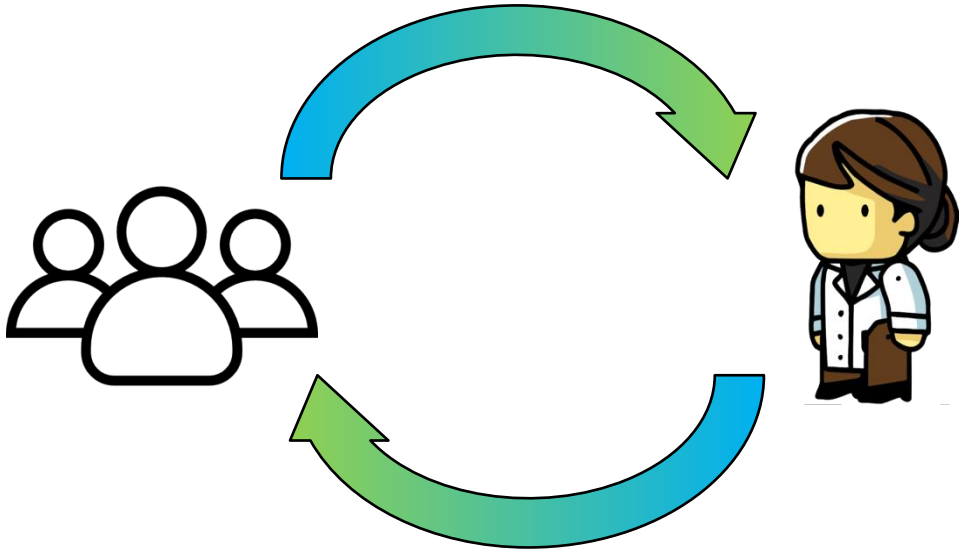
Closing the loop in healthcare: Bringing the benefits of machine learning back to the people



Hsin-Yu (Jane) Lai

NIH/NIDA P50DA054039
NIH/NIBIB and OD P41EB028242

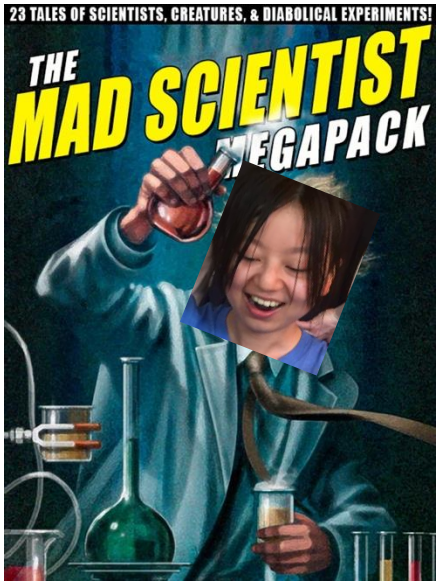
Self introduction?



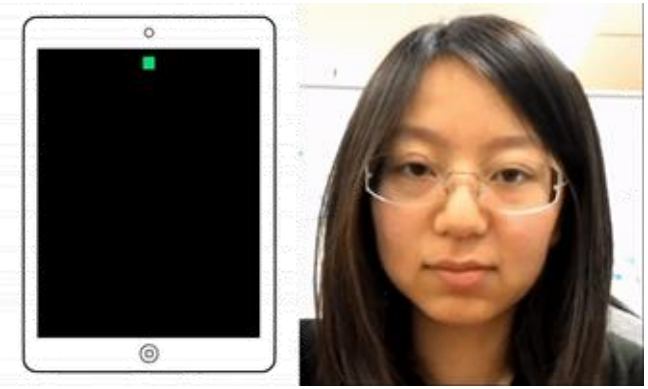
What my friends think



What my parents think



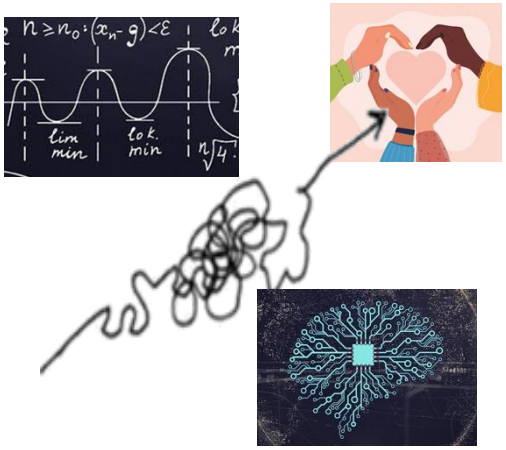
What society think



What my advisors think



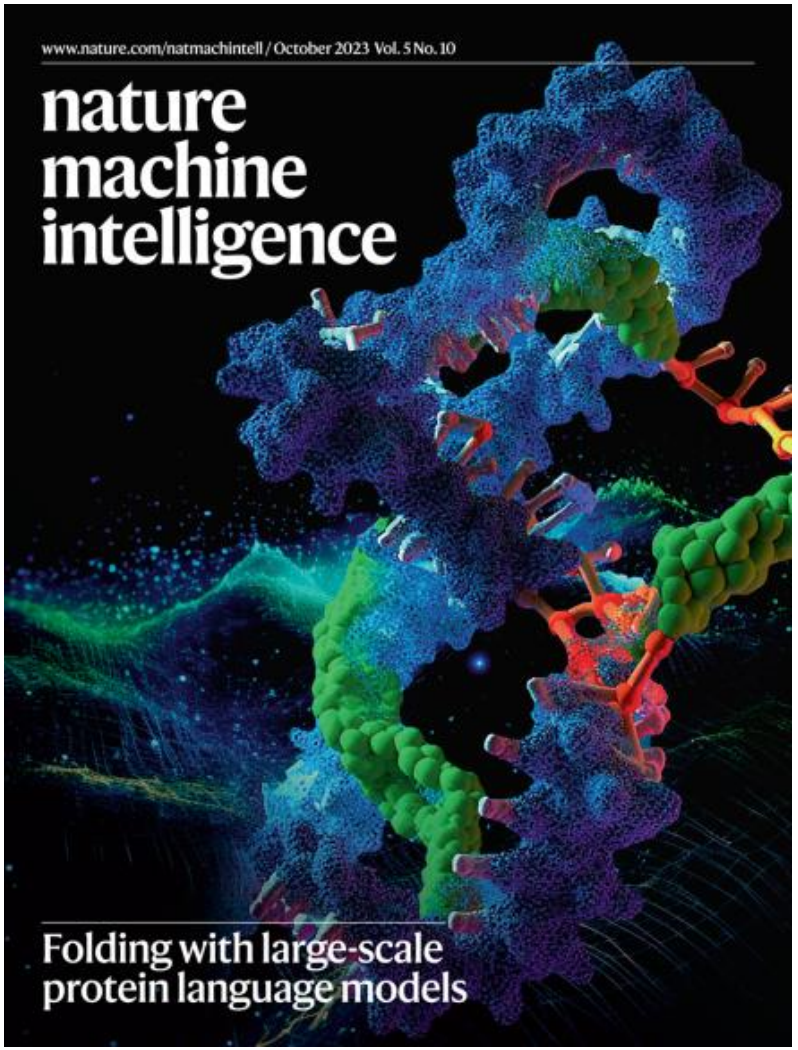
What I think



Who I actually am(?)

