating nearly all patients receive empirical treatment after consultation.

Women may relapse or cycle through treatments, and all can transition to menopause or death at any time based on age-specific probabilities. Menopause was modeled as a semi-absorbing state, after which symptoms typically resolve and no further transitions occur apart from death. Full state definitions and a full list of transitional decision nodes are included in Supplementary Tables 2 and 3.

Model assumptions

The model incorporates a series of structural, clinical, and economic assumptions. These assumptions balance model simplicity and real-world relevance, ensuring that key drivers of diagnostic delay, treatment initiation, and healthcare costs are accurately represented. Women enter the model in a healthy state, with the potential to develop symptoms over time based on age-specific endometriosis onset rates. Once symptomatic, women are

classified as either having endometriosis or just endometriosis-like symptoms, reflecting real-world diagnostic uncertainty.

Diagnostic decisions are modeled as a cumulative process rather than a single test, consistent with the NICE approach³⁴. Instead of applying fixed sensitivity and specificity for a specific diagnostic test, the model incorporates the cumulative accuracy of primary care and gynecologist assessments over multiple visits, based on the empirical estimates of diagnostic delay from epidemiological studies²¹. Women misdiagnosed as having endometriosis are assumed to accrue costs due to misdiagnosis and potential treatment, but do not gain health benefits from formal treatment for endometriosis.

In the model, women who transition to the “No Endometriosis—Diagnosed” state (see Supplementary Table 2) are assumed not to develop endometriosis later in life. While rare cases of late-onset endometriosis may occur, this state is treated as final in the model to simplify the care pathway, similar to the approach taken by NICE³⁴. Additionally, women can transition to menopause (see Supplementary Table 4) or death at any time, based on age-specific probabilities. Menopause serves as a semi-absorbing state, where disease progression and treatment cease, reflecting the natural decline in endometriosis symptoms after reproductive years.

All symptomatic women seeking medical care receive empirical treatment, regardless of whether they have endometriosis. Empirical treatments include non-steroidal anti-inflammatory drugs (NSAIDs) or hormonal therapy, which are commonly prescribed for pelvic pain management in primary care. The effectiveness of empirical treatment is assumed to be similar across all symptomatic women, with formal treatment only initiated upon confirmed diagnosis.

The model incorporates average effectiveness estimates for treatment, rather than differentiating between specific hormonal therapies or surgical procedures. Women who experience symptom relapse transition from the

managed state back to active treatment rather than progressing through specific therapy lines.

The model does not explicitly incorporate comorbid conditions, assuming that endometriosis is the primary driver of healthcare utilization and health outcomes. Costs in the model are assigned based on healthcare utilization, diagnostic tests, and treatment interventions, reflecting real-world expenditures. Treatment costs are averaged over different therapeutic options rather than modeled separately for practical purposes.

Model input parameters

The model incorporates key input parameters grouped into three categories: transition probabilities, healthcare costs and health utilities (see Supplementary Fig. 1 for an overview of key parameters used in the model). To minimize selection bias, parameter choices were anchored in the comprehensive systematic review conducted by NICE (2017)⁶³, which remains the most rigorous evidence synthesis in this area. Where appropriate, we retained baseline parameter values from the NICE model, and for parameters that appeared outdated or not suitable for the U.S. setting, we conducted targeted PubMed literature searches (past 10 years) for English-language, peer-reviewed U.S. studies, as well as consulting the latest clinical guidelines and national statistics. Final values were reviewed with two U.S.-based clinicians for face validity.

Tables 3–5 present the parameter values, distributions, and sources for transition probabilities, costs, and health utilities.

Transition probabilities reflect the likelihood of progressing through health states as described above. Women enter the model as healthy and develop symptoms over time, with age-specific probabilities of endometriosis onset and prevalence of endometriosis among symptomatic individuals informed by epidemiological data.

Assumptions on care-seeking behaviors differ by pathway. In the standard of care cohort, care-seeking is modeled based on women's spontaneous doctor-visiting patterns. In the intervention cohort, symptomatic users may engage with Flo SC at each cycle, which guides subsequent care-seeking depending on self-assessment outcomes and user compliance with the care-seeking advice. These dynamics are governed by Flo SC uptake among the symptomatic population, its triage accuracy, and user compliance rates, informed by empirical studies on general use patterns of digital SCs and internal Flo SC usage data (see Table 3).

