TBM

ORIGINAL RESEARCH

Predicting healthcare-seeking behavior based on stated readiness to act: development and validation of a prediction model

Eric P. Green,[1,2] Shyam Pradheep,[1] Jessica Heinzelman,[2] Anne Nyanchoka,[3] Daphine Achieng,[3] Siddhartha Goyal,[2] Laura Cusson,[2] A. Solomon Kurz,[4] Benjamin Bellows[2]

[1]Duke Global Health Institute, Durham, NC 27708, USA
[2]Nivi, Inc., Boston, MA 02115, USA
[3]Nivi, Inc., Nairobi, Kenya
[4]VISN 17 Center of Excellence for Research on Returning War Veterans, Central Texas Veterans Health Care System, Waco, TX 76711, USA

Correspondence to: EP Green, eric.green@duke.edu

Cite this as: *TBM* 2021;XX:XX–XX
https://doi.org/10.1093/tbm/ibab096

Abstract

A starting point of many digital health interventions informed by the Stages of Change Model of behavior change is assessing a person's readiness to change. In this paper, we use the concept of readiness to develop and validate a prediction model of health-seeking behavior in the context of family planning. We conducted a secondary analysis of routinely collected, anonymized health data submitted by 4,088 female users of a free health chatbot in Kenya. We developed a prediction model of (future) self-reported action by randomly splitting the data into training and test data sets (80/20, stratified by the outcome). We further split the training data into 10 folds for cross-validating the hyperparameter tuning step in model selection. We fit nine different classification models and selected the model that maximized the area under the receiver operator curve. We then fit the selected model to the full training dataset and evaluated the performance of this model on the holdout test data. The model predicted who will visit a family planning provider in the future with high precision (0.93) and moderate recall (0.75). Using the Stages of Change framework, we concluded that 29% of women were in the "Preparation" stage, 21% were in the "Contemplation" stage, and 50% were in the "Pre-Contemplation" stage. We demonstrated that it is possible to accurately predict future healthcare-seeking behavior based on information learned during the initial encounter. Models like this may help intervention developers to tailor strategies and content in real-time.

Keywords

Chatbot, Family planning, Prediction model, Stage of change, Digital health

Implications

Practice: Intervention designers should consider how to assess readiness to act and study how tailoring impacts outcomes.

Policy: Behavior change is a process, and our policies and programs should reflect the reality that most people are not ready to act at the time of their first encounter with the health system.

Research: Future research should validate this model with new cohorts and evaluate the impact of intervention tailoring on health behavior and outcomes.

With half of the world's population unable to access essential health services, there is a growing recognition of the importance of self-care, a people-centered approach that complements the traditional provider-based model and empowers individuals to take an active role in their own health [1]. Self-care interventions include "drugs, devices, and diagnostics" that people can use with or without the help of health care providers [2]. The proliferation of digital tools, such as smartphones and wearables, along with expanding access to the internet and advances in artificial intelligence, is creating new ways for people to engage in self-care from the privacy and convenience of their homes [3]. COVID-19 has only amplified the importance of creating remote options for accessing care.

One class of digital health interventions that can scale self-care to millions, at least in principle, is the automated conversational agent, or chatbot. Chatbots work by automatically engaging users in conversations about health topics. Popular channels of delivery include standalone apps, messaging platforms (e.g., SMS, Facebook Messenger, and WhatsApp), websites, interactive voice response, and smart speaker virtual assistants such as Amazon Alexa and Google Assistant. There are hundreds of chatbot interventions that specialize in self-care health journeys related to mental health, sexual and reproductive health, weight loss, and other topics [4–6]. Most chatbots are rule-based with conversations designed as decision trees, but increasingly conversation designers are turning to machine learning to integrate natural language processing to identify user intent and serve appropriate content.

As the digital self-care market evolves, there is an opportunity to integrate behavior change theory,

Published online: 20 July 2021

evidence-based behavior strategies, and clinical guidelines into conversational agents [7, 8]. The public health and business case for this is promising: interventions based on social and behavioral science theories appear to be more efficacious than atheoretical approaches [9], though the evidence is mixed [10, 11]. In the broader field of digital health, commonly applied theoretical models include the Health Belief Model, Theory of Planned Behavior, Social Cognitive Theory, the Transtheoretical Model, and Self-Determination Theory [12].

The Transtheoretical Model, also known as the Stages of Change Model [13, 14], is especially relevant for digital self-care interventions that emphasize the importance of repeated engagement over time. This model posits that behavior change is a process, and individuals can move through several stages: (a) *Pre-Contemplation*, no intention to act; (b) *Contemplation*, awareness of the need to act but no firm commitment to action; (c) *Preparation/Determination*, preparing to follow through on intention to act in the near future; (d) *Action*, acting on intentions; (e) *Maintenance/Relapse*, sustaining behavior change; and (f) *Termination*, a phase characterized by no risk of relapse. Knowing a person's stage of change can help to tailor intervention content [15], and clinical techniques such as Motivational Interviewing are commonly used to help people move through these stages [16].

A starting point of many interventions informed by the Stages of Change Model is assessing a person's stage—their readiness to change. Often this construct is quantified by asking a person to draw a vertical line that intersects a visual analog scale at the point that represents their readiness to change, or by asking the person to rate their readiness on a scale of 0 (e.g., no thought of changing) to 10 (e.g., taking initial steps toward change). While conceptually face valid, researchers have questioned whether self-reported readiness to change predicts *actual* change [17]. Typically, the benchmark for "prediction" in studies of readiness to change is a statistically significant correlation between an indicator of readiness and a clinical outcome, regardless of the magnitude of the association. In addition to confusing statistical significance for practical or clinical significance [18], this practice also stops short of actual prediction of future behavior, conflating explanation with prediction [19].

Inspired by the concept of readiness to change but aware of these limitations, our objective was to develop a (prognostic) prediction model of action for applied use. Our use case was a digital self-care intervention called askNivi. This free and automated service aims to help people learn about family planning, identify suitable methods of contraception based on their goals, and find nearby providers. In this paper, we describe how we used self-reported data on readiness to act, along with other self-reported characteristics and engagement data, to predict who would later report visiting a family planning provider.

METHODS

We conducted a secondary analysis of routinely collected, anonymized health data submitted by askNivi users in Kenya (nationwide) from January 2019 through May 2020.