Machine learning (ML) has made groundbreaking progress in healthcare applications





ML helps doctors/scientists analyze data

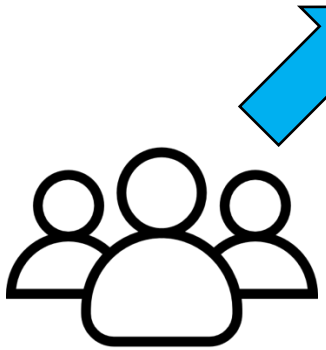
‘Omic data,
Sensor data,
Medical imaging,
Medical records

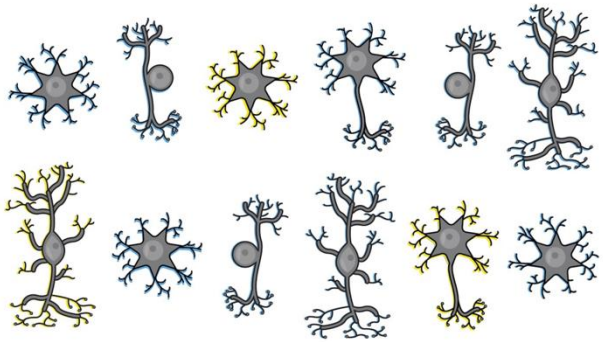
Signal processing,
Deep learning,
Statistics