After diagnosis, women can begin treatment. Those who respond well enter a stable managed state, while others may experience symptom relapse and return to active treatment. The model also includes age-based transitions to menopause and death to reflect the natural course of endometriosis and capture long-term outcomes.

Since the model includes 11 health states, there are 121 possible transitions in the full transition matrix. However, only a small proportion of these transitions occur in practice, as most are non-reversible. Table 3, therefore, lists only non-zero transitions and omits self-transitions, where individuals remain in the same health state across cycles.

All costs, presented in Table 4, were expressed in 2024 USD, inflated where necessary using the U.S. Consumer Price Index⁶⁴.

As we adopted a societal perspective, the model captures both direct healthcare costs and indirect productivity losses, which was estimated to account for up to 75% of total endometriosis-related costs²⁵. Direct costs sourced from existing literature include symptomatic care, diagnosis, treatment, and long-term management of endometriosis. These costs cover primary and specialist consultations, diagnostic procedures, pharmacological therapies, surgeries, and follow-up care. The cost of Flo SC was modeled as a semi-annual subscription fee. For false positives, the estimated cost included travel expenses and productivity losses associated with the doctor visit. The direct medical cost of the visit itself was already captured in the model through the transition to the Doctor Visit health state, ensuring consistent application across all relevant patients and avoiding double-counting of diagnostic or treatment costs. No additional diagnostic or treatment costs are applied to avoid double-counting, as these are already captured through the transition to the Doctor Visit health state.

Indirect costs account for productivity losses due to absenteeism and presenteeism across different disease states, particularly before and during treatment. These costs were sourced from existing literature or derived using a human capital approach⁶⁵, which values lost productivity based on average wage (see Supplementary Table 5 for detailed cost parameters derivation and calculation process).

Health utility values, presented in Table 5, were primarily sourced from EQ-5D-based studies on endometriosis (see Supplementary Table 6 for more detail). They were assigned to each health state to reflect how endometriosis symptoms, diagnosis, treatment options and treatment response affect women's quality of life over time.

Women in the undiagnosed symptomatic state—whether due to endometriosis or similar conditions—had lower utility scores to reflect the physical discomfort and psychological distress from unresolved symptoms. The diagnosed and formally treated state also carried lower utility, acknowledging the temporary reduction in quality of life due to invasive procedures, side effects, or surgical recovery. This is consistent with findings that hormonal or surgical treatment often leads to short-term disability despite targeting long-term relief³⁶.

A separate transitional utility was applied to women receiving first-line or empirical treatment, such as NSAIDs or hormonal therapy, before formal diagnosis⁶⁶. This utility was not tied to a specific state but applied during treatment transitions for both women with and without endometriosis.

Women who responded well to treatment and entered the managed state experienced higher utility, reflecting sustained symptom control. The menopause state was assigned a population norm value for U.S. women aged 55–64⁶⁷. The death state was modeled as absorbing, with a utility of zero.

Cost-effectiveness and sensitivity analyses

To evaluate cost-effectiveness, we calculated the incremental cost-effectiveness ratio (ICER) and incremental net monetary benefit (INMB) between the Flo SC intervention and standard of care. These were computed using the following formula:

$$ICER = \frac{C_{Flo\ SC} - C_{standard\ care}}{Q_{Flo\ SC} - Q_{Standard\ care}} \text{ and } INMB = (\Delta Q \times \lambda) - \Delta C$$

where ΔC is the difference in costs, ΔQ is the difference in QALYs, and λ is the willingness-to-pay (WTP) threshold. The base case analysis used two WTP thresholds: \$100,000/QALY, in line with the threshold for economic evaluations conducted in the US⁶⁸, and a more conservative estimate of \$50,000/QALY.

In addition to cost-effectiveness, we modeled time to diagnosis to assess how Flo SC affects diagnostic delays, using outputs from the Markov trace. In each cycle, we calculated the number of women in the *undiagnosed endometriosis* state and the theoretical minimum time to diagnosis (assuming diagnosis would occur in the next cycle). The difference between these was used to estimate diagnostic delay, which was summed across cycles over the specified time horizon. This approach allowed us to estimate the average diagnostic delay per woman, reported in undiscounted years, for both the Flo SC cohort and the standard of care cohort.