Intervention and data

askNivi is a free service that enables anyone with a mobile phone to ask questions about their health and get information and recommendations through an automated, text-based helper named Nivi. The service is marketed through social media, print media, and face-to-face advertising. The following sections describe how askNivi functioned during the analysis window. Data collection was automated; thus, ascertainment of predictors and the outcome was blinded.

Model inputs (predictors)

Onboarding. People began their engagement with askNivi by asking a question via a dedicated SMS shortcode or Facebook Messenger, by sending an advertised keyword to either channel or by tapping on a Facebook ad (which launches the Facebook Messenger app). Users sent and received messages at no cost through both channels.

After someone accepted the askNivi terms and conditions, Nivi asked them to provide their age, sex, and location to enable referrals. During the period covered by this analysis, Nivi determined a user's location by asking the person to submit the free text and used Google's Geocoding API to map each entry to a set of coordinates. For this analysis, we used these coordinates to identify the user's constituency (a local voting unit in Kenya) and overlaid a high-resolution population raster dataset [20] to classify user locations by density. Within each constituency, we calculated the mean population density value and then divided constituencies into deciles based on mean density. We classified a user as living in a "high density" constituency if their location was mapped to one of the top three highest population density deciles.

Intent classification. After a user answered the demographic questions, Nivi prompted them to ask a question (if they had not done so yet), attempted to automatically classify their intent, and routed them to the best-automated conversation. When the person's intent was not clear or there was not a matching conversation module in the askNivi library, Nivi put them into a queue to chat with a live human agent who could route them to the best automated conversation. For instance, a user might have asked, "What is the best method of family planning?," and the service would have classified the user's intent as wanting a method recommendation. A user who asked a question like this would be automatically routed to Nivi's family planning screening module. People bypassed the intent detection step if they engaged with a marketing campaign that was connected directly to a specific conversation. When this happened, the user's intent was defined by the marketing campaign. For instance, if a user clicked on an advertisement about methods of family planning, Nivi inferred that their intent was to get a method recommendation.

Family planning screening module. People interested in family planning topics could complete an automated screening to find out which methods of contraception fit their personal goals and preferences. As part of this screening, Nivi collected user responses to questions about marital status, pregnancy status, current and former contraceptive use, and method preferences.

Readiness to act. After issuing method recommendations, Nivi asked users two questions: (a) "How important to you is preventing pregnancy?" and (b) "How ready are you to visit a facility for family planning in the next 2 weeks?" Users responded to the *importance* question on an ordinal scale: "Not important," "Less important," "Useful but not a primary goal," "Important," or "Very important." Users also responded to the *readiness* question on an ordinal scale: "Not ready at all," "Not sure," "Ready," "Very ready," or "Extremely ready."

Outcome (action): visiting a family planning provider

Within 2 weeks of someone completing a family planning screening, Nivi sent them up to two automated check-in messages asking if they had visited a family planning provider. Users could reply (still at no cost to them) and indicate "yes" they had or "no" they had not. If a user reported a visit, Nivi asked them details about the visit such as the name of the health facility and whether or not they adopted a method. If they indicated they had not yet visited a provider, Nivi asked them to identify the main barrier. Among the standard response options was an option to say that they still planned to go. For this analysis, we assumed that no reply (69%, 2,801/4,088) means that the user has not visited a provider.

Analysis cohort

For this secondary analysis, we queried the askNivi Kenya database and created a cohort of users who met the following criteria: (a) female; (b) 18 to 49 years old at onboarding; (c) not currently pregnant; (d) engaged with askNivi for the first time between January 1, 2019 and May 30, 2020; (e) completed the automated family planning screening; and (f) were sent a check-in prompt within 2 weeks of completing the screening. To use the service, women needed access to a mobile phone and had to be able to read and write in English or Swahili. This includes most women in Kenya. According to the 2014 Kenya DHS, 88% of women ages 15 to 49 are literate, and 86% of households own a mobile phone [21]. A more recent survey of adults estimates that 82% of women themselves own a mobile phone [22].

Empirical approach

We used R version 4.0.2 for all analyses [23]. We began by describing the analysis cohort, women's self-reported readiness to visit a family planning provider, and their perceptions about the importance of preventing pregnancy. Then we used the {brms} package (version 2.13.5) to fit two Bayesian models (see the Appendix for details about our choice of priors) [24]. The first model was a cumulative ordinal regression [25] of women's stated readiness to act. The objective was to explore the correlates of stated readiness to act. The second model was a logistic regression of action that modeled stated readiness, an ordinal variable, as a monotonic effect [26] and adjusted for age based on a causal directed acyclic graph (see Appendix Fig. A3) [27] that identified age as a potential confounding variable. The objective was to estimate the effect of stated readiness on actions taken within the next two weeks.

We then developed and evaluated a prediction model of (future) self-reported action using the {tidymodels} suite of machine learning packages [28]. Continuous predictors were centered, rescaled, and checked for large absolute correlations. Nominal predictors were converted into binary dummy variables for each level (one hot encoding). We included a zero variance filter to ensure that no input variables contained only a single value.

We split the data into training and test data sets (80/20) and further split the training data into 10 folds for cross-validating the hyperparameter tuning step in model selection. We fit nine different classification models and selected the model that maximized the area under the receiver operator curve. We then fit the selected model to the full training dataset and evaluated the performance of this model on the holdout test data. Classes were predicted with the default threshold of 0.5. Reporting followed the TRIPOD guidelines [29].

Ethical review

The Duke University Institutional Review Board reviewed and approved a study protocol to conduct this secondary data analysis of anonymized data.

RESULTS

Characteristics of the analysis cohort

Table 1 summarizes the key characteristics of the analysis cohort (N = 4,088). Just over half of women in the cohort said they were not married or living with a partner, approximately 1 out of 3 have children and live in a high-density constituency, and the average age is 22.5 years. In terms of family planning, two-thirds of women said they are currently using a method (36%) or used a method in the past but discontinued (27%), and most said they wanted to delay or prevent pregnancy for more than one year (75%).

Descriptive exploration of self-reported readiness to act

Figure 1 displays the distributions of responses to Nivi's *importance* and *readiness* prompts according to the characteristics of users. When asked, "How important to you is preventing pregnancy?," 73% of users (2,991/4,088) said it was "important" or "very important," with little variation by measured characteristics. One exception: compared to women seeking short-term protection, a higher percentage of women who said they wanted to prevent pregnancy for at least one year also indicated that preventing pregnancy was important (77% vs. 61%).

