# Human-Centered Reinforcement Learning for Personalized Self-Management Strategies

IÑIGO URTEAGA, Applied Mathematics and Applied Physics, Data Science Institute, Columbia University, USA
NOÉMIE ELHADAD, Biomedical Informatics, Columbia University, USA

In this position paper, we discuss the challenges and exciting research directions for augmenting personal health informatics systems with AI-driven recommendations for self-management strategies. Because self-management is more successful when aligned with an individual's goals and context of daily living, as well as with their own health status and physiological responses, we argue that the promise of automated recommendations hinges on their personalization. We posit that reinforcement learning is a promising technique for learning and delivering such personalized self-management recommendations, if designed in a human-centered fashion.

**ACM Reference Format:**
Iñigo Urteaga and Noémie Elhadad. 2022. Human-Centered Reinforcement Learning for Personalized Self-Management Strategies . In *Proceedings of (CHI 22).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 INTRODUCTION

Personal health informatics solutions have been proposed to support self-management and scaffold problem solving for individuals [6, 8, 11, 15, 17, 24–26, 35, 38], and to promote experimentation that help identify potential triggers of disease flares [9, 10, 19], across a range of health conditions [2–4, 7, 18, 20, 23, 32, 37]. In many chronic diseases, there is strong evidence of person-to-person variation in treatment responses and associated symptoms. In addition, there are often no predetermined policy guidelines for self-management, and if there are, individuals are left with the burden of translating them into their day-to-day lives [5, 12, 28, 29, 31]. For instance, the recommendation to "*engage in regular exercise*" is left up to individuals to determine how to implement (e.g., which exercises would make sense to them, how often, and how intense). Human-centered AI has the potential to empower individuals in self-managing their chronic diseases better, but there are a number of unresolved challenges in marrying human-centered AI with health informatics systems for personalized self-management recommendations.

We envision self-tracking personal informatics tools augmented with automated, effective, safe, and actionable self-management recommendations. These recommendations would lower the guesswork for individuals in selecting which strategies to experiment with (either alone or in concert of each other), how to go about experimenting with them, and determining their efficacy. We posit that Reinforcement Learning (RL) [40] is a promising approach for personal informatics systems to learn and deliver these recommendations, if designed in a human-centered fashion.

We instantiate this position paper in our ongoing work with Phendo, a self-tracking app for endometriosis —a poorly understood, complex chronic condition with no cure or reliable treatment— that was designed to capture the experience and perspective of patients [13, 27, 43]. When eliciting the needs of patients and providers in caring for endometriosis, both stakeholders demanded tools that scaffold the trial-and-error self-experimentation process to support them figure

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2022 Association for Computing Machinery.
Manuscript submitted to ACM

out what self-management strategies to try and for how long [34]. Analysis of self-tracked data from a 10,500 Phendo users cohort further justifies the need for personalized self-management recommendations [33]. We find that users are heterogenous in which strategies they use for different health states and, furthermore, that users' response to the same management strategy is heterogeneous, both at the population and individual levels. This heterogeneity in effects applies to different self-management strategies, as well as to different self-management goals. Therefore, we argue for RL based personal self-management recommendations, as these techniques allow for individual-level optimization, rather than focusing on population-level estimates. We hereby discuss the feasibility of delivering RL-supported recommendations within a personal informatics system for chronic disease self-management. We then identify three avenues for research within the grand challenge of human-centered RL for personalized recommendations.

## 2 FEASIBILITY: USER, DATA, AND ALGORITHMIC PERSPECTIVES

Before designing an algorithmic solution for the personalization of self-management strategies, feasibility questions must be addressed: under which conditions would users accept such automated recommendations and whether the data and algorithmic requirements can ensure safe and personalized RL solutions.

*Acceptability.* It is critical to elucidate whether individuals want and would use an RL tool for self-management. Probing patients, as well as their caregivers and providers, helps map out how individuals managing a specific condition develop their own regimens and identify potential hindrances and bottlenecks in management, as well as elucidate practical scenarios of use for RL-based self-management recommendations. When prompted with the goal of an AI-augmented personal health system, individuals provide valuable information about opportunities and constraints they foresee. In our work, focus groups and semi-structured interviews with Phendo users established the acceptability of an RL-based self-management recommendation tool in principle, as long as individuals retained control over different aspects of the algorithmic recommendations [33].