We performed a series of sensitivity analyses to assess the robustness of model findings. We did one-way deterministic sensitivity analysis (DSA) to explore the influence of individual input parameters on cost-effectiveness by varying key variables one at a time within their plausible ranges. We also conducted a probabilistic sensitivity analysis (PSA) using 1000 Monte Carlo simulations to account for uncertainty across all model inputs. Additionally, we carried out scenario analyses to test alternative model assumptions regarding screening uptake, compliance, Flo SC accuracy, and time horizon to assess real-world implementation conditions and policy relevance.

Table 3 | Model inputs of transition probabilities, their corresponding distributions, and sources

Transition from	Transition to	Variable name	Description	Point estimate ^a	Distribution	Source
Healthy	Symptomatic	incidence_endo[age]	Age-specific incidence of endometriosis onset	See S-Table 4	Constant; Symptom onset age mean = 21.58 y/o, SD = 8.61	Hadfield et al. ⁶⁸
Symptomatic	Endometriosis	prevalence	Prevalence of endometriosis among the symptomatic	0.074	Beta ($\alpha = 35.29$, $\beta = 441.59$)	Zondervan et al. ⁷⁰
Symptomatic	Endometriosis-like symptoms	1 – prevalence	Complementary probability of endometriosis prevalence	0.926	Derived	Zondervan et al. ⁷⁰
Endometriosis	Doctor visit	tp/visitDoctor	Probability of visiting a doctor	0.36	Beta ($\alpha = 78.12$, $\beta = 138.88$)	NICE ³⁴
Endometriosis-like symptoms	Doctor visit	tp/visitDoctor	Probability of visiting a doctor	0.36	Beta ($\alpha = 78.12$, $\beta = 138.88$)	NICE ³⁴
Symptomatic	Screening	uptake	Uptake of the digital SC among the symptomatic sub population	0.80	Beta ($\alpha = 4.2$, $\beta = 1.05$)	Assumed, informed by use patterns of digital SCs ^{51,71} and Flo internal data
Screening Endometriosis	Screened positive	sensitivity	Probability of screened positive when having endometriosis	0.73	Beta ($\alpha = 346.75$, $\beta = 128.25$)	Wickham et al. ³¹
Screening Endometriosis	Screened negative	1 – sensitivity	Probability of screened negative when having endometriosis	0.27	Derived	Wickham et al. ³¹
Screening Endometriosis-like symptoms	Screened negative	specificity	Probability of screened negative when not having endometriosis	0.73	Beta ($\alpha = 346.75$, $\beta = 128.25$)	Wickham et al. ³¹
Screening Endometriosis-like symptoms	Screened positive	1 – specificity	Probability of screened positive when not having endometriosis	0.27	Derived	Wickham et al. ³¹
Screened positive	Doctor visit	compliance	Compliance with SC recommendation	0.7	Beta ($\alpha = 151.90$, $\beta = 65.10$)	Pairon et al. ³⁷ ; Flo internal data (2024)
Screened negative	Doctor visit	1 – compliance	Non-compliance with SC recommendation	0.3	Derived	Pairon et al. ³⁷ ; Flo internal data (2024)
Screened negative Endometriosis-like symptoms	No endometriosis	compliance	Compliance with SC recommendation	0.7	Beta ($\alpha = 151.90$, $\beta = 65.10$)	As above
Doctor visit Endometriosis	Diagnosed	doctor_sensitivity	Probability of correctly diagnosing endometriosis patients as positive	0.25	Beta ($\alpha = 54.25$, $\beta = 162.75$)	NICE ³⁴
Doctor visit Endometriosis	Endometriosis	1 – doctor_sensitivity	Complementary probability	0.75	Derived	As above
Doctor visit Endometriosis-like symptoms	No endometriosis	doctor_specificity	Probability of correctly diagnosing endometriosis-like patients as negative	0.85	Beta ($\alpha = 184.45$, $\beta = 32.55$)	NICE ³⁴
Doctor visit Endometriosis-like symptoms	Diagnosed	1 – doctor_specificity	Complementary probability	0.15	Derived	NICE ³⁴
Diagnosed	Treated	tpTreated	Probability of starting treatment after diagnosis	0.9	Beta ($\alpha = 89.1$, $\beta = 9.9$)	Clinician inputs
Treated	Relapsed	tpRelapsed	Probability of relapse after treatment	0.30	Beta ($\alpha = 29.7$, $\beta = 63.9$)	NICE ³⁴
Treated	Managed	tpManaged	Probability of being managed after treatment	0.70	Derived	As above
Doctor visit Endometriosis	Diagnosed	doctor_sensitivity	Probability of correctly diagnosing endometriosis patients as positive	0.25	Beta ($\alpha = 54.25$, $\beta = 162.75$)	NICE ³⁴
Any state	Death	natDeath [age]	Age-specific natural mortality	see S-Table 4	Constant	Life tables for US: 2021 (females) ⁷²