When asked, "How ready are you to visit a facility for FP in the next two weeks?," 48% of users (1,962/4,088) reported being ready to take action. Readiness appears to vary somewhat by age, channel, contraceptive history, and family planning preferences. For instance, compared to adolescents, a higher percentage of women over age 30 said they were ready to visit a provider (64% vs. 34%). Readiness was also more common among Facebook Messenger users (54% vs. 43% among SMS users), women currently using contraception (58% vs. 47% among never users), and women wanting to prevent pregnancy for at least 1 year (52% vs. 35% among women seeking shorter-term protection).

In Fig. 2, we cross-tabulate self-reported importance and readiness to understand how these perceptions interact among users. Among the 2,991 users who said preventing pregnancy is "important" or "very important," only 54% also signaled some degree of readiness to visit a provider. Thus, almost half of women who said preventing pregnancy was an important goal were not yet ready to take action. Motivation to change often preceded stated readiness to change.

Table 1 | Characteristics of analysis cohort

Characteristics at initial app encounter	Total (N = 4,088)	Later reported visiting a FP provider: No (n = 3,683)	Later reported visiting a FP provider: Yes (n = 405)
Mean age (SD)	22.5 (4.0)	22.5 (4.0)	22.7 (3.6)
Married or living with partner: yes	1,836 (44.9%)	1,621 (44%)	215 (53.1%)
Has children: yes	1,355 (33.1%)	1,170 (31.8%)	185 (45.7%)
Lives in high density constituency (top 3 decile): yes	1,510 (36.9%)	1,376 (37.4%)	134 (33.1%)
Channel SMS or Messenger: SMS	2,324 (56.8%)	2,057 (55.9%)	267 (65.9%)
Mean number of askNivi conversation modules engaged (SD)	2.9 (1.0)	2.9 (1.0)	3.3 (1.3)
Use of FP			
Never used	1,533 (37.5%)	1,425 (38.7%)	108 (26.7%)
Currently using	1,452 (35.5%)	1,259 (34.2%)	193 (47.7%)
Not currently using, but used in the past	1,103 (27%)	999 (27.1%)	104 (25.7%)
Expressed FP-related intent at onboarding: yes	834 (20.4%)	743 (20.2%)	91 (22.5%)
Wants to delay/prevent pregnancy for >1 year: yes	3,084 (75.4%)	2,759 (74.9%)	325 (80.2%)
Perceived importance of preventing pregnancy			
Not important	279 (6.8%)	252 (6.8%)	27 (6.7%)
Less important	211 (5.2%)	197 (5.3%)	14 (3.5%)
Useful but not a primary goal	607 (14.8%)	565 (15.3%)	42 (10.4%)
Important	1,026 (25.1%)	922 (25%)	104 (25.7%)
Very important	1,965 (48.1%)	1,747 (47.4%)	218 (53.8%)
Stated readiness to visit provider (next 2 weeks)			
Not ready at all	797 (19.5%)	764 (20.7%)	33 (8.1%)
Not sure	1,329 (32.5%)	1,243 (33.7%)	86 (21.2%)
Ready	1,055 (25.8%)	932 (25.3%)	123 (30.4%)
Very ready	416 (10.2%)	339 (9.2%)	77 (19%)
Extremely ready	491 (12%)	405 (11%)	86 (21.2%)

FP family planning.

Fig 1 | (A) Importance rating and (B) readiness rating by user characteristics. "Unmet need for FP" is constrained to "important" or "very important" responses by definition since the importance variable is part of the unmet need construction. *FP* family planning.

Correlates of self-reported readiness to act

To further explore the correlates of women's stated readiness to act, we fit a Bayesian ordinal regression model (cumulative). Figure 3 displays the Markov chain Monte Carlo draws from the posterior distribution of the parameters. Holding constant all other variables, women reported greater stated readiness to act (standard deviations on the latent readiness scale) if they said preventing pregnancy was important to them, had children, said they wanted to prevent pregnancy for more than 1 year, expressed an intent related to family planning, contacted askNivi via Facebook Messenger (vs. SMS), were older, were married or living with a partner, or engaged more deeply with askNivi conversation modules. The perceived importance of preventing pregnancy is the strongest correlate of stated readiness.

Association between self-reported readiness to act and visiting a provider

With a better understanding of which measured characteristics of users are linked to greater self-reported readiness to act, we explored the association between stated readiness and a future action: self-reported visits to a family planning provider. We fit a Bayesian linear regression model, modeling the ordered categorical variable stated readiness as a monotonic effect and adjusting for age. The decision to adjust for age, and only age, was made through the use of a causal directed acyclic graph (see Appendix Fig. A3). Figure 4 displays the conditional effects of stated readiness on the probability of visiting a family planning provider. On average, the predicted probability of visiting a provider increases 3.5% points per increase in one stated readiness category. However, stated readiness alone is not sufficient for accurately predicting this future action. Only 15% of women who said they were ready to act later reported visiting a family planning provider.

Predicting who will take action

To improve our ability to predict who will go on to take action, we incorporated additional information women shared during their initial engagement

Fig 2 | Distribution of readiness to visit family planning provider by perceived importance of preventing pregnancy (N = 4,088).