*Data quality.* Wearable streams and self-tracked data are promising sources for monitoring an individual's physiological, physical, and mental wellbeing. The quality of RL-based recommendations hinges on the quality of these data and the signals inferred from them. Aligning raw and low-level data from wearables with constructs of interest (e.g., intensity of physical activity, presence of an acute symptom) is an active field of research. Smart processing of self-tracked data to disentangle observed data from constructs of interest is an emerging area of research. An individual's engagement with self-tracking through time intervenes with inferring the presence or absence of a symptom on a particular day, for instance [39]. Mechanisms to support engagement with a personal health informatics system are necessary, but might not be sufficient to ensure continued, daily engagement. Machine learning methods can help disentangle self-tracking engagement artifacts from true physiological and physical phenomena [22, 42].

*Algorithmic constraints.* When aiming for optimal RL policy learning in the wild, many algorithmic challenges arise. Beyond handling the high dimensionality of the state and action spaces typical of real-world scenarios and ensuring that effect-sizes are meaningful enough for policy learning, one must foresee the number of interactions the RL algorithm will require to find the right exploration-exploitation tradeoff. Critically, these constraints must be aligned and coupled with the users' goals and preferences; i.e., how often and for how long will the RL recommend strategies, and is the user willing to follow through? Towards real-life online deployment of RL-based recommendations, offline data analysis and pilot studies are a necessary starting point. For instance, our analysis of how long and how often do Phendo users engage with different self-management strategies in the wild, i.e., without algorithmic supervision, enabled us to asses the feasibility of the envisioned self-management tool from both the human and computational perspectives [33].

## 3 PERSONALIZATION CHALLENGES: CONJOINING HUMAN AND COMPUTATIONAL PERSPECTIVES

An RL-augmented personal informatics tool that will be used and is useful to individuals with chronic diseases must meet their needs, operate acceptably and within expectations, and fit into their lives and self-management routines. We argue that personalization is a necessary step towards the success of these future tools, and that incorporating human feedback into the learning process can help deliver an effective and safe human-centered RL solution.

*Personalized user-model.* For the state (context) and action (recommendation) spaces to be meaningful for each individual, human and algorithmic inputs shall be reconciled. The range of activities (RL actions) individuals experiment with is large in the self-management of chronic diseases. For endometriosis, we found that individuals are willing to explore a broad action space, as long as they can set some boundaries [33]. Therefore, the state space that describes users' context is complex, varies across individuals, and must incorporate personal data and explicit individual's input. The computational challenge is to translate this human input into informative RL state representations, so that self-management recommendations are fine-tuned and adapted to each individuals' context. These representations will require semi-supervised learning approaches, not only with input from each individual's preferences (e.g., "*I am not willing to run for more than 1 hour today*"), but with each users' goals (e.g., learn the computational representation that is most informative for achieving pain reduction in the short term). Furthermore, this user-model is likely to evolve as circumstances and health status fluctuate, which requires updating and revising the learned representations over time.

*Personalized recommendations.* A personalized RL algorithm must align its reward function with the user's notion of success. To that end, uncertainty in users' goal preferences (e.g., short term Vs long term) needs to be considered when defining the RL reward function. This, in turn, impacts not only what to recommend, but also the way recommendations are delivered, and the mechanisms for incorporating user feedback. Explainable reward function design [1] and inverse reward learning [16] could reduce the gap between the RL output and the users' understanding of the recommended self-management strategies, as well as to increase users' adherence to them. Another computational challenge pertains to finding personalized policies in the wild that maximize self-management "effect" (e.g., reduction in pain) and are acceptable to a user at a given time. For instance, if the algorithm recommends running but the user is not willing to run that day because of snow, a mechanism is needed to identify alternatives. A solution is to avoid single-item recommendations, and instead resort to more than one, ranked recommendations (e.g., via ranked bandits [21, 36]). Finally, the algorithm must accommodate users' time-varying preferences and effects, as what works for a user today might not be as effective tomorrow. Solutions to adapt to the time-evolving nature of individuals can be incorporated from the dynamic or restless bandit literature [14, 41], and the broader field of Markov Decision Processes [30].

*Personalized, safe policies.* A challenge with personalized RL recommendations is that —contrary to general guidelines decided at the population level— users have full control of not only the execution of the recommendations, but also the reporting of outcomes and symptoms. As such, the RL tool may recommend potentially unsafe options that the user is unaware of, e.g., to fast or to exercise for too long, or by recommending strategies that help in the short-term, but are harmful in the long-term. Because the user is in control of reporting the direct outcomes of the recommendations and other relevant health symptoms, it might be unclear how to capture the necessary information that, even if not directly related to the policy in itself, indicates harm, and how to flag it: for example, beyond the immediate effect of a recommendation (e.g., running) towards an individual's pain reduction goal, whether secondary, longer-term health outcomes (e.g., a foot stress fracture) are observed. Domain experts who provide additional, external knowledge and information over the timeline of RL recommended self-management strategies might be critical to guarantee these safety considerations.