Scientific discovery
Diagnosis
Treatment design
Disease progression monitoring

Data

Downstream tasks

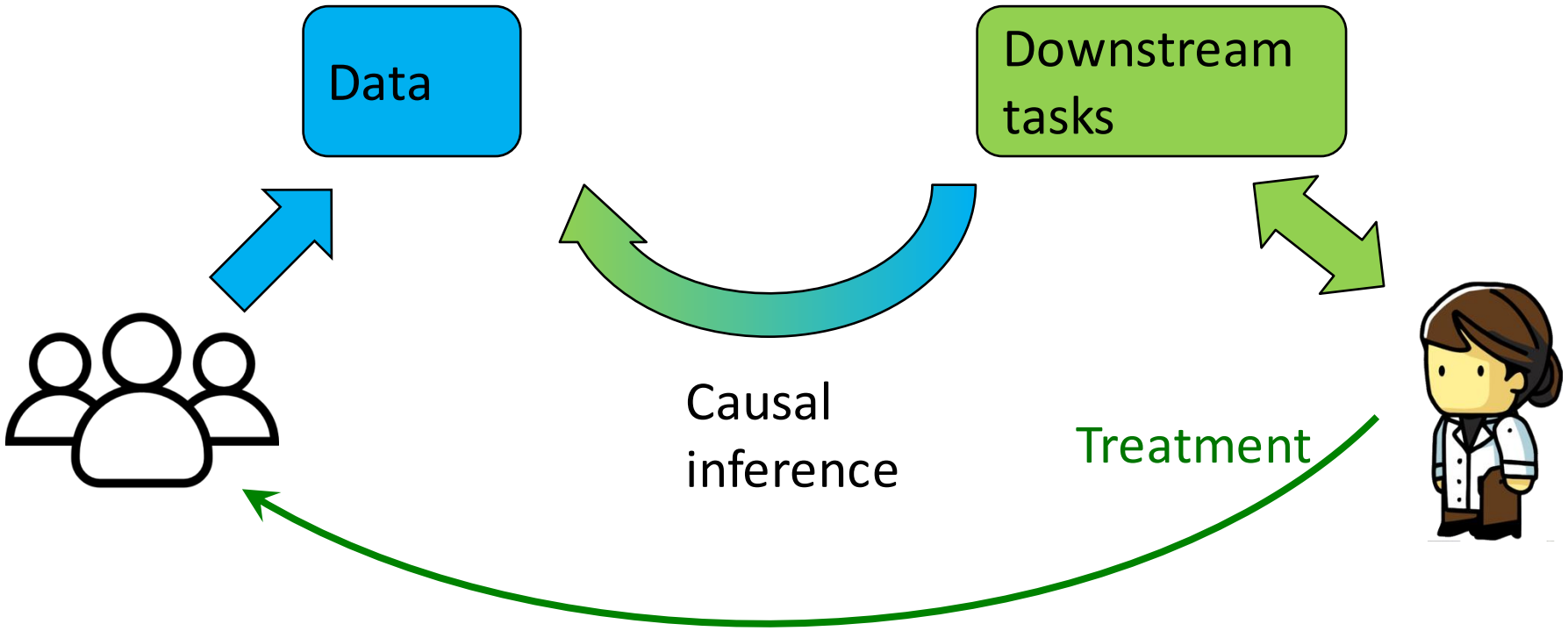




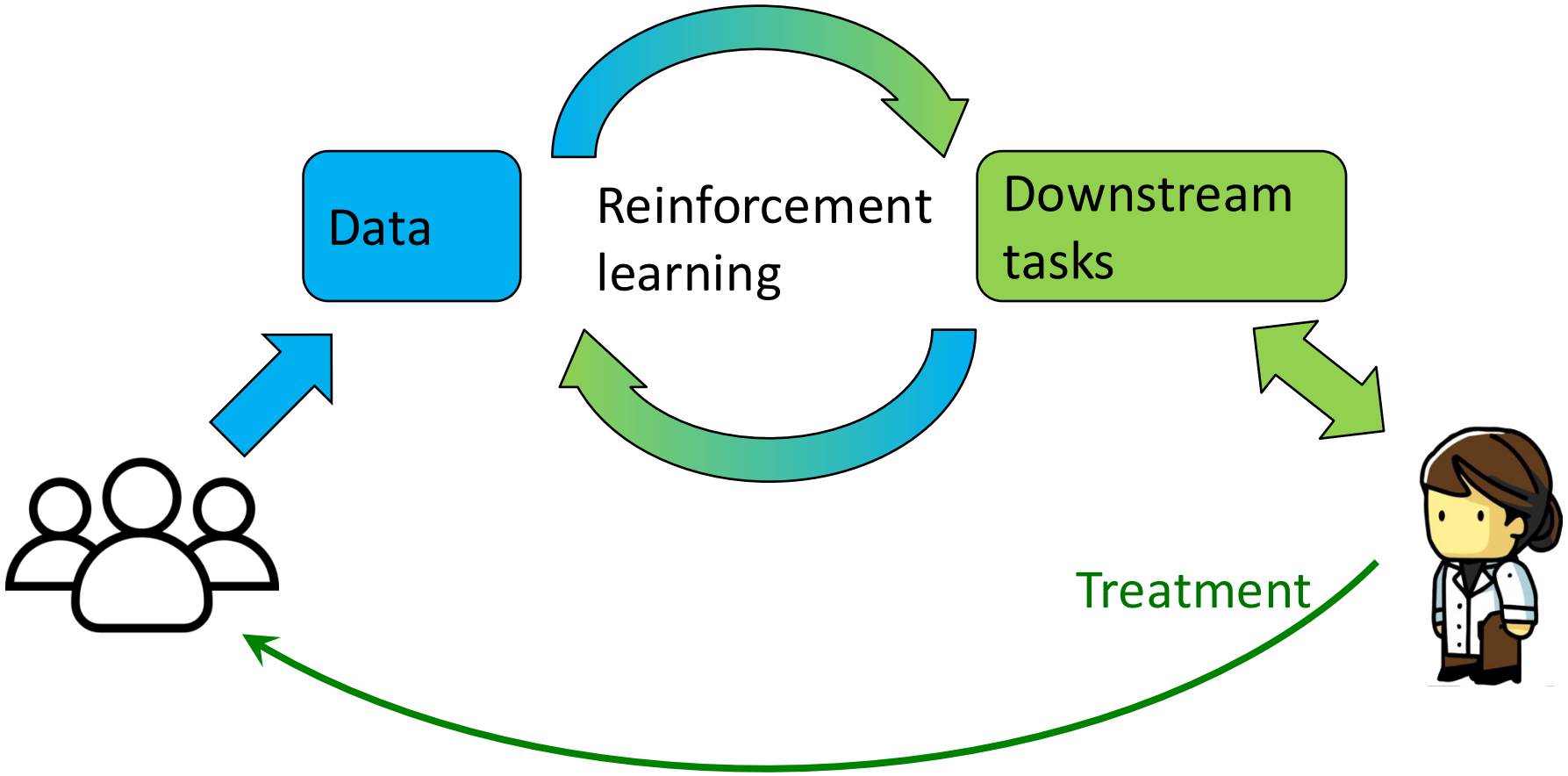
From 08.05.20. Allen Institute News



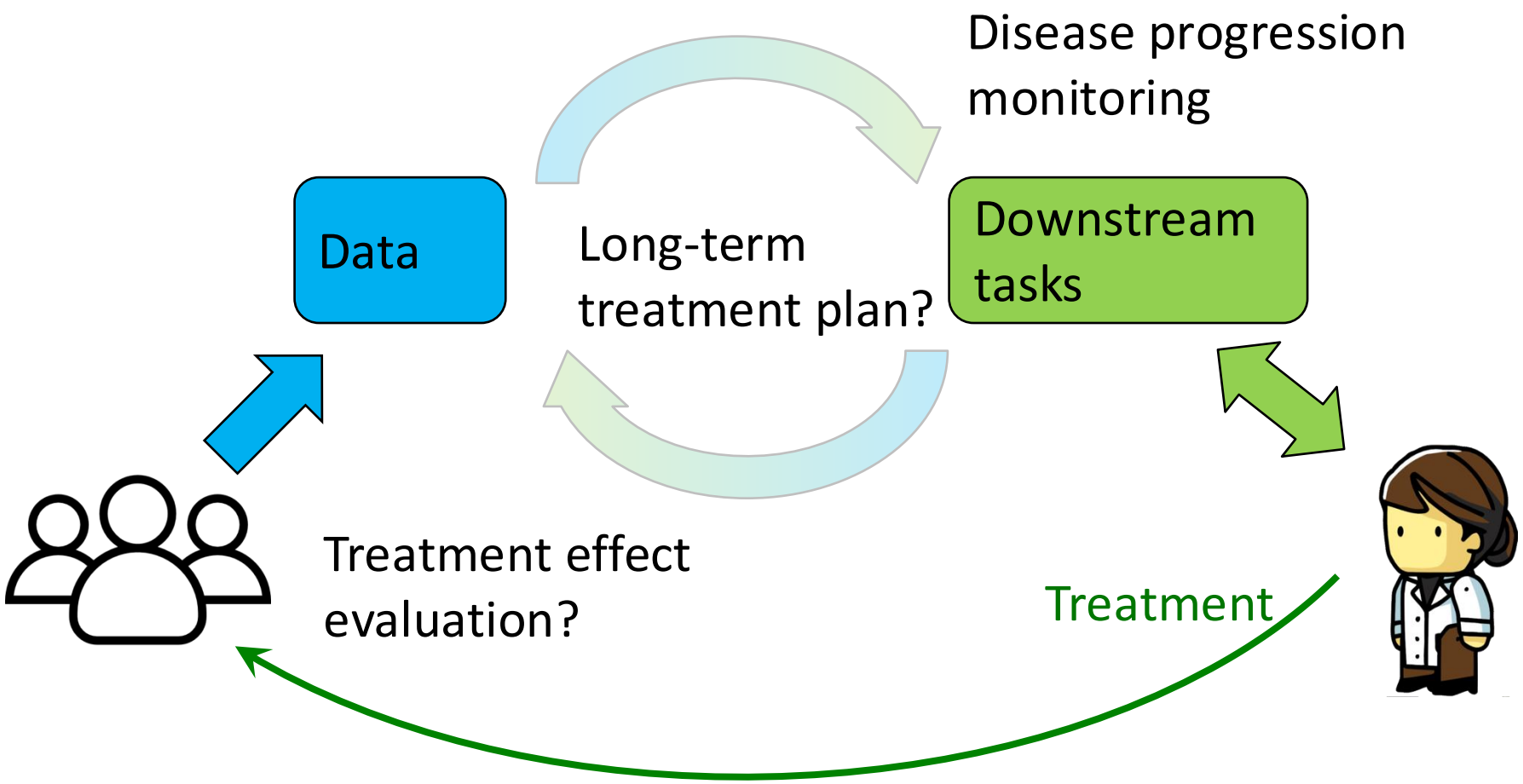
ML helps doctors/scientists evaluate causal relationships