^aThe standard error was assumed to be 20% of the mean. Transition probabilities are converted to a 6-month cycle where applicable. The vertical bar symbol denotes conditional probability. For instance, 'Doctor visit | Endometriosis' represents the probability of visiting a doctor given that the person has endometriosis.

Table 4 | Model inputs of costs (2024 USD) and their corresponding distributions

Variable name	Description	Point estimates ^a	Distribution	Source
cSymptomatic	Incremental 6-monthly healthcare cost in the Endometriosis—undiagnosed state	\$997.02	Gamma ($\alpha = 25.00, \beta = 39.88$)	Fuldeore et al. ²⁴
cSubscription	6-monthly subscription cost of Flo SC	\$30.00	Constant	Flo Health 2024 price
cDiagnosis	Incremental cost associated with endometriosis diagnosis	\$6828.98	Gamma ($\alpha = 25.00, \beta = 273.16$)	Soliman et al. ²⁶
cTreatment	Incremental healthcare cost attributed to post-diagnosis treatment for endometriosis	\$2895.56	Gamma ($\alpha = 25.00, \beta = 115.82$)	Soliman et al. ²⁶
cTreatment_addon	Incremental 6-monthly healthcare cost in the endometriosis—treated state	\$768.93	Gamma ($\alpha = 25.00, \beta = 30.76$)	Soliman et al. ²⁶
cManaged	Incremental healthcare costs in the Endometriosis—Managed state	\$1603.14	Gamma ($\alpha = 25.00, \beta = 64.13$)	Fuldeore et al. ²⁴
cPL_treated	Incremental 6-monthly productivity loss in the Endometriosis—Diagnosed and Treated state	\$1391.65	Gamma ($\alpha = 25.00, \beta = 55.67$)	Soliman et al. ²⁶
cPL_symptomatic	Incremental 6-monthly productivity loss in all Endometriosis states before treatment	\$865.08	Gamma ($\alpha = 25.00, \beta = 34.60$)	Soliman et al. ²⁶ ; Simoens et al. ²⁷
cDoctor_visit	Cost per visit to a primary care physician	\$134.09	Gamma ($\alpha = 25.00, \beta = 5.36$)	Centers for Medicare & Medicaid Services (CMS) ⁷³ , adjusted by factor from Lopez et al. ⁷⁴
cObG_visit	Cost per visit to an obstetrician-gynecologist	\$266.21	Gamma ($\alpha = 25.00, \beta = 10.65$)	CMS ⁷³ , adjusted by factor from Lopez et al. ⁷⁴
cFalse_positive	Cost of receiving false positive result from Flo SC	\$100.00	Gamma ($\alpha = 25.00, \beta = 4.00$)	Derived, human capital approach, see S-Table 5
cDeath	Cost associated with the death state	\$0.00	Constant	Assumed

^aThe standard error was assumed to be 20% of the mean.

