

Barriers to Implementing an Artificial Intelligence Model for Unplanned Readmissions

Sally L. Baxter^{1,2} Jeremy S. Bass^{1,3} Amy M. Sitapati^{1,4}

¹Health Department of Biomedical Informatics, University of California San Diego, La Jolla, California, United States

²Viterbi Family Department of Ophthalmology and Shiley Eye Institute, University of California San Diego, La Jolla, California, United States

³Department of Psychiatry, University of California San Diego, La Jolla, California, United States

⁴Department of Medicine, University of California San Diego, La Jolla, California, United States

Address for correspondence Sally L. Baxter, MD, MSc, 9415 Campus Point Drive, MC 0946, La Jolla, CA 92093-0946, United States (e-mail: s1baxter@health.ucsd.edu).

ACI Open 2020;4:108–113.

Abstract

Background Electronic health record (EHR) vendors now offer “off-the-shelf” artificial intelligence (AI) models to client organizations. Our health system faced difficulties in promoting end-user utilization of a new AI model for predicting readmissions embedded in the EHR.

Objectives The aim is to conduct a case study centered on identifying barriers to uptake/utilization.

Methods A qualitative study was conducted using interviews with stakeholders. The interviews were used to identify relevant stakeholders, understand current workflows, identify implementation barriers, and formulate future strategies.

Results We discovered substantial variation in existing workflows around readmissions. Some stakeholders did not perform any formal readmissions risk assessment. Others accustomed to using existing risk scores such as LACE+ had concerns about transitioning to a new model. Some stakeholders had existing workflows in place that could accommodate the new model, but they were not previously aware that the new model was in production. Concerns expressed by end-users included: whether the model’s predictors were relevant to their work, need for adoption of additional workflow processes, need for training and change management, and potential for unintended consequences (e.g., increased health care resource utilization due to potentially over-referring discharged patients to home health services).

Conclusion AI models for risk stratification, even if “off-the-shelf” by design, are unlikely to be “plug-and-play” in health care settings. Seeking out key stakeholders and defining clear use cases early in the implementation process can better facilitate utilization of these models.

Keywords

- readmissions
- case management
- predictive models
- predictive analytics
- electronic health records
- artificial intelligence
- machine learning
- health system
- clinical informatics

received
April 14, 2020
accepted
June 30, 2020

DOI <https://doi.org/10.1055/s-0040-1716748>.
ISSN 2566-9346.

© 2020 Georg Thieme Verlag KG
Stuttgart · New York

License terms



Background and Significance

The promise of artificial intelligence (AI)-based technologies, including machine learning (ML) algorithms, to improve health outcomes via predictive analytics has been widely touted.¹ The performance of predictive models has improved thanks to increasing electronic health record (EHR) data. However, best practices for implementing these models for point-of-care risk stratification in real-world settings are still evolving.^{2–4} Other industries have described challenges in a plug-and-play adoption model including inadequate consideration of strategy, workflow, and training.⁵

Recognizing the need for governance, the University of California San Diego (UCSD) formed an AI Committee of clinical informaticists, data scientists, and information technology professionals to standardize validation processes, ensure appropriate infrastructure, and provide oversight of predictive models licensed from the enterprise EHR vendor (Epic Systems, Verona, Wisconsin, United States). The vendor's models calculate risk scores from patients' data and can be built into EHR interfaces to support clinical decision-making. This case study describes our committee's efforts to analyze barriers to operationalize a model predicting unplanned readmissions.

Reducing unplanned readmissions is a key health care quality metric.^{6–9} The cost to Medicare of unplanned rehospitalization was estimated to be \$17.4 billion in a single year.¹⁰ UCSD licensed a model from the vendor that calculated risk of unplanned 30-day readmissions using LASSO-penalized logistic regression. Predictors included age, diagnoses, laboratory values, medication types, order types, and utilization data (e.g., current length of stay, past emergency department visits, past hospitalizations). The vendor reported the model's C-statistic to range from 0.69 to 0.74, exceeding the 0.63 to 0.69 range for LACE+ (length of stay, acuity of admission, Charlson comorbidity index score, and Emergency department visits in the past 6 months), a standard tool for predicting 30-day readmissions.^{11–14} Using institutional data for validation, the C-statistic of the model was 0.73. The available data indicated that the vendor's readmissions model outperformed LACE+. Despite the potential for improved risk stratification, initial efforts to encourage utilization by case managers at our institution failed, and no end-users were utilizing the new model in clinical practice even after several months of engagement.