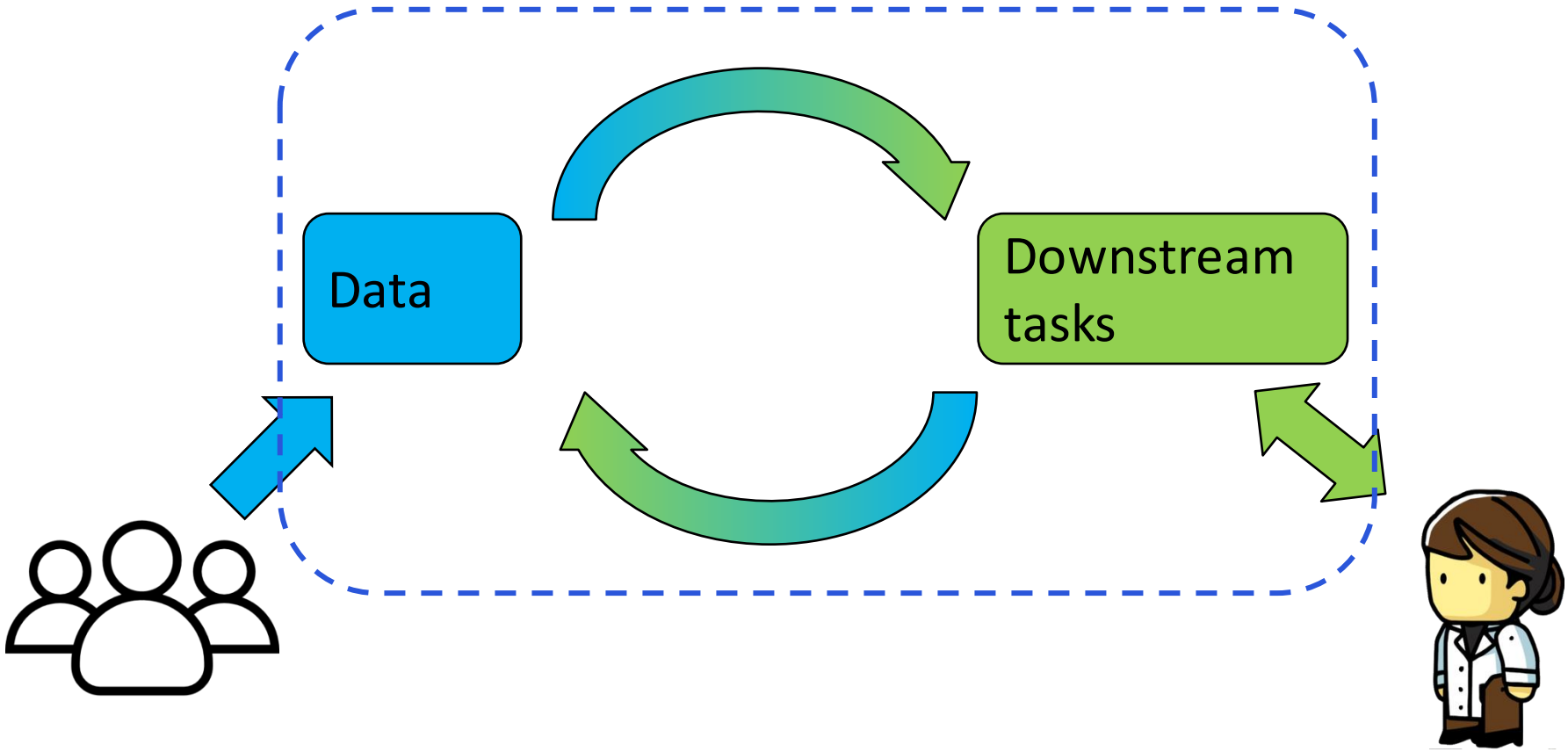
ML helps doctors/scientists make long-term plans



Could we close the loop?



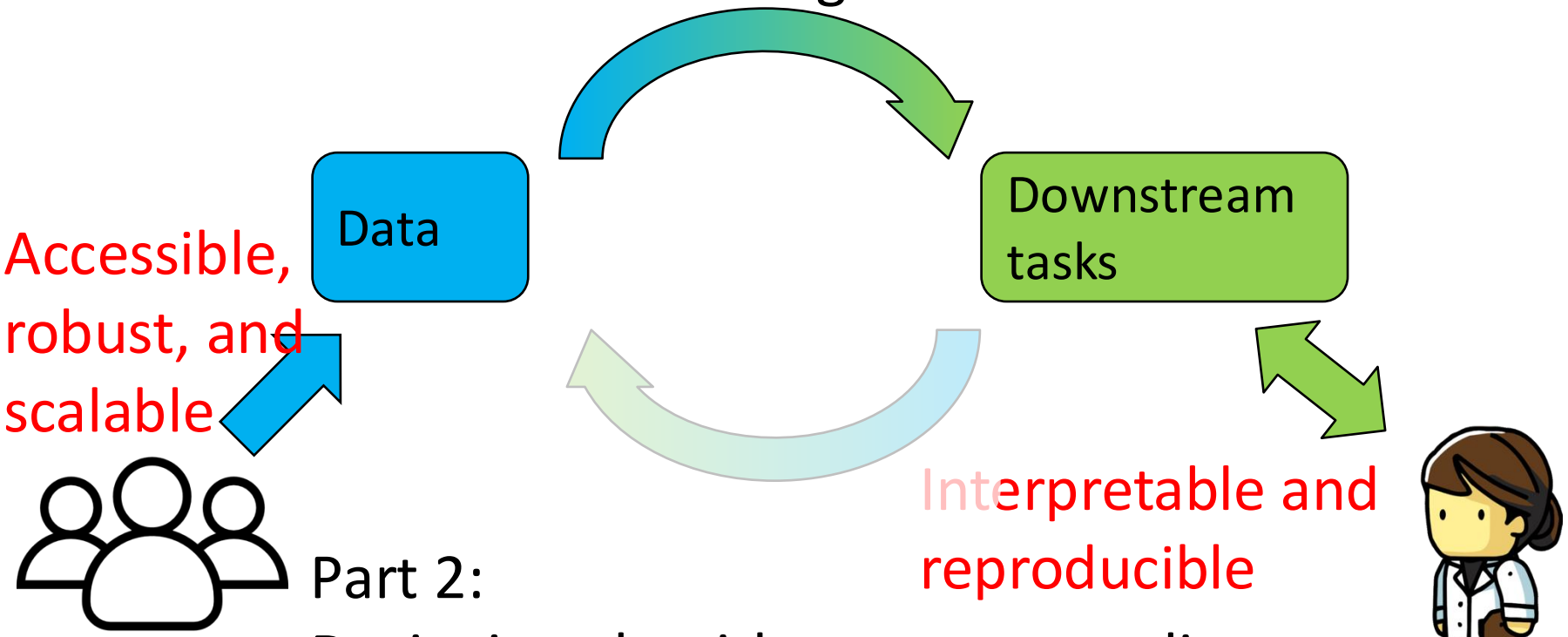
While many ML scientists focus on one of these domains,