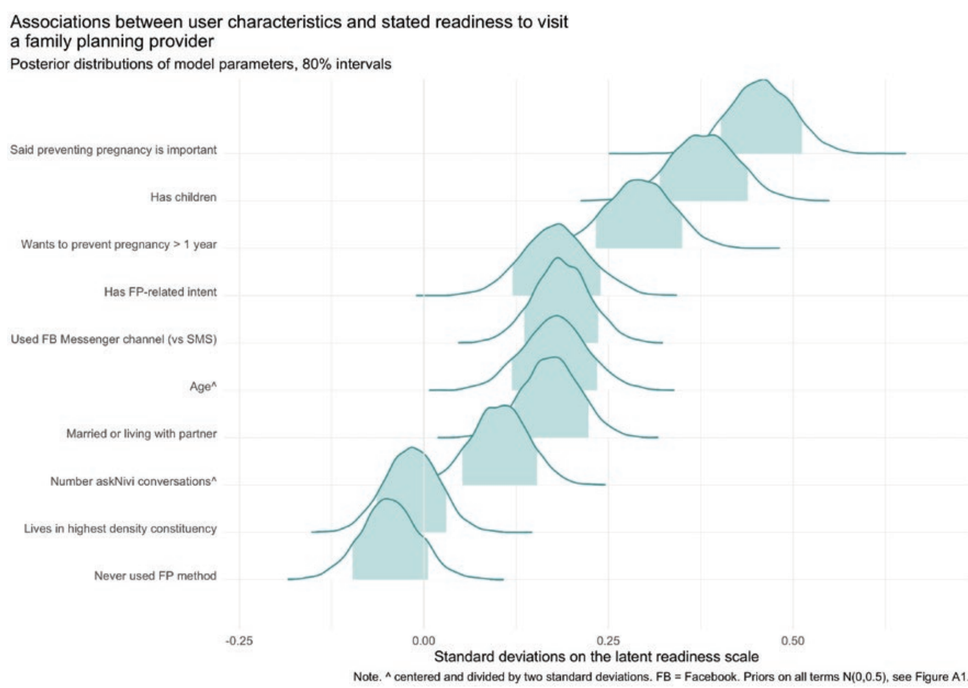

Fig 3 | Results of a Bayesian ordinal regression model (cumulative) of readiness to act (N = 3,329). Plot shows Markov chain Monte Carlo draws from the posterior distribution of the parameters.

with askNivi into a prediction model. The model included 10 features: stated readiness plus the covariates shown in Fig. 3 (excluding density).

We split the data (N = 4,088) into training (N = 3,271, 80%) and test (N = 817, 20%) sets (see Appendix Table A1) and used 10-fold cross-validation on the training set to tune model hyperparameters and fit nine different classification models. Given the class imbalance (12% later reported visiting a provider), we up-sampled the training data based on the outcome (the ratio of the majority-to-minority frequencies was set empirically through cross-validation on the training data). We then selected the model (and the model hyperparameters) with the best overall area under the receiver operator curve in the cross-validated

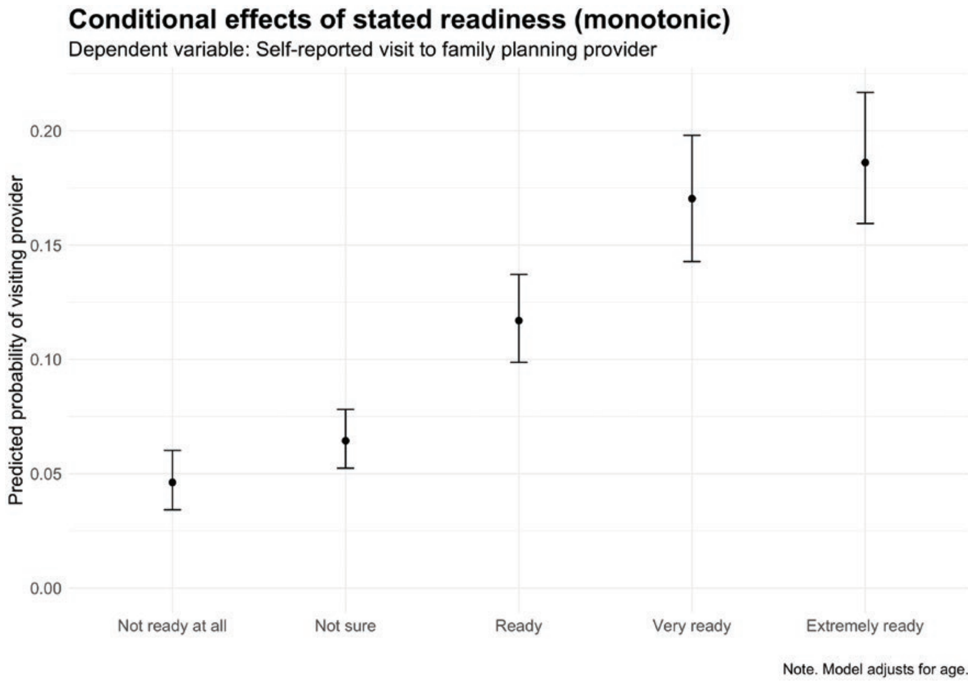

Fig 4 | Results of a Bayesian logistic regression model of self-reported visits to family planning providers (N = 4,088). Plot shows conditional effects of stated readiness (monotonic) on visits, adjusted for age.

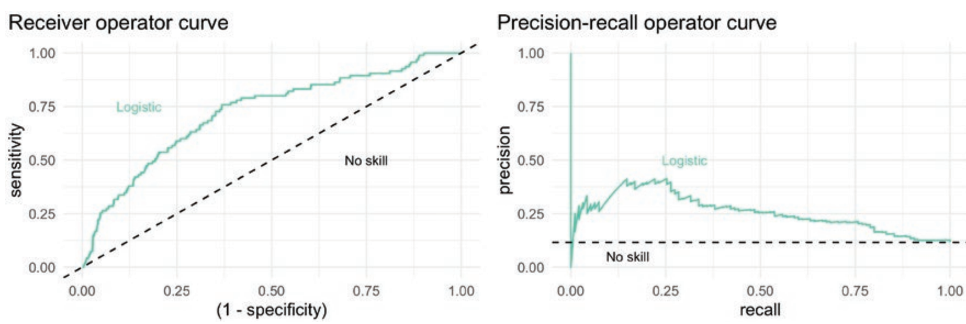

Fig 5 | Performance of the trained model of action on the unseen test data (N = 817).

training data, fit the best model (logistic regression) to the full training dataset, and evaluated performance on the holdout (unseen) test data.

Figure 5 shows the receiver operator curve and the precision-recall curve for the final model fit to the test data. Both curves are compared to a no-skill classifier and show the model's relative advantage. Given the class imbalance, the precision-recall curve is preferred for its focus on predicting the minority class. The area under the precision recall curve is 0.25 (compared to a baseline of 0.12). The area under the receiver operator curve is 0.72. The model has high precision of 0.93 and moderates 0.75 recall, with an F1 score of 0.83. This means 93% of women classified by the model as future action takers were correctly classified (precision, positive predictive value), and 75% of true future action takers were correctly classified (recall, sensitivity). To assess mean calibration (calibration-in-the-large), we compared the average predicted probability of action with the event rate in the validation dataset. The prevalence of visiting a family planning provider was 12%, whereas the average predicted event given by the final model was 29%. This means that the model overestimates action in general.