Table 5 | Model inputs of utilities and their corresponding distributions

Variable name	Description	Point estimates ^a	Distribution	Source
uNo_endo	Utility in the No endometriosis—diagnosed state	0.81	Beta ($\alpha = 79.87, \beta = 19.13$)	Derived from Nnoaham et al. ²¹ and Ara and Brazier ⁷⁵
uSymptomatic_noendo	Utility in the No endometriosis—undiagnosed state	0.79	Beta ($\alpha = 77.84, \beta = 21.16$)	Derived from Nnoaham et al. ²¹ and Ara and Brazier ⁷⁵
uSymptomatic	Utility in the Endometriosis—undiagnosed & untreated state	0.76	Beta ($\alpha = 75.05, \beta = 23.95$)	Derived from Wang et al. ³⁶ and Nnoaham et al. ²¹
uFirstline_treatment		0.02	Beta ($\alpha = 97.98, \beta = 4801.02$)	NICE ³⁴
uTreated_initial	Utility in the Endometriosis—empirically treated state (diagnosed or undiagnosed)	0.78	Derived	Derived from NICE ³⁴ and Wang et al. ³⁶
uTreated_further	Utility in the Endometriosis—diagnosed & formally treated state	0.68	Beta ($\alpha = 67.72, \beta = 31.28$)	Derived from Soliman et al. ²⁶ and Wang et al. ³⁶
uManaged	Utility in the Endometriosis—diagnosed & managed state	0.84	Beta ($\alpha = 83.36, \beta = 15.64$)	Wang et al. ³⁶
uNormal	Utility in the normal healthy state	By age group 18–24: 0.924 25–34: 0.912 35–44: 0.889 45–54: 0.855 55–64: 0.83	Constant	Janssen and Szende ⁶⁷
uMenopause	Utility in the menopause state	0.83	Constant	Janssen and Szende ⁶⁷
uDeath	Utility in the death state	0.00	Constant	Assumed

^aThe standard error was assumed to be 20% of the mean.

Ethics approval and consent to participate

Not applicable. This study did not involve human participants, human data, or human tissue.

Data availability

This study is an economic modelling study. All model input data were sourced from published literature or clinical expert input, with references

cited in the manuscript. The model structure, input parameters, and assumptions are fully detailed in the Methods section. Further information is available from the corresponding author upon reasonable request.

Code availability

The economic model was implemented in Microsoft Excel. R (version 4.4.1) was used for data visualization and plotting. The model code and R scripts

are available from the corresponding author upon reasonable request, subject to approval by Flo Health Ltd.

Received: 6 May 2025; Accepted: 29 December 2025;