My vision is to close the loop and bring the benefits back to the people

Part 1:

Tracking Neurodegenerative-Disease-Affected Eye-Movement Features Using Mobile Devices



Part 2:

Designing algorithms to personalize digital interventions for long-term goals

Acknowledgement



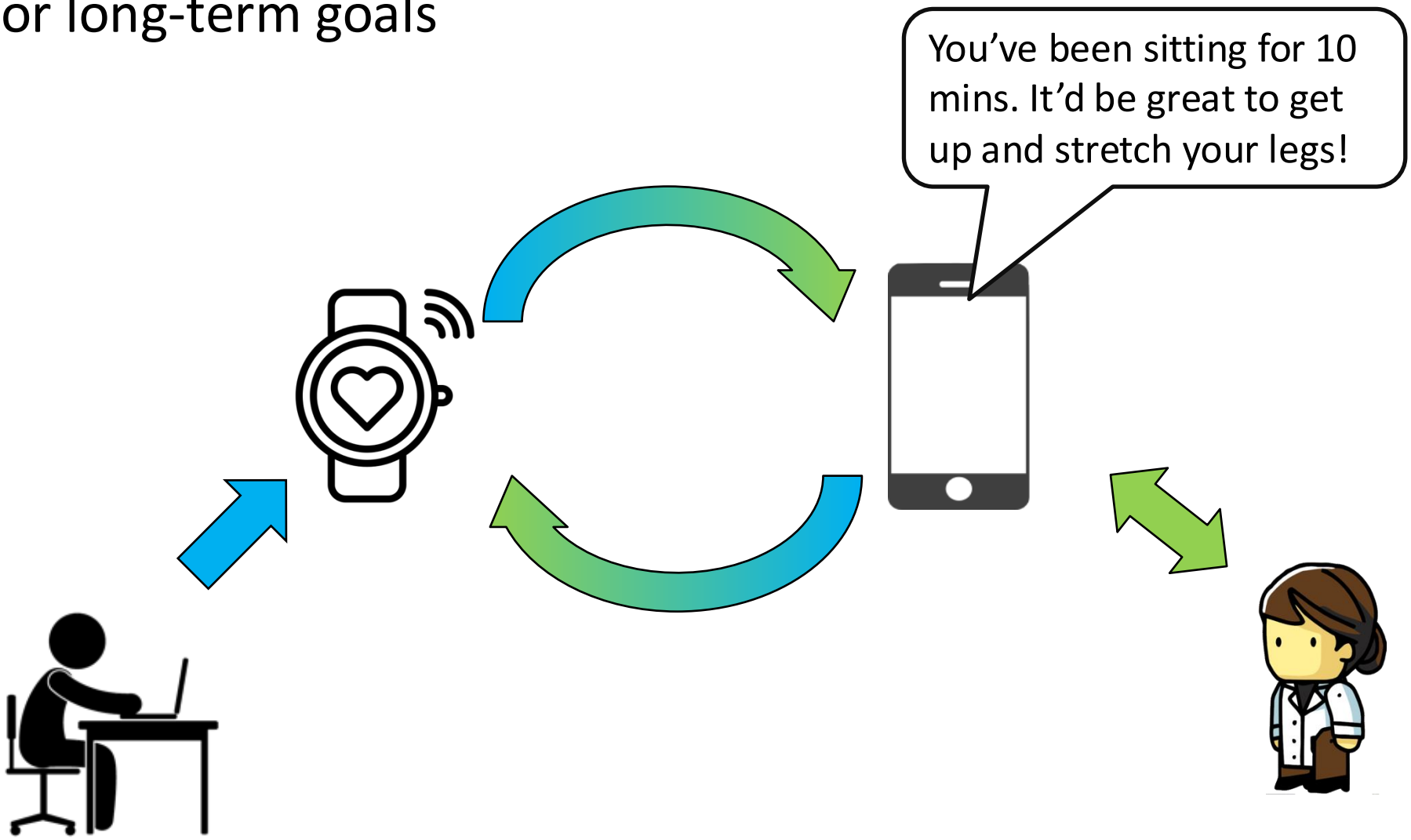








Part 2:
Designing algorithms to personalize digital interventions for long-term goals

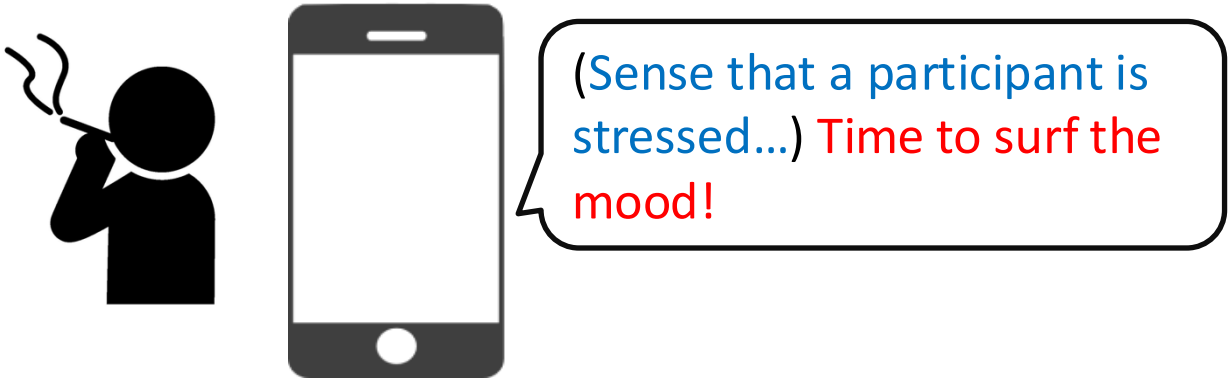


Mobile health enables Just-in-Time Adaptive Interventions (JITAIs)