Identifying the stage of change

Next, we used the model's predictions of future action to aid in classifying women in the test dataset (N = 817) according to their stage of change (see Fig. 6). If the model predicted a woman was going to visit a provider (based on her data from her initial encounter with askNivi), we classified her as being in the "Preparation" stage (N = 233, 29%). Women predicted not to act were classified as being in the "Contemplation" stage if they stated they were ready to visit a provider or if they said preventing pregnancy was important (N = 172, 21%). All other women were classified as being in the "Pre-Contemplation" stage (N = 412, 50%).

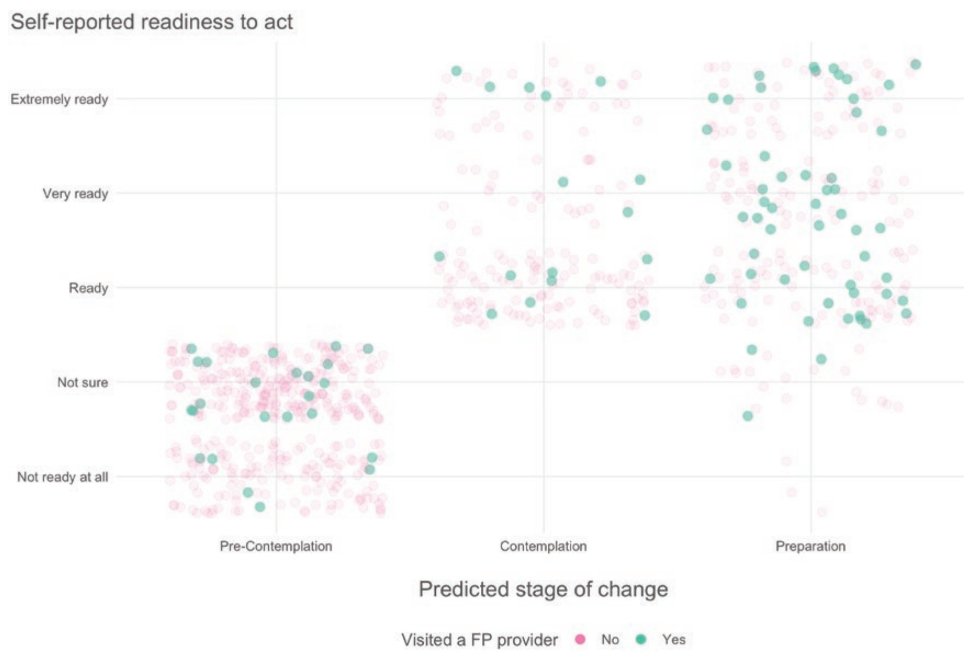

Fig 6 | Predicted stage of change by stated readiness to act. Test dataset, $N = 817$.

This approach used only data available at the initial engagement to classify a woman's stage of change.

DISCUSSION

Using data from a sample of 4,088 Kenyan women who completed a family planning screening through the automated askNivi conversational agent, we developed and internally validated a prediction (prognostic) model of healthcare seeking. The inputs to this model are several individual characteristics ascertained at the time of the digital screening, including women's stated readiness to act. Applied to unseen test data, the model predicted who will visit a family planning provider in the future with high precision (0.93) and moderate recall (0.75).

In this paper, we also demonstrate how model predictions can be framed in the Transtheoretical (or Stages of Change) Model of behavior change. We classified women who were predicted to act to be in the "Preparation" stage. For women not predicted to act, we labeled their stage of change ("Contemplation" or "Pre-Contemplation") based on their self-reported data on readiness to act and perceived importance of acting. Applied to the test data, we concluded that 29% of women were in the "Preparation" stage, 21% were in the "Contemplation" stage, and 50% were in the "Pre-Contemplation" stage.

This reinforces a takeaway offered by Norcross and colleagues from their original 2011 meta-analysis: Most patients are not ready to act *today*, and our interventions should reflect this reality [14]. There is evidence suggesting that tailoring interventions to someone's stage of readiness is effective [15, 30, 31], and we believe prediction models like the one we describe in this paper can provide an empirically derived guide for such tailoring. This is a key hypothesis to test in future research [4].

Like many other studies of behavioral intention, we also show that a person's stated readiness is correlated with future outcomes. It is common in the readiness to change literature for authors to describe a statistically significant correlation as evidence that readiness "predicts" future outcomes. For instance, a 2018 meta-analysis of 76 psychotherapy studies involving more than 25,000 patients aimed to "assess the ability of stages of change and related readiness measures to *predict* psychotherapy outcomes" [emphasis added] [32]. That work updated a prior meta-analysis by the same authors [14]. In both analyses, the authors quantified prediction as a standardized effect size (Cohen's *d*) measuring the strength of the association between readiness and outcomes: 0.46 (95% CI 0.35–0.58) in the first meta-analysis and 0.41 (95% CI 0.34–0.48) in the update. (For comparison, we fit a linear model of action with readiness and age and converted the adjusted coefficient for readiness to a Cohen's *d* value of 0.30 with a 95% confidence interval of 0.24 to 0.36.)

Based on these results, the authors claimed that "client stages of change *reliably predicted* psychotherapy outcomes" and "stages of change are robustly associated with and *predictive of* outcomes in psychotherapy" [emphasis added] [32]. This use of the term "predict" is common in the social and behavioral sciences, but Shmueli argues, and we concur, that it conflates causal explanation with empirical prediction [19]. Altman and Royston make a related point in their paper on validating prognostic models: predicting outcomes for groups of patients is not the same as predicting outcomes for individuals [33]. As they state, "Usefulness is determined by how well a model works in practice, not by how many zeros there are in the associated P-values." We make this distinction to highlight what we see as a contribution of our paper: the development and validation of an individual prediction model that incorporates self-reported readiness to act. We are not aware of other papers that use readiness to act in this type of prognostic modeling.

LIMITATIONS

While we provide evidence for internal validation, this analysis lacks temporal and external validation [33]. That is to say, we tested model performance on an unseen holdout sample from the analysis cohort (internal), but we did not test the model on data from a cohort from a different time or setting. Future work will be needed to externally validate the model in different settings. Additionally, misclassification of the outcome (i.e., visiting a family planning provider) was possible. This outcome was self-reported, and we assumed that women who did not respond to our post-referral prompts did not visit a provider. We believe this is a reasonable choice because our prior work has shown that most people do not take action in the short term, and users who have not acted have little incentive to respond. To the extent that this assumption is incorrect, model performance will suffer in practice, and users should be less engaged and report lower satisfaction with app experiences informed by model predictions.