Published online: 02 March 2026

References

- Chambers, D. et al. Digital and online symptom checkers and assessment services for urgent care to inform a new digital platform: a systematic review. *Health Serv. Deliv. Res.* **7**, 1–88 (2019).
- Fraser, H. et al. Comparison of diagnostic and triage accuracy of Ada Health and WebMD Symptom Checkers, ChatGPT, and physicians for patients in an emergency department: clinical data analysis study. *JMIR MHealth UHealth* **11**, e49995 (2023).
- Kujala, S., Hörhammer, I., Hänninen-Ervasti, R. & Heponiemi, T. Health professionals' experiences of the benefits and challenges of online symptom checkers. In *Digital Personalized Health and Medicine* (eds Pape-Haugaard, L. B. et al.) 966–970 (IOS Press, 2020).
- Australia Healthdirect. Symptom Checker. <https://www.healthdirect.gov.au/symptom-checker> (2022).
- Judson, T. J. et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. *J. Am. Med. Inform. Assoc.* **27**, 860–866 (2020).
- Riboli-Sasco, E. et al. Triage and diagnostic accuracy of online symptom checkers: systematic review. *J. Med. Internet Res.* **25**, e43803 (2023).
- Semigran, H. L., Linder, J. A., Gidengil, C. & Mehrotra, A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* **351**, h3480 (2015).
- Wallace, W. et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. *Npj Digital Med.* **5**, 1–9 (2022).
- Müller, R. et al. Ethical, legal, and social aspects of symptom checker applications: a scoping review. *Med. Health Care Philos.* **25**, 737–755 (2022).
- Preiser, C. et al. The doctors, their patients, and the symptom checker app: qualitative interview study with general practitioners in Germany. *JMIR Hum. Factors* **11**, e57360 (2024).
- Schmieding, M. L., Mörgeli, R., Schmieding, M. A. L., Feufel, M. A. & Balzer, F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. *J. Med. Internet Res.* **23**, e24475 (2021).
- Gilbert, S. et al. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. *BMJ Open* **10**, e040269 (2020).
- Gräf, M. et al. Comparison of physician and artificial intelligence-based symptom checker diagnostic accuracy. *Rheumatol. Int.* **42**, 2167–2176 (2022).
- Fraser, H., Coiera, E. & Wong, D. Safety of patient-facing digital symptom checkers. *Lancet* **392**, 2263–2264 (2018).
- Ronicke, S. et al. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. *Orphanet J. Rare Dis.* **14**, 69 (2019).
- Berry, A. C. Online symptom checker applications: syndromic surveillance for international health. *Ochsner J.* **18**, 298–299 (2018).
- Willmen, T. et al. Health economic benefits through the use of diagnostic support systems and expert knowledge. *BMC Health Serv. Res.* **21**, 947 (2021).
- Fuldeore, M. & Soliman, A. M. Prevalence and symptomatic burden of diagnosed endometriosis in the United States: national estimates from a cross-sectional survey of 59,411 women. *Gynecol. Obstet. Invest.* **82**, 453–461 (2017).
- Zondervan, K. T., Becker, C. M. & Missmer, S. A. Endometriosis. *N. Engl. J. Med.* **382**, 1244–1256 (2020).
- Nezhat, C. et al. Use of the free endometriosis risk advisor app as a non-invasive screening test for endometriosis in patients with chronic pelvic pain and/or unexplained infertility. *J. Clin. Med.* **12**, 5234 (2023).
- Nnoaham, K. E. et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. *Fertil. Steril.* **96**, 366–373.e8 (2011).
- Ballard, K., Lowton, K. & Wright, J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertil. Steril.* **86**, 1296–1301 (2006).
- De Graaff, A. A. et al. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum. Reprod.* **28**, 2677–2685 (2013).
- Fuldeore, M. et al. Healthcare utilization and costs in women diagnosed with endometriosis before and after diagnosis: a longitudinal analysis of claims databases. *Fertil. Steril.* **103**, 163–171 (2015).
- Simoens, S. et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Hum. Reprod.* **27**, 1292–1299 (2012).
- Soliman, A. M., Surrey, E., Bonafede, M., Nelson, J. K. & Castelli-Haley, J. Real-world evaluation of direct and indirect economic burden among endometriosis patients in the United States. *Adv. Ther.* **35**, 408–423 (2018).
- Simoens, S., Meuleman, C. & D'Hooghe, T. Non-health-care costs associated with endometriosis. *Hum. Reprod.* **26**, 2363–2367 (2011).
- Edgley, K., Horne, A. W., Saunders, P. T. K. & Tsanas, A. Symptom tracking in endometriosis using digital technologies: knowns, unknowns, and future prospects. *Cell Rep. Med.* **4**, 101192 (2023).
- Sirohi, D. et al. Good-quality mHealth apps for endometriosis care: systematic search. *J. Med. Internet Res.* **27**, e49654 (2025).
- Peven, K. et al. Assessment of a digital symptom checker tool's accuracy in suggesting reproductive health conditions: clinical vignettes study. *JMIR MHealth UHealth* **11**, e46718 (2023).
- Wickham, A. P. et al. Exploring self-reported symptoms for developing and evaluating digital symptom checkers for polycystic ovarian syndrome, endometriosis, and uterine fibroids: exploratory survey study. *JMIR Form. Res.* **8**, e65469 (2024).
- Arakawa, I., Momoeda, M., Osuga, Y., Ota, I. & Koga, K. Cost-effectiveness of the recommended medical intervention for the treatment of dysmenorrhea and endometriosis in Japan. *Cost Eff. Resour. Alloc.* **16**, 12 (2018).
- Grand, T. S., Basarir, H. & Jackson, L. J. The cost-effectiveness of oral contraceptives compared to 'no hormonal treatment' for endometriosis-related pain: an economic evaluation. *PLoS ONE* **14**, e0210089 (2019).
- National Institute for Health and Care Excellence (NICE). Endometriosis: diagnosis and management, Appendix K—Diagnosis and treatment model. <https://www.nice.org.uk/guidance/ng73/documents/guideline-appendix-2> (2017).
- Sanghera, S., Barton, P., Bhattacharya, S., Horne, A. W. & Roberts, T. E. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. *BMJ Open* **6**, e010580 (2016).
- Wang, S. I.-T. et al. Cost-effectiveness of elagolix versus leuprolide acetate for treating moderate-to-severe endometriosis pain in the USA. *J. Comp. Eff. Res.* **8**, 337–355 (2019).
- Pairon, A., Philips, H. & Verhoeven, V. A scoping review on the use and usefulness of online symptom checkers and triage systems: how to proceed?. *Front. Med.* **9**, 1040926 (2023).
- D'Hooghe, T., Dirksen, C. D., Dunselman, G. A. J., De Graaff, A. & Simoens, S. The costs of endometriosis: it's the economy, stupid. *Fertil. Steril.* **98**, S218–S219 (2012).
- Borges do Nascimento, I. J. et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. *Npj Digital Med.* **6**, 1–28 (2023).