HeartSteps



Sense2Stop



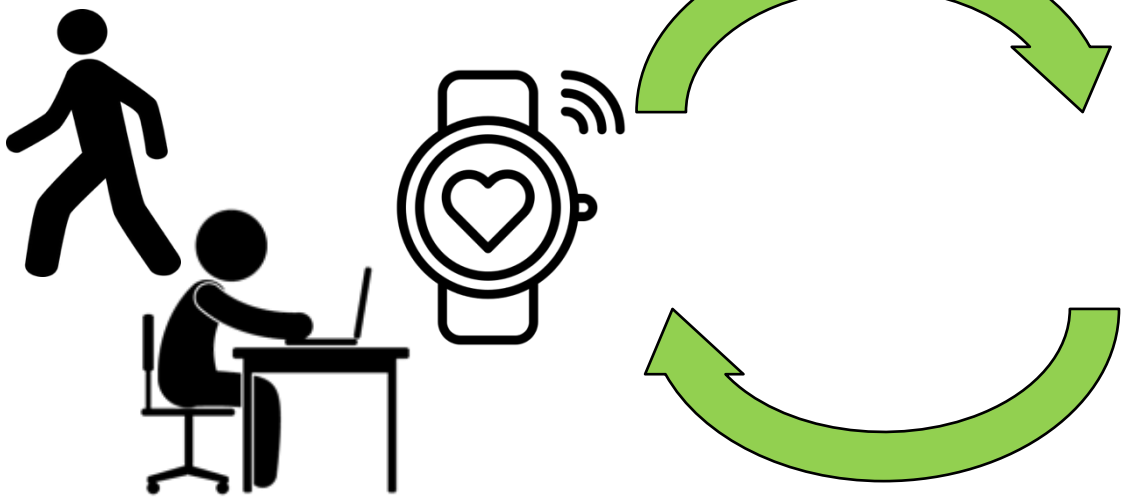
Klansja et al. *Ann. Behave. Med.*, 2019
Battalio et al. *Contemp. Clinical Trials*, 2021

The decision rule of a JITAI is an if-then statement

Use HeartSteps as an example

State: Weather, location, last 30-minute step count, …

IF



You've been sitting for 10 mins. It'd be great to get up and stretch your legs!

THEN

Action: Send an activity suggestion/Do nothing

The decision rule is often designed to help a person reach a long-term behavioral goal (**distal outcome**), e.g., staying physically active.

The intervention is adaptive to the state, but the decision rule is fixed.

Decision rules need to be further personalized because

- The decision rule may not be optimal for everyone.
- The decision rule may not be translatable across populations.
- The decision rule may not be optimal over time.

We can personalize a decision rule using reinforcement learning (RL).

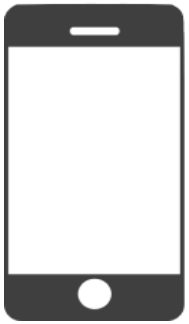
Reinforcement learning updates its decision rule as data accrue

State 1: Weather, location, last 30-minute step count, … => State 2

Decision rule 1 => Decision rule 2



Decision time 1 => 2



Action 1: Send an activity suggestion/Do nothing => Action 2

Reward 1 => Reward 2

The algorithm aims to learn the decision rule that optimizes (cumulative) reward.

How to choose the reward?

Distal outcome may not be ideal…

Distal outcomes (e.g., underlying activeness) are sparsely measured.

$\Rightarrow$Thus, it is hard to learn a decision rule with distal outcomes.

$\Rightarrow$Proximal outcomes (e.g., next 30-minute step count after a decision time) are chosen as a "proxy."

$\Rightarrow$The underlying assumption is that by improving the proximal outcome, the distal outcome can be improved too.

How to choose the RL algorithm?

Online RL in mHealth faces significant challenges

- Noisy data => Robustness
- Autonomy and Stability => Scalability
- Interpretability
- After-study analysis => Reproducibility

In a nutshell, at each decision time, TS chooses an action with the posterior probability that given the history of data, such an action is optimal.

Thompson Sampling (TS), a Bayesian RL, is chosen to tackle these challenges.

Consider a linear TS model for the proximal outcome

$$Y_{i,t} = \alpha_{i,0} + \alpha_{i,1}X_{i,t,1} + \cdots + \alpha_{i,M}X_{i,t,M} + A_{i,t}(\beta_{i,0} + \beta_{i,1}X_{i,t,1} + \cdots + \beta_{i,K}X_{i,t,K}) + \epsilon_{i,t}$$

$$\epsilon_{i,t} \sim N(0, \sigma_i^2) \qquad \theta_i = \{\alpha_{i,0}, \ldots, \alpha_{i,M}, \beta_{i,0}, \ldots, \beta_{i,K}, \sigma_i^2\}$$

i: i-th participant

t: t-th decision time

Y: proximal outcome (next 30-minute step count)

A: action (1: send a morning message, 0: no message)

X: state (weather, location, yesterday's step count, etc…)

$$Y_{i,t}(X_{i,t}, 1) - Y_{i,t}(X_{i,t}, 0) = \beta_{i,0} + \beta_{i,1}X_{i,t,1} + \cdots + \beta_{i,K}X_{i,t,K}$$

If treatment effect > 0, then action 1 is better.

Baseline model => Reduce variation for estimating β

TS sends an action with the probability that it is optimal

$$
\begin{aligned}
&Y_{i,t} \\
&= \alpha_{i,0} + \alpha_{i,1} X_{i,t,1} + \cdots + \alpha_{i,M} X_{i,t,M} \\
&+ A_{i,t}(\beta_{i,0} + \beta_{i,1} X_{i,t,1} + \cdots + \beta_{i,K} X_{i,t,K}) \\
&+ \epsilon_{i,t}
\end{aligned}
$$

$$\epsilon_{i,t} \sim N(0, \sigma_i^2) \qquad \theta_i = \{\alpha_{i,0}, \ldots, \alpha_{i,M}, \beta_{i,0}, \ldots, \beta_{i,K}, \sigma_i^2\}$$

Assign a prior to θ_i.

Say at time t, the history $h_{i,t} = \{x_{i,1}, a_{i,1}, y_{i,1}, \ldots, x_{i,t-1}, a_{i,t-1}, y_{i,t-1}\}$.

Then, send Action 1 with probability:

$$\mathrm{P}(\beta_{i,0} + \beta_{i,1} x_{i,t,1} + \cdots + \beta_{i,K} x_{i,t,K} > 0 | X_{i,t} = x_{i,t}, h_{i,t})$$

TS is a Bayesian method, so how do we choose the prior? How do we calculate the posterior?

Furthermore, how do we incorporate the relationship between proximal outcome and distal outcome into our RL algorithm?