CONCLUSIONS

We demonstrated that it is possible to predict future healthcare-seeking behavior based on information learned during a user's initial interaction with an automated conversational agent on a mobile phone. Prediction models like the one we developed could be applied to digital health applications to help tailor health communication strategies and content in real-time, even in low-resource settings.

According to Kreuter, Strecher, and Glassman [34], tailoring uses unique characteristics of an individual obtained during an assessment to customize communication. The assumed mechanism of tailoring is that it makes health messages more personally relevant to a user [35], and based on the elaboration likelihood model of persuasion [36], users who perceive messages to be more relevant may be more likely to engage.

For instance, using the askNivi example, prediction models could be used to tailor follow-up message content and timing to encourage users to visit a family planning provider. Women predicted to take action in the short term could receive messages aimed at overcoming access barriers and cognitive biases that often lead people to delay or abandon health goals like preventing pregnancy. Conversely, women who are not predicted to act—despite their stated desire to avoid pregnancy—could receive messages intended to persuade them about the benefits of contraception and the chances of unwanted or unintended pregnancy without contraception.

Meta-analyses of intervention tailoring have consistently reported small positive effects of tailoring on a variety of health outcomes, from smoking cessation to breast cancer screening [for a review of reviews, see 35]. For digital health applications where data collection and tailoring can be automated, small effects at scale are potentially very cost-effective means of improving individual and population health. It remains an open question, however, how to most effectively tailor communication and interventions for maximum effect. Prediction models that incorporate information on a person's readiness to act merit further investigation.

Acknowledgment: There was no specific funding to support this secondary analysis.

Compliance with Ethical Standards

Conflicts of Interest: All Nivi-affiliated authors disclose a financial conflict of interest as employees or owners of Nivi, Inc., the company behind the askNivi service that generated the data for this analysis. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Veterans Affairs.

Human Rights: The Duke University Institutional Review Board reviewed and approved a study protocol to conduct this secondary data analysis of anonymized data.

Informed Consent: This was a retrospective analysis of anonymized program data. The askNivi Terms of Service informed potential users that information they provide may be used for research and analytics purposes.

Transparency Statements:

1. This study was not formally registered.
2. The analysis plan was not formally pre-registered.
3. De-identified data, analytic code, and materials from this study are available in a public archive: https://github.com/ericpgreen/readiness-2020

APPENDIX

Table A1 | Characteristics of the development and validation data

Characteristics at initial app encounter	Total (*N* = 4,088)	Data set: Development (*n* = 3,271)	Data set: Validation (*n* = 817)
Mean age (SD)	22.5 (4.0)	22.5 (4.0)	22.4 (3.9)
Married or living with partner: yes	1,836 (44.9%)	1,465 (44.8%)	371 (45.4%)
Has children: yes	1,355 (33.1%)	1,098 (33.6%)	257 (31.5%)
Channel SMS or Messenger: SMS	2,324 (56.8%)	1,842 (56.3%)	482 (59%)
Mean number of askNivi conversation modules engaged (SD)	2.9 (1.0)	2.9 (1.0)	2.9 (1.1)
Use of FP			
Never used	1,533 (37.5%)	1,209 (37%)	324 (39.7%)
Currently using	1,452 (35.5%)	1,172 (35.8%)	280 (34.3%)
Not currently using, but used in the past	1,103 (27%)	890 (27.2%)	213 (26.1%)
Expressed FP-related intent at onboarding: yes	834 (20.4%)	670 (20.5%)	164 (20.1%)
Wants to delay/prevent pregnancy for >1 year: yes	3,084 (75.4%)	2,467 (75.4%)	617 (75.5%)
Perceived importance of preventing pregnancy			
Not important	279 (6.8%)	225 (6.9%)	54 (6.6%)
Less important	211 (5.2%)	175 (5.4%)	36 (4.4%)
Useful but not a primary goal	607 (14.8%)	481 (14.7%)	126 (15.4%)
Important	1,026 (25.1%)	813 (24.9%)	213 (26.1%)
Very important	1,965 (48.1%)	1,577 (48.2%)	388 (47.5%)
Stated readiness to visit provider (next 2 weeks)			
Not ready at all	797 (19.5%)	637 (19.5%)	160 (19.6%)
Not sure	1,329 (32.5%)	1,059 (32.4%)	270 (33%)
Ready	1,055 (25.8%)	860 (26.3%)	195 (23.9%)
Very ready	416 (10.2%)	337 (10.3%)	79 (9.7%)
Extremely ready	491 (12%)	378 (11.6%)	113 (13.8%)
Self-reported visit to family planning provider: Yes	405 (9.9%)	329 (10.1%)	76 (9.3%)

FP family planning.