Objectives

We aimed to conduct a detailed analysis of barriers to utilization with the perspectives of a framework focused on ML implementation in health care developed by Shaw et al¹⁵ and by the Nonadoption, Abandonment, and Challenges to the Scale-up, Spread, and Sustainability of Health and Care Technologies (NASSS) framework by Greenhalgh et al.¹⁶ Because both of these frameworks, as well as the more general framework of diffusion of innovations by Everett Rogers¹⁷ include a strong focus on end-users/adopters, we focused our analyses in this domain.

Methods

This quality improvement study was certified as nonhuman subjects research by the UCSD IRB. We employed a qualitative approach, gathering data via interviews at UCSD Health, an academic medical center with clinical sites in San Diego, CA, United States and La Jolla, CA, United States, from December 2019 to March 2020.

Using the NASSS framework,¹⁶ our first objective was to identify key stakeholders relevant to the following domains: condition (readmissions), adopters, and organization. Health system leadership identified key stakeholders. Subsequent stakeholder identification was iterative as we were referred to additional relevant personnel. In total, six stakeholder groups were identified. We conducted one 45- to 90-minute interview per stakeholder group. Fifteen individuals participated in interviews, with a mean of 4.7 years in their current leadership roles, although more than half (8/15, 53%) had a decade or more of cumulative clinical experience (→Table 1). Questions pertaining to current readmissions workflows and concerns utilizing the new model are shown in →Fig. 1.

We analyzed interview comments and workflow descriptions to conceptualize reasons for utilization barriers. Finally, we summarized lessons learned to help inform future implementation of EHR-based predictive models. We interpreted our findings within the context of the designated frameworks.

Results

Stakeholders

The stakeholder groups we identified relevant to the indicated NASSS framework domains included case management, social work, inpatient and outpatient pharmacy, population health management, transitional telephonic nursing, inpatient nursing, and physician leaders in clinical informatics.

Current Workflows and Readiness for Adoption

→Table 2 details each stakeholder, current methods used for risk stratification around readmissions, and workflow summaries. There were substantial variations in existing workflows. Several services did not have any formal risk assessments or existing workflows developed around readmissions risk scores in the EHR. Physicians indicated that they would be less likely to use an EHR-generated readmissions model given their existing familiarity with their patients' risk factors. In contrast, case management, transitional telephonic nursing, and population health utilized an array of existing risk scores and had well-developed workflows. However, they expressed varying degrees of enthusiasm for adopting the new model.