40. Morse, K. E., Ostberg, N. P., Jones, V. G. & Chan, A. S. Use characteristics and triage acuity of a digital symptom checker in a large integrated health system: population-based descriptive study. *J. Med. Internet Res.* **22**, e20549 (2020).
41. Mardon, A. K. et al. The efficacy of self-management strategies for females with endometriosis: a systematic review. *Reprod. Sci.* **30**, 390–407 (2023).
42. O'Hara, R., Rowe, H. & Fisher, J. Self-management in condition-specific health: a systematic review of the evidence among women diagnosed with endometriosis. *BMC Womens Health* **19**, 80 (2019).
43. Assaf, A. R. et al. The relative burden of menopausal and postmenopausal symptoms versus other major conditions: a retrospective analysis of the medical expenditure panel survey data. *Am. Health Drug Benefits* **10**, 311–321 (2017).
44. Cunningham, A. C. et al. Perimenopause symptoms, severity, and healthcare seeking in women in the US. *Npj Womens Health* **3**, 1–8 (2025).
45. Faubion, S. S. et al. Impact of menopause symptoms on women in the workplace. *Mayo Clin. Proc.* **98**, 833–845 (2023).
46. Saavolainen, L. et al. Mortality of midlife women with surgically verified endometriosis—a cohort study including 2.5 million person-years of observation. *Hum. Reprod.* **34**, 1576–1586 (2019).
47. D'Alterio, M. N. et al. Medical and surgical interventions to improve the quality of life for endometriosis patients: a systematic review. *Gynecol. Surg.* **18**, 13 (2021).
48. As-Sanie, S. et al. Endometriosis: a review. *JAMA* **334**, 64–78 (2025).
49. Vercellini, P. et al. The effect of surgery for symptomatic endometriosis: the other side of the story. *Hum. Reprod. Update* **15**, 177–188 (2009).
50. Becker, C. M., Gattrell, W. T., Gude, K. & Singh, S. S. Reevaluating response and failure of medical treatment of endometriosis: a systematic review. *Fertil. Steril.* **108**, 125–136 (2017).
51. Coney, A. et al. Accuracy of online symptom checkers and the potential impact on service utilisation. *PLoS ONE* **16**, e0254088 (2021).
52. McDuff, D. et al. Towards accurate differential diagnosis with large language models. *Nature* 1–7 <https://doi.org/10.1038/s41586-025-08869-4> (2025).
53. Tu, T. et al. Towards conversational diagnostic artificial intelligence. *Nature* 1–9 <https://doi.org/10.1038/s41586-025-08866-7> (2025).
54. Mahlknecht, A., Engl, A., Piccoliori, G. & Wiedermann, C. J. Supporting primary care through symptom checking artificial intelligence: a study of patient and physician attitudes in Italian general practice. *BMC Prim. Care* **24**, 174 (2023).
55. Weik, L., Fehring, L., Mortsiefer, A. & Meister, S. Understanding inherent influencing factors to digital health adoption in general practices through a mixed-methods analysis. *Npj Digital Med.* **7**, 1–16 (2024).
56. Pagoto, S. & Bennett, G. G. How behavioral science can advance digital health. *Transl. Behav. Med.* **3**, 271–276 (2013).
57. Husereau, D. et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *BMC Med.* **20**, 23 (2022).
58. Christ, J. P. et al. Incidence, prevalence, and trends in endometriosis diagnosis: a United States population-based study from 2006 to 2015. *Am. J. Obstet. Gynecol.* **225**, 500.e1–500.e9 (2021).
59. Caro, J. J., Briggs, A. H., Siebert, U. & Kuntz, K. M. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Value Health* **15**, 796–803 (2012).
60. Siebert, U. et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value Health* **15**, 812–820 (2012).
61. Gomes, M., Murray, E. & Raftery, J. Economic evaluation of digital health interventions: methodological issues and recommendations for practice. *Pharmacoeconomics* **40**, 367–378 (2022).
62. Haacker, M., Hallett, T. B. & Atun, R. On discount rates for economic evaluations in global health. *Health Policy Plan* **35**, 107–114 (2020).
63. NICE. Endometriosis: diagnosis and management, Appendix G—Evidence tables. <https://www.nice.org.uk/guidance/ng73/documents/guideline-appendix> (2017).
64. U.S. Bureau of Labor Statistics. Consumer Price Index, 1913, Federal Reserve Bank of Minneapolis. <https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1913> (2025).
65. Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L. & Torrance, G. W. *Methods for the Economic Evaluation of Health Care Programmes* (Oxford University Press, 2015).
66. McCarus, S. et al. Treatment of moderate-severe pain associated with early stage endometriosis; a consensus panel summary: endometriosis pain treatment. *Sci. Prepr.* <https://doi.org/10.14293/PR2199.000538.v2> (2024).
67. Janssen, B. & Szende, A. Population norms for the EQ-5D. In *Self-Reported Population Health: An International Perspective based on EQ-5D* (eds Szende, A., Janssen, B. & Cabases, J.) (Springer, 2014).
68. Pichon-Riviere, A., Drummond, M., Palacios, A., Garcia-Marti, S. & Augustovski, F. Determining the efficiency path to universal health coverage: cost-effectiveness thresholds for 174 countries based on growth in life expectancy and health expenditures. *Lancet Glob. Health* **11**, e833–e842 (2023).
69. Hadfield, R., Mardon, H., Barlow, D. & Kennedy, S. Delay in the diagnosis of endometriosis: a survey of women from the USA and the UK. *Hum. Reprod.* **11**, 878–880 (1996).
70. Zondervan, K. T. et al. Chronic pelvic pain in the community—symptoms, investigations, and diagnoses. *Am. J. Obstet. Gynecol.* **184**, 1149–1155 (2001).
71. You, Y., Ma, R. & Gui, X. User experience of symptom checkers: a systematic review. *AMIA. Annu. Symp. Proc.* **2022**, 1198–1207 (2023).
72. Centres for Disease Control and Prevention. *Life Tables for US: 2021*. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/72-12/ (2023).
73. Centers for Medicare & Medicaid Services (CMS). *Medicare Physician Fee Schedule Look-Up Tool: HCPCS Code 99213*. <https://www.cms.gov/medicare/physician-fee-schedule/search> (2024).
74. Lopez, E., Neuman, T., Jacobson, G. & Levitt, L. How much more than medicare do private insurers pay? A review of the literature. *KFF* <https://www.kff.org/medicare/issue-brief/how-much-more-than-medicare-do-private-insurers-pay-a-review-of-the-literature/> (2020).
75. Ara, R. & Brazier, J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). *Value Health J. Int. Soc. Pharmacoeconomics Outcomes Res.* **11**, 1131–1143 (2008).