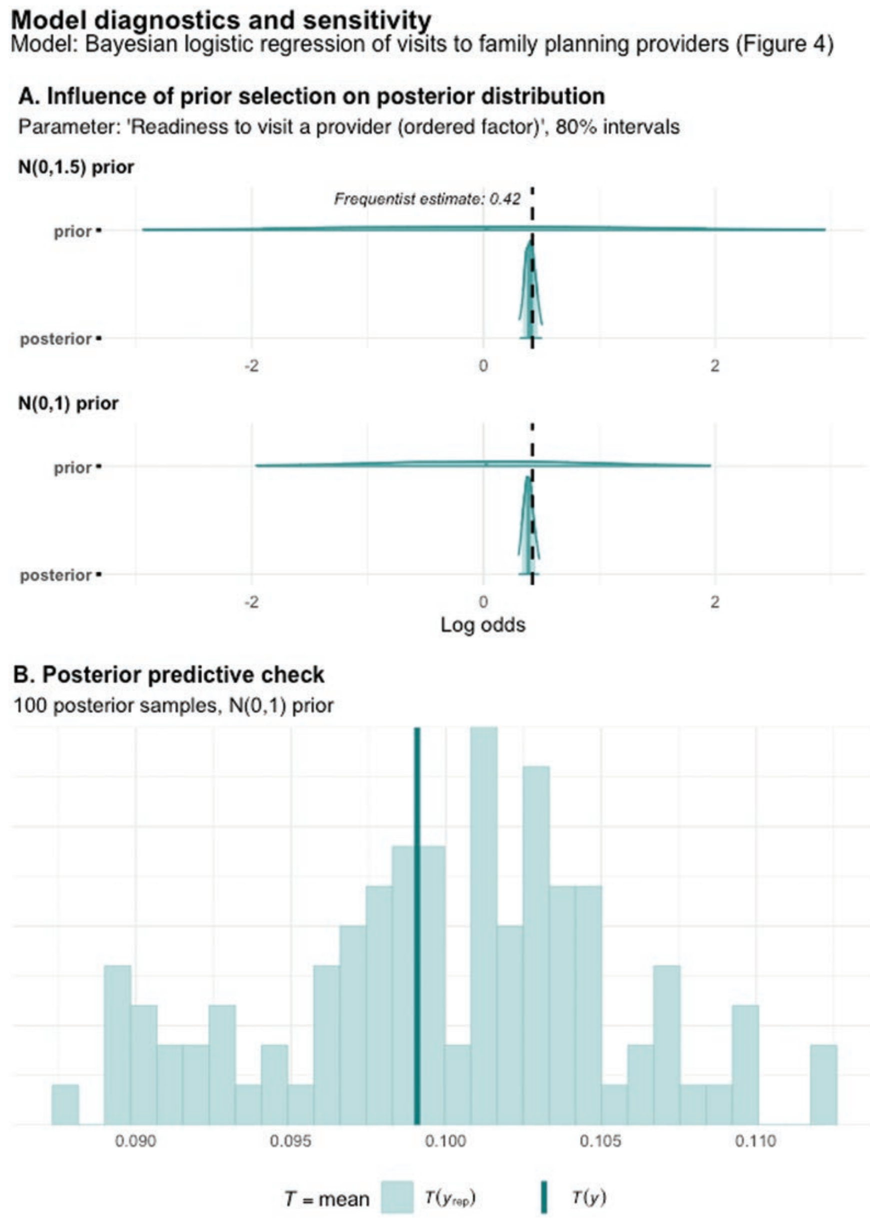

Fig A2 | This figure corresponds to the Bayesian logistic regression model of visits to family planning providers presented in Fig 4. Panel A shows how the choice between two different priors—a flat N(0,1.5) on the probability scale to a weakly regularizing N(0,1) prior—also has very little influence on the posterior of one of the key model parameters. We set all parameters in the model to N(0,1). The posterior predictive check in Panel B shows that the model generates synthetic data that captures the essential characteristics of the original data.

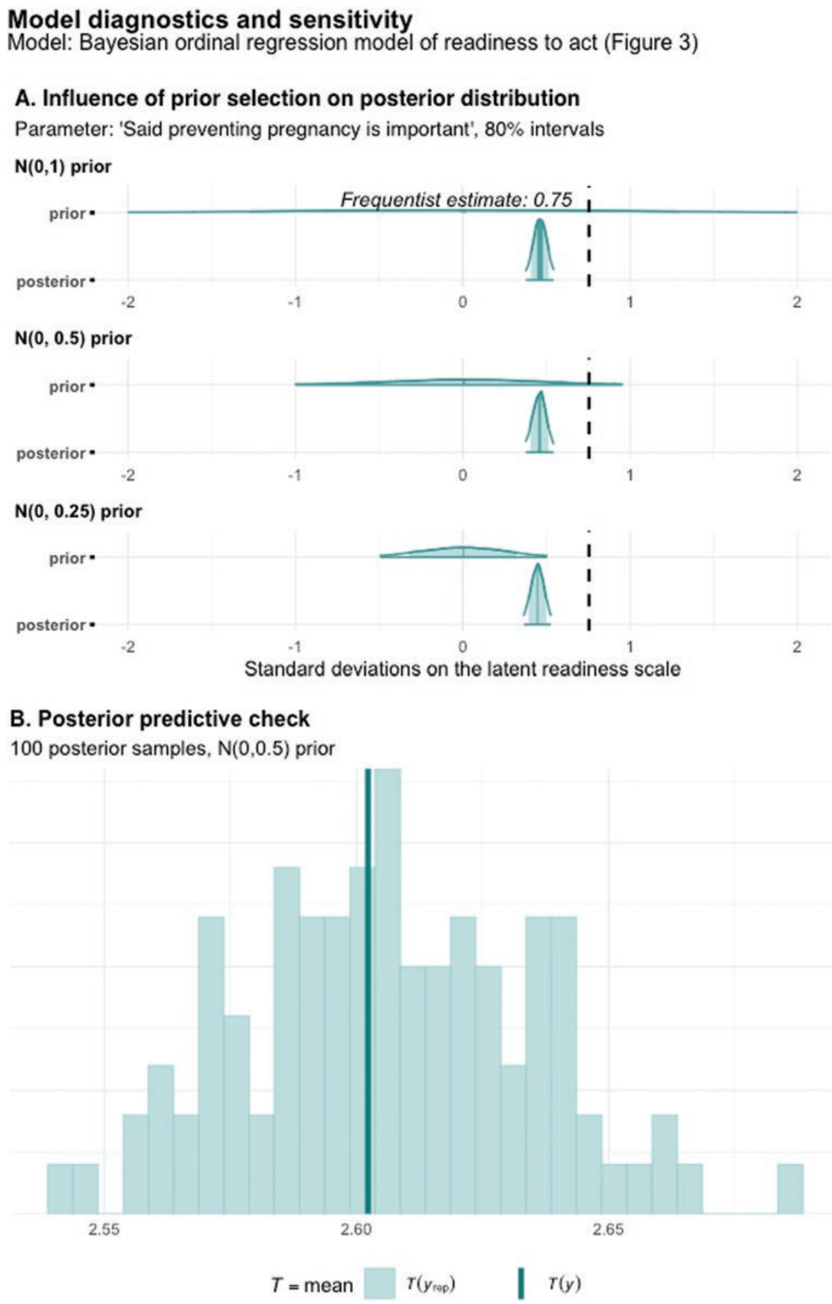

Fig A1 | This figure corresponds to the Bayesian ordinal regression model of readiness to act presented in Fig 3. Panel A shows how three different priors, from almost flat N(0,1) to weakly regularizing N(0,0.25), have very little influence on the posterior of one of the key model parameters. We set all parameters in the model to N(0,0.5). The posterior predictive check in Panel B shows that the model generates synthetic data that captures the essential characteristics of the original data.