## REFERENCES


[1] [n.d.]. ([n. d.]).

[2] ABL Berry, CY Lim, T Hirsch, AL Hartzler, LM Kiel, ZA Bermet, and JD Ralston. 2019. Supporting Communication About Values Between People with Multiple Chronic Conditions and Their Providers. In *Proc ACM CHI Conf* (New York, NY, USA) *(CHI '19)*. ACM, 470:1–470:14.

[3] Ross Brown, Bernd Ploderer, Leonard Si Da Seng, Peter Lazzarini, and Jaap van Netten. 2017. MyFootCare: a mobile self-tracking tool to promote self-care amongst people with diabetic foot ulcers. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (New York, NY, USA) *(OZCHI '17)*. Association for Computing Machinery, 462–466. https://doi.org/10.1145/3152771.3156158

[4] Adrian Bussone, Simone Stumpf, and Stephanie Wilson. 2019. Designing for Reflection on Shared HIV Health Information. In *Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the Next Interaction* (New York, NY, USA) *(CHItaly '19)*. ACM, 3:1–3:10. https://doi.org/10.1145/3351995.3352036

[5] Helen L Cameron-Tucker, Richard Wood-Baker, Christine Owen, Lyn Joseph, and E Haydn Walters. 2014. Chronic disease self-management and exercise in COPD as pulmonary rehabilitation: a randomized controlled trial. 9 (2014), 513–523. https://doi.org/10.2147/COPD.S58478

[6] Wonchan Choi, Shengang Wang, Yura Lee, Hyunkyoung Oh, and Zhi Zheng. 2020. A systematic review of mobile health technologies to support self-management of concurrent diabetes and hypertension. 27, 6 (2020), 939–945. https://doi.org/10.1093/jamia/ocaa029

[7] CF Chung, Q Wang, J Schroeder, A Cole, J Zia, J Fogarty, and SA Munson. 2019. Identifying and Planning for Individualized Change: Patient-Provider Collaboration Using Lightweight Food Diaries in Healthy Eating and Irritable Bowel Syndrome. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 3, 1 (2019), 7:1–7:27.

[8] Lori E. Crosby, Naomi E. Joffe, James Peugh, Russell E. Ware, and Maria T. Britto. 2017. Pilot of the Chronic Disease Self-Management Program for Adolescents and Young Adults With Sickle Cell Disease. 60, 1 (2017), 120–123. https://doi.org/10.1016/j.jadohealth.2016.08.022

[9] Nediyana Daskalova, Karthik Desingh, Alexandra Papoutsaki, Diane Schulze, Han Sha, and Jeff Huang. 2017. Lessons Learned from Two Cohorts of Personal Informatics Self-Experiments. 1, 3 (2017), 46:1–46:22. https://doi.org/10.1145/3130911

[10] Nediyana Daskalova, Eindra Kyi, Kevin Ouyang, Arthur Borem, Sally Chen, Sung Hyun Park, Nicole Nugent, and Jeff Huang. 2021. Self-E: Smartphone-Supported Guidance for Customizable Self-Experimentation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan). ACM, 1–13. https://doi.org/10.1145/3411764.3445100

[11] Pooja M. Desai, Elliot G. Mitchell, Maria L. Hwang, Matthew E. Levine, David J. Albers, and Lena Mamykina. 2019. Personal Health Oracle: Explorations of Personalized Predictions in Diabetes Self-Management. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 1–13.

[12] Tanja Effing, Gerhard Zielhuis, Huib Kerstjens, Paul van der Valk, and Job van der Palen. 2011. Community based physiotherapeutic exercise in COPD self-management: A randomised controlled trial. 105, 3 (2011), 418–426. https://doi.org/10.1016/j.rmed.2010.09.017

[13] Ipek Ensari, Adrienne Pichon, Sharon Lipsky-Gorman, Suzanne Bakken, and Noémie Elhadad. 2020. Augmenting the Clinical Data Sources for Enigmatic Diseases: A Cross-Sectional Study of Self-Tracking Data and Clinical Documentation in Endometriosis. *Applied Clinical Informatics* 11, 05 (2020), 769–784.

[14] Steffen Grünewälder and Azadeh Khaleghi. 2019. Approximations of the restless bandit problem. *The Journal of Machine Learning Research* 20, 1 (2019), 514–550.