Acknowledgements

The study was funded by Flo Health Ltd. We thank Dr. Erica Stockwell (AdventHealth, Florida, US) and Dr. Sarina B. Schragger (University of Wisconsin–Madison, US) for their clinical validation of the model assumptions and pathways. Their expert input significantly strengthened the medical accuracy and real-world relevance of this study. We also thank Kym Jacks-Bryant (former Director of Medical Affairs and Partnerships, Flo Health) for her support in coordinating clinician engagement.

Author contributions

Y.X., L.Z., and A.M. conceived the study. Y.X. conceptualized the model and conducted the cost-effectiveness analyses. C.P. and Y.X. performed the literature review to inform model parameters. S.T.-R. and J.M. validated the model structure and assumptions and provided feedback on parameter

selection. A.W. provided efficacy data on the intervention and user compliance to inform model assumptions. H.S. contributed product insights and supporting data. A.M. and A.K. reviewed the clinical pathways. L.K. performed independent model technical validation. Y.X. interpreted the results and prepared the first draft. L.Z. and S.T.-R. critically revised the manuscript. All authors contributed to the interpretation of the findings, reviewed, and approved the final manuscript.

Competing interests

Y.X., A.M., H.S., A.W., and A.K. are employees of Flo Health. C.P. and L.Z. are contractors for Flo Health. S.T.-R. and J.M. are external consultants for Flo Health. A.M., A.W., H.S., L.Z., and A.K. hold equity interests in Flo Health.

Additional information

Supplementary information The online version contains supplementary material available at

<https://doi.org/10.1038/s41746-025-02332-4>.

Correspondence and requests for materials should be addressed to Yihan Xu.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026