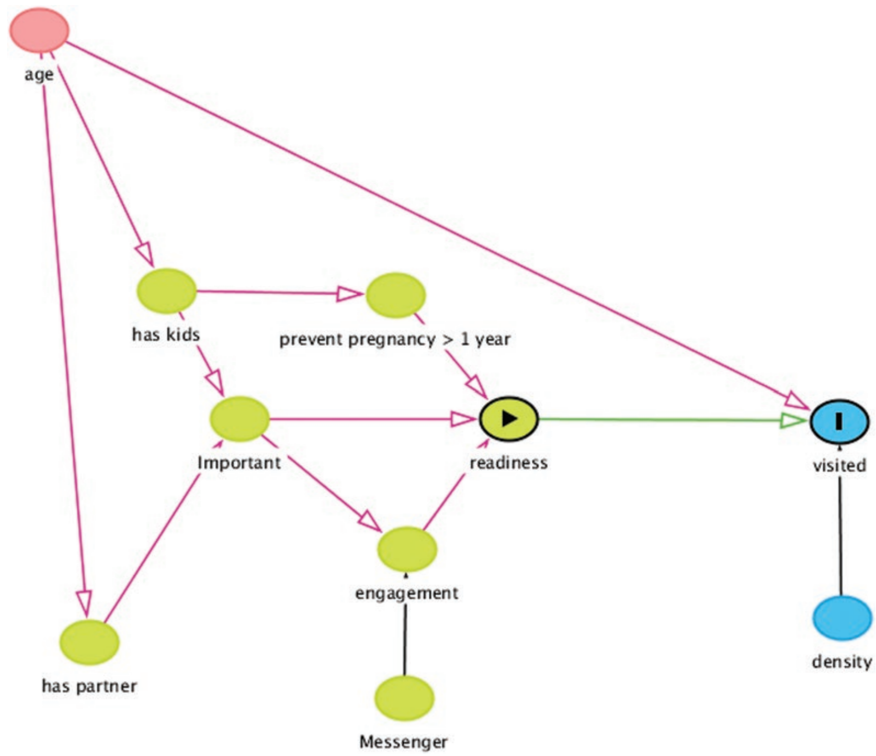

Fig A3 | Causal directed acyclic graph to inform the statistical adjustment for the linear regression of action.

References

1. WHO. WHO consolidated guideline on self-care interventions for health: sexual and reproductive health and rights; 2019. https://tinyurl.com/y4258wuj
2. WHO. Self-care health interventions. 2020. https://tinyurl.com/y5nnfpx9
3. Self-Care Trailblazer Group. Digital self-care: a framework for design, implementation, & evaluation; 2020. https://www.psi.org/project/self-care/digital-self-care/
4. Gibson DG, Tamrat T, Mehl G. The state of digital interventions for demand generation in low-and middle-income countries: considerations, emerging approaches, and research gaps. *Glob Health Sci.* 2018:6(suppl 1);S49–S60.
5. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. *J Med Internet Res.* 2017;19(8):e267.
6. Kocaballi AB, Berkovsky S, Quiroz JC, et al. The personalization of conversational agents in health care: systematic review. *J Med Internet Res.* 2019;21(11):e15360.
7. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. *J Behav Med.* 2019;42(1):67–83.
8. Pagoto S, Bennett GG. How behavioral science can advance digital health. *Transl Behav Med.* 2013;3(3):271–276.
9. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. *Annu Rev Public Health.* 2010;31(1):399–418.
10. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev.* 2015;9(3):323–344.
11. Aung B, Mitchell JW, Braun KL. Effectiveness of mHealth interventions for improving contraceptive use in low-and middle-income countries: a systematic review. *Glob Health Sci Pract.* 2020;8(4):813–826.
12. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med.* 2011;1(1):53–71.
13. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol.* 1983;51(3):390–395.
14. Norcross JC, Krebs PM, Prochaska JO. Stages of change. *J Clin Psychol.* 2011;67(2):143–154.
15. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med.* 2010;51(3–4):214–221.
16. Miller WR, Rollnick S. *Motivational Interviewing: Preparing People to Change Addictive Behavior.* New York: Guilford Publications; 1991.
17. Merrill JE, Wardell JD, Read JP. Is readiness to change drinking related to reductions in alcohol use and consequences? a week-to-week analysis. *J Stud Alcohol Drugs.* 2015;76(5):790–798.
18. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol.* 2016;31(4):337–350.
19. Shmueli G. To explain or to predict? *Stat Sci* 2010;25(3):289–310.
20. Facebook Connectivity Lab, & Center for International Earth Science Information Network. High resolution settlement layer; 2019. https://tinyurl.com/y34howpr
21. KNBS. *Kenya Demographic and Health Survey 2014.* Rockville, MD: Kenya National Bureau of Statistics; 2015.
22. Rowntree O. *The Mobile Gender Gap Report 2019.* London: GSMA; 2019.
23. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; 2020. https://www.R-project.org/
24. Bürkner PC. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw.* 2017;80(1):1–28. https://doi.org/10.18637/jss.v080.i01
25. Bürkner PC, Vuorre M. Ordinal regression models in psychology: a tutorial. *Adv Methods Pract Psychol Sci.* 2019;2(1):77–101.
26. Bürkner PC, Charpentier E. Modelling monotonic effects of ordinal predictors in Bayesian regression models. *Br J Math Stat Psychol.* 2020;73(3):420–451.
27. Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol.* 2016;45(6):1887–1894.
28. Kuhn M, Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles; 2020. https://www.tidymodels.org

29. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol.* 2015;68(2):134–143.
30. Dijkstra A, Conijn B, De Vries H. A match-mismatch test of a stage model of behaviour change in tobacco smoking. *Addiction.* 2006;101(7):1035–1043.
31. Prochaska JO, Norcross JC. *Systems of Psychotherapy: A Transtheoretical Analysis.* New York: Oxford University Press; 2018.
32. Krebs P, Norcross JC, Nicholson JM, Prochaska JO. Stages of change and psychotherapy outcomes: A review and meta-analysis. *J Clin Psychol.* 2018;74(11):1964–1979.
33. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med.* 2000;19(4):453–473.
34. Kreuter MW, Strecher VJ, Glassman B. One size does not fit all: the case for tailoring print materials. *Ann Behav Med.* 1999;21(4):276–283.
35. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun.* 2013;18(9):1039–1069.
36. Petty RE, Cacioppo JT. Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *J Pers Soc Psychol.* 1979;37(10):1915.
[dataset] Green E. *Data and Code to Reproduce. 'Predicting Healthcare Seeking Behavior Based on Stated Readiness to Act'.* Github; 2020. https://github.com/ericpgreen/readiness-2020