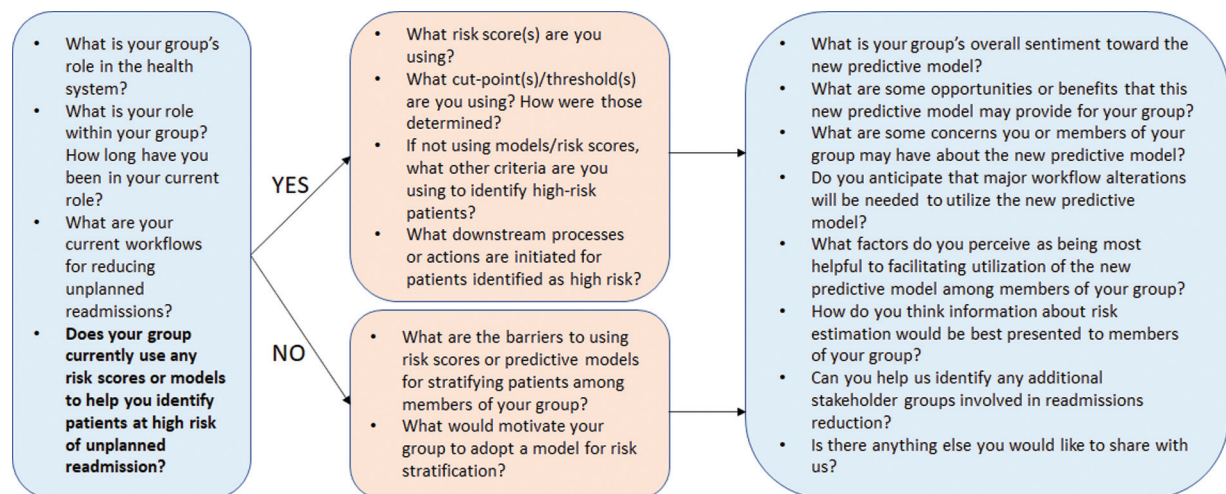
Concerns Regarding Model Implementation

The interviews not only clarified system use and acceptance of a new score but also identified potential utilization barriers (→Fig. 2). A concern expressed by multiple stakeholders was whether the score itself was useful. This was best exemplified by physicians and inpatient nursing who did not have any existing workflows around readmissions risk scores and indicated low likelihood of use. They anticipated that significant

Table 1 Characteristics of stakeholders interviewed about an updated predictive model for unplanned readmissions

Stakeholder group	Number of individuals interviewed	Gender distribution of individuals interviewed, number (%)	Mean (standard deviation) years in current role ^a of individuals interviewed (Years)
Case management	2	2 (100%) Female	4.3 (1.1)
Pharmacy (inpatient, outpatient, and transitions of care)	3	2 (67%) Female, 1 (33%) Male	4 (0)
Population health	1	1 (100%) Female	3
Transitional telephonic nursing	1	1 (100%) Female	5
Inpatient nursing	1	1 (100%) Female	14
Physicians/clinical informaticists	7	3 (43%) Female, 4 (57%) Male	4 (1.5)
<i>Total</i>	<i>15</i>	<i>10 (67%) Female, 5 (33%) Male</i>	<i>4.7 (2.7)</i>

^aNumbers indicate years in their current leadership roles at the institution under study, not cumulative working experience, which in many cases was much longer.

**Fig. 1** Flow diagram of interview questions for stakeholders regarding workflows around reducing unplanned readmissions and the potential role of a new predictive model.

effort would be required to develop completely new workflows, and it was unclear whether the institution would provide adequate resources and time. Some physicians also felt their personal knowledge of patients' histories and circumstances were more accurate than risk score estimations. Several stakeholders were also concerned about the relevance of the new model's predictors, particularly because social determinants of health were not included. Predictors also needed to be actionable. For example, several years prior to the study, inpatient pharmacy had created a custom pharmacy-specific risk score, citing that LACE+ and the General Risk Score (the existing risk scores at the time) lacked medication-related predictors and were not actionable for them.

Another concern from case management was training staff around a new risk score. Their leadership had recently invested considerable time and education into the use of LACE+ (including training around specific numeric cutoffs) for the current workflow, and managers were concerned about the additional time investment to replace LACE+. A solution that the stakeholders were open to was adding another column in the EHR interface that would display the new model's risk

score next to LACE+ with the same red/yellow/green color indications for high/medium/low risk. They also wanted the new model to have the option of displaying component predictors upon hovering over the score, a feature frequently used by their staff when using the LACE+ score.