Roadmap of part 2:

Careful design of a conjugate prior

A toolbox to enable behavioral scientists to design RL and run their design in their studies

---

An algorithm that incorporates both proximal outcome and distal outcome

Roadmap of part 2:

	Previous work	This work
Careful design of the prior	Used impoverished data to learn empirical prior	Provided a warm-start interpretation to allow for both subjective and empirical prior
A toolbox to enable behavioral scientists to design RL and run their design in their studies	No previous work **X**	Developed a toolbox to help behavioral scientists across disciplines design personalizing JITAIs
An algorithm that incorporates both proximal outcome and distal outcome	Only focus on the proximal outcome	Incorporating domain knowledge by considering a causal directed acyclic graph (DAG)

Roadmap of part 2:

	Previous work	This work
Careful design of a conjugate prior	Used impoverished data to learn empirical prior	Provided a warm-start interpretation to allow for both subjective and empirical prior

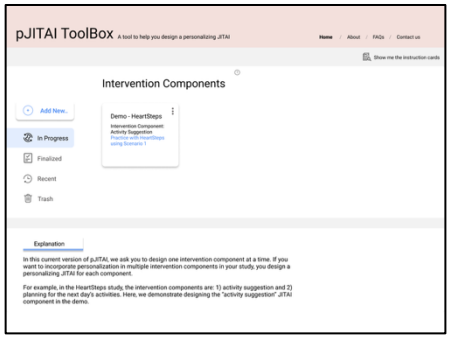
---

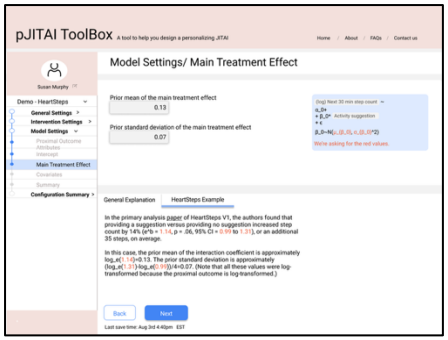
Key Contributions:

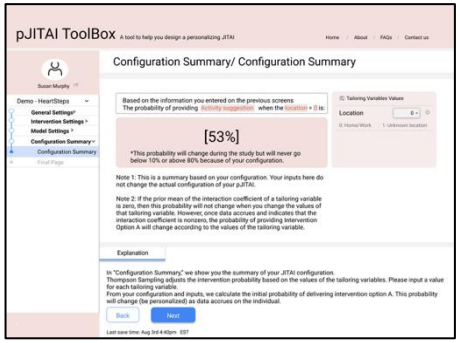
- Explain “prior” to behavioral scientists
- Interpret prior as a warm-start model
- Guide behavioral scientists on how to incorporate domain knowledge

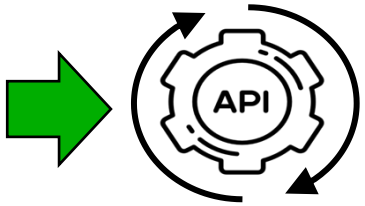
Interpretable and reproducible

Roadmap of part 2:









Previous work | **This work**

A toolbox to enable behavioral scientists to design RL and run their design in their studies	No previous work **X**	Developed a toolbox to help behavioral scientists across disciplines design personalizing JITAIs **In collaboration with Yuxuan Li**

Key Contributions:

Reproducible

- Enable behavioral scientists to design their RL algorithms for their personalizing JITAIs
- Provide an API to allow behavioral scientists to run their design

Roadmap of part 2:

Key Contribution: Interpretable and reproducible

- Design an RL framework for JITAIs that utilizes the causal relationship between proximal and distal outcomes to aim for long-term goals.

Ongoing Efforts: Scalable and robust

- Ensure the algorithm is automated and stable
- Ensure the algorithm is robust to a miss-specified causal DAG

	Previous work	**This work**
An algorithm that incorporates both proximal outcome and distal outcome (Ongoing)	Only focus on the proximal outcome	Incorporating domain knowledge by considering a causal directed acyclic graph (DAG) **In collaboration with Daiqi Gao**

Roadmap of part 2:

	Previous work	This work
Careful design of the prior	Used impoverished data to learn empirical prior	Provided a warm-start interpretation to allow for both subjective and empirical prior
A toolbox to enable behavioral scientists to design RL	No previous work **X**	Developed a toolbox to help behavioral scientists across disciplines design personalizing JITAIs **In collaboration with Yuxuan Li**
An algorithm that incorporates both proximal outcome and distal outcome	Only focus on the proximal outcome	Incorporating domain knowledge by considering a causal DAG **In collaboration with Daiqi Gao**

Consider a linear TS model for the proximal outcome

$$
\begin{aligned}
&Y_{i,t} \\
&= \alpha_{i,0} + \alpha_{i,1} X_{i,t,1} + \cdots + \alpha_{i,M} X_{i,t,M} \\
&+ A_{i,t}(\beta_{i,0} + \beta_{i,1} X_{i,t,1} + \cdots + \beta_{i,K} X_{i,t,K}) \\
&+ \epsilon_{i,t}
\end{aligned}
$$

$$\epsilon_{i,t} \overset{\text{i.i.d.}}{\sim} N(0, \sigma_i^2)$$

i: i-th participant

t: t-th day

Y: proximal outcome (total step counts per day)

A: action (1: send a morning message, 0: no message)

X: weather, location, yesterday's step count, etc..

To ensure automated and stable computation, a conjugate prior is assumed

$$Y_{i,t} = \alpha_{i,0} + \alpha_{i,1}X_{i,t,1} + \cdots + \alpha_{i,M}X_{i,t,M} + A_{i,t}(\beta_{i,0} + \beta_{i,1}X_{i,t,1} + \cdots + \beta_{i,K}X_{i,t,K}) + \epsilon_{i,t}$$

i: i-th participant

t: t-th day

$$\epsilon_{i,t} \overset{\text{i.i.d.}}{\sim} N(0, \sigma_i^2) \qquad \theta_i = \{\{\alpha_{i,m}\}, \{\beta_{i,k}\}, \sigma_i^2\}$$

Assign a conjugate prior for θ_i

$$\alpha_{i,m}|\sigma_i^2 \sim N(\mu_{\alpha,m}, \sigma_{\alpha,m}^2\sigma_i^2)$$
$$\beta_{i,k}|\sigma_i^2 \sim N(\mu_{\beta,k}, \sigma_{\beta,k}^2\sigma_i^2)$$

$$\sigma_i^2 \sim \textit{Scale-Inv-}\chi^2(\sigma_0^2, \nu)$$

At day t, given the observed data $h_{i,t}$, TS sends a morning message with probability: $P(\beta_{i,0} + \beta_{i,1}x_{i,t,1} + \cdots + \beta_{i,K}x_{i,t,K} > 0|h_{i,t})$