[15] Ankit Gupta, Xin Tong, Chris Shaw, Linda Li, and Lynne Feehan. 2017. FitViz: A Personal Informatics Tool for Self-management of Rheumatoid Arthritis. In *HCI International 2017 – Posters' Extended Abstracts* (Cham) *(Communications in Computer and Information Science)*, Constantine Stephanidis (Ed.). Springer International Publishing, 232–240. https://doi.org/10.1007/978-3-319-58753-0_35

[16] Dylan Hadfield-Menell, Smitha Milli, Pieter Abbeel, Stuart J Russell, and Anca Dragan. 2017. Inverse Reward Design. In *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/32fdab6559cdfa4f167f8c31b9199643-Paper.pdf

[17] Chi Yan Hui, Robert Walton, Brian McKinstry, Tracy Jackson, Richard Parker, and Hilary Pinnock. 2017. The use of mobile applications to support self-management for people with asthma: a systematic review of controlled studies to identify features associated with clinical effectiveness and adherence. 24, 3 (2017), 619–632. https://doi.org/10.1093/jamia/ocw143

[18] Ravi Karkar, Jasmine Zia, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, and Roger Vilardaga. 2017. TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17* (Denver, Colorado, USA). ACM Press, 6850–6863. https://doi.org/10.1145/3025453.3025480

[19] Ravi Karkar, Jasmine Zia, Roger Vilardaga, Sonali R. Mishra, James Fogarty, Sean A. Munson, and Julie A. Kientz. 2016. A framework for self-experimentation in personalized health. 23, 3 (2016), 440–448. https://doi.org/10.1093/jamia/ocv150

[20] Sung-In Kim, Eunkyung Jo, Myeonghan Ryu, Inha Cha, Young-Ho Kim, Heejung Yoo, and Hwajung Hong. 2019. Toward Becoming a Better Self: Understanding Self-Tracking Experiences of Adolescents with Autism Spectrum Disorder Using Custom Trackers. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare* (New York, NY, USA) *(PervasiveHealth'19)*. Association for Computing Machinery, 169–178. https://doi.org/10.1145/3329189.3329209

[21] Anisio Lacerda. 2017. Multi-Objective Ranked Bandits for Recommender Systems. *Neurocomputing* 246 (2017), 12–24. https://doi.org/10.1016/j.neucom.2016.12.076 Brazilian Conference on Intelligent Systems 2015.

[22] Kathy Li, Iñigo Urteaga, Amanda Shea, Virginia J Vitzthum, Chris H Wiggins, and Noémie Elhadad. 2021. A predictive model for next cycle start date that accounts for adherence in menstrual self-tracking. *Journal of the American Medical Informatics Association* 29, 1 (09 2021), 3–11. https://doi.org/10.1093/jamia/ocab182

[23] CY Lim, ABL Berry, AL Hartzler, T Hirsch, DS Carrell, ZA Bermet, and JD Ralston. 2019. Facilitating Self-reflection About Values and Self-care Among Individuals with Chronic Conditions. In *Proc ACM CHI Conf* (New York, NY, USA) *(CHI '19)*. ACM, 655:1–655:12.

[24] Haley MacLeod, Anthony Tang, and Sheelagh Carpendale. 2013. Personal Informatics in Chronic Illness Management. In *Proceedings of Graphics Interface 2013* (Toronto, Ont., Canada, Canada) *(GI '13)*. Canadian Information Processing Society, 149–156.

[25] Lena Mamykina, Elizabeth M Heitkemper, Arlene M Smaldone, Rita Kukafka, Heather Cole-Lewis, Patricia G Davidson, Elizabeth D Mynatt, Jonathan N Tobin, Andrea Cassells, Carrie Goodman, and George Hripcsak. 2016. Structured scaffolding for reflection and problem solving in diabetes self-management: qualitative study of mobile diabetes detective. 23, 1 (2016), 129–136. https://doi.org/10.1093/jamia/ocv169

[26] Lena Mamykina, Arlene M. Smaldone, and Suzanne R. Bakken. 2015. Adopting the sensemaking perspective for chronic disease self-management. 56 (2015), 406–417. https://doi.org/10.1016/j.jbi.2015.06.006

[27] M McKillop, L Mamykina, and N Elhadad. 2018. Designing in the dark: eliciting self-tracking dimensions for understanding enigmatic disease. In *Proc ACM CHI Conf.*

[28] Carla K. Miller, Jean L. Kristeller, Amy Headings, and Haikady Nagaraja. 2014. Comparison of a Mindful Eating Intervention to a Diabetes Self-Management Intervention Among Adults With Type 2 Diabetes: A Randomized Controlled Trial. 41, 2 (2014), 145–154. https://doi.org/10.1177/1090198113493092

[29] Katy E. Mitchell, Vicki Johnson-Warrington, Lindsay D. Apps, John Bankart, Louise Sewell, Johanna E. Williams, Karen Rees, Kate Jolly, Michael Steiner, Mike Morgan, and Sally J. Singh. 2014. A self-management programme for COPD: a randomised controlled trial. 44, 6 (2014), 1538–1547. https://doi.org/10.1183/09031936.00047814

[30] George E Monahan. 1982. State of the art—a survey of partially observable Markov decision processes: theory, models, and algorithms. *Management science* 28, 1 (1982), 1–16.