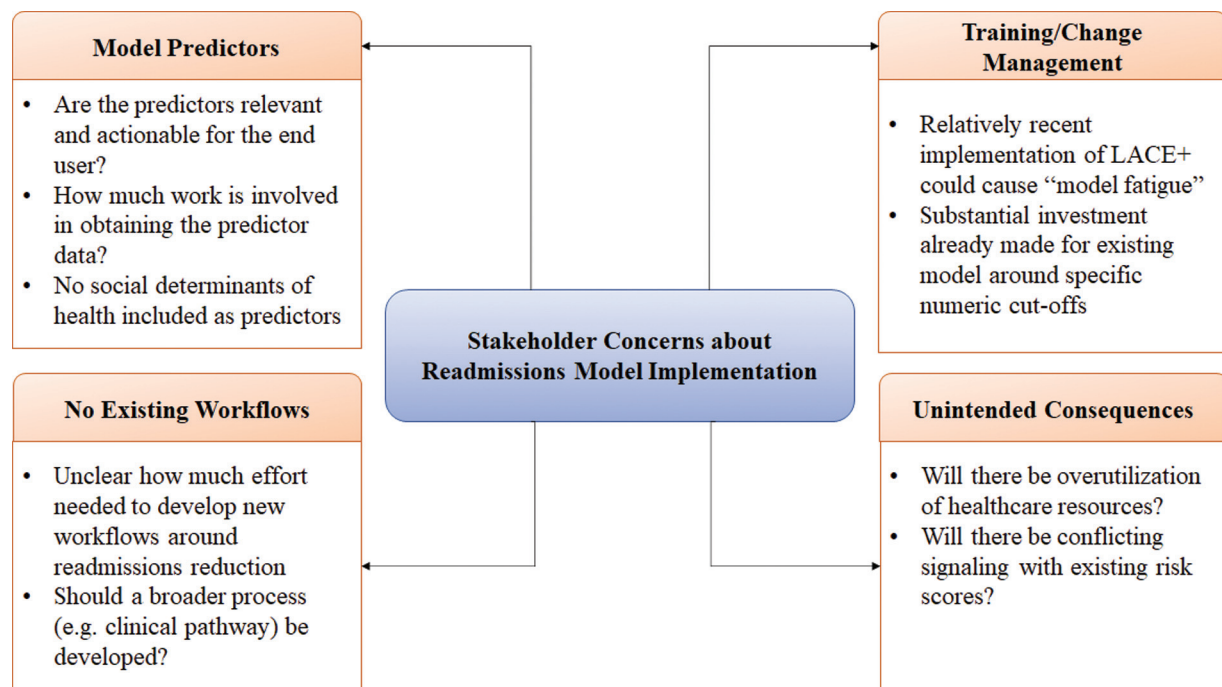
Finally, some stakeholders expressed concern about the possibility of over-utilizing health care resources if the new score lacked adequate specificity. They were also worried that with multiple risk scores already in use, adding an additional score could potentially cause confusion if risk estimation differed from existing models.

Discussion

Implementation of an "off-the-shelf" AI model from our EHR vendor for predicting readmissions failed to launch at our institution, despite evidence that it would perform better than existing models. In this case study, we aimed to better understand the failures of the initial implementation and provide lessons learned to improve future efforts. We believe that health care organizations should include governance for

Table 2 Readmissions risk estimation and summary of associated workflows for various stakeholder groups at the University of California San Diego

Stakeholder	Current method(s) of assessing readmissions risk	Existing workflows around reducing readmissions
Case management	LACE+	LACE+ used to guide post-discharge appointment scheduling, detailed chart reviews, triage for skilled nursing facility referrals, and prioritization of discussion during rounds focused on barriers to discharge and interdisciplinary care.
Social work	LACE+	Consult-driven, so assessments were based on consultation requests, rather than on risk scores. However, high LACE+ scores did prompt more detailed chart reviews.
Inpatient pharmacy	Custom pharmacy institutional risk score	The custom risk score was used to allocate pharmacy technician resources and prioritize discussions on interdisciplinary rounds.
Outpatient pharmacy	No formal assessment	No formal workflows
Population health services organization (PHSO)	Institutional general risk score (custom risk score previously developed by our health system), LACE + , patients with new diagnoses, elderly	Interventions for high-risk patients included assignment to a complex case management team, remote patient monitoring, home visits, mobile outreach (e.g., text reminders), and detailed chart reviews for identification of additional needs.
Transitional telephonic nursing	LACE + , health literacy, social determinants of health	Initial outreach did not involve risk stratification, as all patients were contacted by the telephonic nursing team after hospital discharge. However, high-risk patients identified by the listed criteria received additional outreach to connect them to community resources.
Inpatient nursing	No formal assessment	No formal workflows
Physicians	No formal assessment	No formal workflows

**Fig. 2** Concerns expressed by stakeholders regarding implementation of a new predictive model for unplanned readmissions.

an AI Committee to provide oversight, organizational alignment, and long-term centralized knowledge related to performance, workflow, and training.

First, we found that the initial implementation lacked sufficient buy-in from end-users and clear use cases. Prior work by Benda et al has indicated obtaining support from organizational leadership, end-users, and informaticists at

the beginning of a project is critical.¹⁸ A key insight from that study was, “Start informatics innovations by identifying a problem, not the data.” In many current practices, end-users are not engaged early enough. This was a relevant issue in this implementation, where clear use cases had not been defined at the outset. This was also identified by the Shaw framework on implementing ML in health care, where ML

algorithms “must have a meaningful entryway into decision making...algorithms that perform isolated risk prediction may be less useful.”¹⁵ Using the NASSS framework,¹⁶ in this case the organization was ready for the change, but the adopters (staff/caregivers) were not.