[31] Nelda New. 2010. Teaching so they hear: Using a co-created diabetes self-management education approach. 22, 6 (2010), 316–325. https://doi.org/10.1111/j.1745-7599.2010.00514.x

[32] T Owen, J Pearson, H Thimbleby, and G Buchanan. 2015. ConCap: Designing to Empower Individual Reflection on Chronic Conditions Using Mobile Apps. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (New York, NY, USA) *(MobileHCI '15)*. ACM, 105–114.

[33] Adrienne Pichon, I nigo Urteaga, Lena Mamykina, and Noémie Elhadad. [n.d.]. Informing the Design of Individualized Self-management Regimens from the Human, Data, and Algorithmic Perspectives. ([n. d.]).

[34] Adrienne Pichon, Kayla Schiffer, Emma Horan, Bria Massey, Suzanne Bakken, Lena Mamykina, and Noémie Elhadad. 2020. Divided We Stand: The Collaborative Work of Patients and Providers in an Enigmatic Chronic Disease. 4 (2020), 261:1–261:24. Issue CSCW3. https://doi.org/10.1145/3434170

[35] Ari H. Pollack, Uba Backonja, Andrew D. Miller, Sonali R. Mishra, Maher Khelifi, Logan Kendall, and Wanda Pratt. 2016. Closing the Gap: Supporting Patients' Transition to Self-Management after Hospitalization. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA) *(CHI '16)*. Association for Computing Machinery, 5324–5336. https://doi.org/10.1145/2858036.2858240

[36] Filip Radlinski, Robert Kleinberg, and Thorsten Joachims. 2008. Learning Diverse Rankings with Multi-Armed Bandits. In *Proceedings of the 25th International Conference on Machine Learning* (Helsinki, Finland) *(ICML '08)*. Association for Computing Machinery, New York, NY, USA, 784–791. https://doi.org/10.1145/1390156.1390255

[37] Jessica Schroeder, Chia-Fang Chung, Daniel A. Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A. Munson. 2018. Examining Self-Tracking by People with Migraine: Goals, Needs, and Opportunities in a Chronic Health Condition. In *Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY, USA) *(DIS '18)*. ACM, 135–148. https://doi.org/10.1145/3196709.3196738

[38] Ian A. Scott, Paul Scuffham, Deepali Gupta, Tanya M. Harch, John Borchi, and Brent Richards. 2018. Going digital: a narrative overview of the effects, quality and utility of mobile apps in chronic disease self-management. 44, 1 (2018), 62–82. https://doi.org/10.1071/AH18064

[39] Mollie McKillop Sharon Lipsky-Gorman, Iñigo Urteaga and Noémie Elhadad. 2021. User Engagement Metrics and Patterns in Phendo, an Endometriosis Research Mobile App. *Journal of the American Medical Informatics Association* (2021). (Under review).

[40] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement learning: an introduction* (second edition ed.). The MIT Press.

[41] Cem Tekin and Mingyan Liu. 2012. Online Learning of Rested and Restless Bandits. *IEEE Transactions on Information Theory* 58, 8 (2012), 5588–5611. https://doi.org/10.1109/TIT.2012.2198613

[42] Inigo Urteaga, Kathy Li, Amanda Shea, Virginia J. Vitzthum, Chris H. Wiggins, and Noemie Elhadad. 2021. A Generative Modeling Approach to Calibrated Predictions: A Use Case on Menstrual Cycle Length Prediction. In *Proceedings of the 6th Machine Learning for Healthcare Conference (Proceedings of Machine Learning Research)*, Ken Jung, Serena Yeung, Mark Sendak, Michael Sjoding, and Rajesh Ranganath (Eds.), Vol. 149. PMLR, 535–566.

[43] Iñigo Urteaga, Mollie McKillop, and Noémie Elhadad. 2020. Learning endometriosis phenotypes from patient-generated data. *NPJ digital medicine* 3, 1 (2020), 1–14.