Although the organization had an overarching goal of preventing unplanned readmissions, different areas of the health system had differing focus areas. Aligning different organizational components prior to creating and adopting new data models is critical in complex adaptive health systems.^{18,19} As an illustration of disjointed communication, we found that the stakeholders’ varying risk assessments could result in different patients being labeled “high-risk.” Additionally, stakeholders were interested in different predictors based on what interventions they could provide. For example, the pharmacists highlighted that medication-related predictors were not included in LACE+, which was the primary score used by case managers. As a result, sometimes case managers were not aware of medication-related issues, such as gaps in insurance coverage for certain medications or home health needs such as intravenous antibiotic therapy. Therefore, using a common model across different end-users may improve alignment in readmission workflows (i.e., common identification of high-risk patients, awareness of the same issues). The initial implementation focused primarily on case management and failed to take a global view with engagement of other stakeholders. Future efforts may present opportunities for improved communication and alignment among these different groups.

Future trainings should not be centered around a specific cut-point (as had been done by case management around LACE+), but rather on an overall process, where different models could be substituted. Updating models in the future would incur less change management difficulties if there was more focus on the surrounding workflows. Another change management strategy was to create an additional column next to LACE+ with the new risk score in the same style and color coding. This obfuscated the need for retraining and would allow the end-users a side-by-side comparison to ease the transition. With the increasing number of predictive models in use in EHR systems, working toward improved standardization or normalization would be helpful, so that end-users could easily identify low, medium, or high-risk patients without needing to memorize specific cut-offs. EHR vendors should also consider supplying confidence information about predictions from these models, which is currently lacking.

Stakeholders noted that actionable predictors relevant to their specific service was an important consideration for adoption. This highlights the importance of explainability in implementing ML in health care, identified as a key issue in the Shaw framework¹⁵ and in the technology domain of the NASSS framework.¹⁶ The EHR vendor provided specific predictors for this model, but other ML or AI models based on unsupervised learning or deep learning may be less interpretable, without clear indication of what features are driving clinical predictions (i.e., “black-box” predictions).²⁰ Lack of explainability may decrease trust among adopters. Therefore, ongoing work to improve algorithmic transparency in AI models is critically important.

Although implementing AI and specifically ML carries some specific considerations such as the importance of explainability and the need for scaled-up computational resources, one of our key observations was that principles of traditional clinical support still apply. The improved performance of the ML model did not obviate the need for workflow analysis, appropriate training, and end-user engagement.

Limitations of the study included its observational nature and limited generalizability given its focus on the experience of a single academic center. However, we anticipate that many of the workflows, concerns, and barriers are not unique to our institution. Lessons learned through this analysis may help inform other organizations’ efforts to operationalize AI models into their EHRs.

Conclusion

AI will have widespread influence in health care by improving the performance of data models, risk scores, and predictive analytics. These will have the potential to improve patient health, decrease provider burnout, and improve efficient use of resources. However, given the complex sociotechnical aspects of health care organizations, it is unlikely that new models will just be “plug-and-play.” Barriers to utilizing a readmissions model at our institution included wide variation in existing workflows and stakeholder readiness for adoption, lack of perceived relevance, unclear resource requirements for training efforts, and lack of awareness. Although AI models may have improved performance, applying principles of traditional clinical decision support such as defining workflows and engaging end-users at an early stage is still highly relevant. Doing so may improve the speed and feasibility of AI adoption in the future.

Clinical Relevance Statement

The clinical relevance of this study is to highlight difficulties in operationalizing predictive models in electronic health records systems for clinical use, even when they are likely to outperform existing models. We provide lessons learned in an implementation failure at our institution to inform future adoption efforts.

Protection of Human and Animal Subjects

The study was performed in compliance with the Declaration of Helsinki and was reviewed by the UCSD Institutional Review Board, which declared the study as a quality improvement protocol and certified that the study did not qualify as human subjects research according to the Code of Federal Regulations, Title 45, part 46 and UCSD Standard Operating Policies and Procedures.

Funding

This study was supported by the National Institutes of Health/National Library of Medicine (grant T15LM011271). The funding organization had no role in the design or conduct of the study.

Authors' Contributions

S.L.B., J.B., and A.M.S. conceived and designed the study. S.L.B. and J.B. conducted interviews and data collection. S.L.B., J.B., and A.M.S. participated in data analysis and interpretation. S.L.B. and J.B. drafted the manuscript. All authors provided critical review of the manuscript for important intellectual content and approved the final version of the manuscript.

Conflict of Interest

None declared.

Acknowledgments

The authors wish to thank the UCSD Health AI Committee for input and feedback throughout the stakeholder engagement process described in this case study.

References

- 1 Parikh RB, Kakad M, Bates DW. Integrating predictive analytics into high-value care: the dawn of precision delivery. *JAMA* 2016; 315(07):651–652
- 2 Harris AH. Path from predictive analytics to improved patient outcomes: a framework to guide use, implementation, and evaluation of accurate surgical predictive models. *Ann Surg* 2017;265 (03):461–463
- 3 Amarasingham R, Patzer RE, Huesch M, Nguyen NQ, Xie B. Implementing electronic health care predictive analytics: considerations and challenges. *Health Aff (Millwood)* 2014;33(07): 1148–1154
- 4 He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019;25(01):30–36
- 5 Fountaine T, McCarthy B, Saleh T. Building the AI-powered organization. *Harvard Business Review*. 2019. Available at: <https://hbr.org/2019/07/building-the-ai-powered-organization>. Accessed April 13, 2020
- 6 Joynt KE, Jha AK. A path forward on Medicare readmissions. *N Engl J Med* 2013;368(13):1175–1177
- 7 Institute of Medicine (U.S.) *Rewarding Provider Performance: Aligning Incentives in Medicare*. Washington, DC: National Academies Press; 2007
- 8 Zuckerman RB, Sheingold SH, Orav EJ, Ruhter J, Epstein AM. Readmissions, observation, and the hospital readmissions reduction program. *N Engl J Med* 2016;374(16):1543–1551
- 9 Nuckols TK, Keeler E, Morton S, et al. Economic evaluation of quality improvement interventions designed to prevent hospital readmission: a systematic review and meta-analysis. *JAMA Intern Med* 2017;177(07):975–985
- 10 Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med* 2009;360(14):1418–1428
- 11 Ibrahim AM, Koester C, Al-Akchar M, et al. HOSPITAL Score, LACE Index and LACE+ Index as predictors of 30-day readmission in patients with heart failure. *BMJ Evid Based Med* 2019. Doi: 0.1136/bmjebm-2019-111271
- 12 Caplan IF, Sullivan PZ, Kung D, et al. LACE+ index as predictor of 30-day readmission in brain tumor population. *World Neurosurg* 2019;127:e443–e448
- 13 Caplan IF, Zadnik Sullivan P, Glauser G, et al. The LACE+ index fails to predict 30-90 day readmission for supratentorial craniotomy patients: a retrospective series of 238 surgical procedures. *Clin Neurol Neurosurg* 2019;182:79–83
- 14 Etttyreddy AR, Kao WTK, Roland LT, Rich JT, Chi JJ. Utility of the LACE scoring system in predicting readmission following tracheotomy and laryngectomy. *Ear Nose Throat J* 2019;98(04):220–222
- 15 Shaw J, Rudzicz F, Jamieson T, Goldfarb A. Artificial intelligence and the implementation challenge. *J Med Internet Res* 2019;21 (07):e13659
- 16 Greenhalgh T, Wherton J, Papoutsis C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017;19(11):e367
- 17 Rogers EM. *Diffusion of Innovations*. 5th ed New York, NY: Free Press; 2003
- 18 Benda NC, Das LT, Abramson EL, et al. “How did you get to this number?” Stakeholder needs for implementing predictive analytics: a pre-implementation qualitative study *J Am Med Inform Assoc* 2020;27(05):709–716
- 19 Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010;19(Suppl 3):i68–i74
- 20 Sussillo D, Barak O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput* 2013;25(03):626–649