The hyperparameters should be interpreted in the context of a warmstart

i: i-th participant
t: t-th day

$$Y_{i,t} = \Phi(X_{i,t}, A_{i,t}, \pi_{i,t})\eta_i + \epsilon_{i,t},$$

$$\epsilon_{i,t} \sim N(0, \sigma_i^2),$$

$$\eta_i|\sigma_i^2 \sim N(\mu, \sigma_i^2\Sigma),$$

$$\sigma_i^2 \sim \text{Scale-Inv-}\chi^2(\nu, \sigma_0^2),$$

Given the observed data $h_{i,t+1}$, the posterior distributions are:

$$\eta_i|\sigma_i^2, h_{i,t} \sim N(\mu_{i,t}, \sigma_i^2\Sigma_{i,t}),$$

$$\sigma_i^2|h_{i,t} \sim \text{Scale-Inv-}\chi^2(\nu_{i,t}, \sigma_{i,t}^2),$$

$$\sigma_{i,t+1}^2 = \frac{1}{\nu + t}\{\nu\sigma_0^2 + t\frac{(y_{prev,i,t} - \phi_{i,t}\mu)^T[\phi_{i,t}\Sigma\phi_{i,t}^T + I]^{-1}(y_{prev,i,t} - \phi_{i,t}\mu)}{t}\}$$

Prior noise variance

Sample noise variance

Interpretation

σ_0^2: prior noise variance

ν: the weight on the prior noise variance

Larger ν, longer warmup

Less exploration, more exploitation

The hyperparameters should be interpreted in the context of a warmstart

i: i-th participant
t: t-th day

$$Y_{i,t} = \Phi(X_{i,t}, A_{i,t}, \pi_{i,t})\eta_i + \epsilon_{i,t},$$

$$\epsilon_{i,t} \sim N(0, \sigma_i^2),$$

$$\eta_i|\sigma_i^2 \sim N(\mu, \sigma_i^2\Sigma),$$

$$\sigma_i^2 \sim \text{Scale-Inv-}\chi^2(\nu, \sigma_0^2),$$

Given the observed data $h_{i,t}$, the posterior distributions are:

$$\eta_i|\sigma_i^2, h_{i,t} \sim N(\mu_{i,t}, \sigma_i^2\Sigma_{i,t}),$$

$$\sigma_i^2|h_{i,t} \sim \text{Scale-Inv-}\chi^2(\nu_{i,t}, \sigma_{i,t}^2),$$

Prior mean

$$\mu_{i,t+1} = (\Sigma^{-1} + \phi_{i,t}^T\phi_{i,t})^{-1}[\Sigma^{-1}\mu + \phi_{i,t}^T y_{prev,i,t}]$$

Sample mean

$$= \phi_{i,t}^T\phi_{i,t}(\phi_{i,t}^T\phi_{i,t})^{-1}\phi_{i,t}^T y_{prev,i,t}$$

Interpretation

μ: prior mean

Σ^{-1}: the weight on the prior mean

Larger Σ^{-1}, longer warmup

Less exploration, more exploitation

Summary

By interpreting the hyperparameters as a warm-start model, we enable behavioral scientists to:

- incorporate domain knowledge (using μ and σ_0^2)
- control the length of the warm-up period (using ν and Σ^{-1}) (which can depend on the confidence of the domain knowledge and the risk in the study)

I am writing a paper with behavioral scientists to submit to Psychological Methods, which should assist more behavioral scientists in designing their personalizing JITAIs.

Roadmap of part 2:

	Previous work	This work
Careful design of the prior	Used impoverished data to learn empirical prior	Provided a warm-start interpretation to allow for both subjective and empirical prior
A toolbox to enable behavioral scientists to design RL	No previous work **X** **High Level Overview**	Developed a toolbox to help behavioral scientists across disciplines design personalizing JITAIs
An algorithm that incorporates both proximal outcome and distal outcome (Ongoing)	Only focus on the proximal outcome	Incorporating domain knowledge by considering a causal DAG **In collaboration with Daiqi Gao**

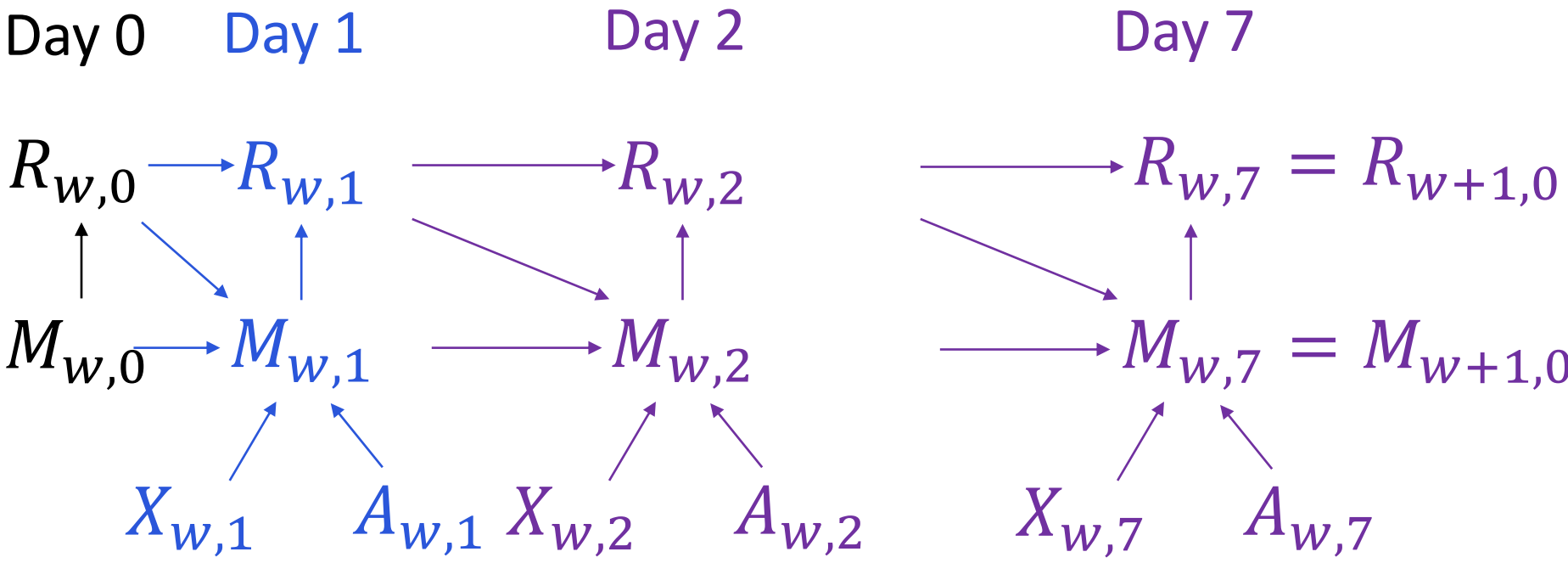
We consider a simplified ”HeartSteps” to help the discussion

HeartSteps



- Tailoring variables (State): Weather, location, …
- Intervention options (Actions): Send a morning motivational message/Do nothing
- Decision points (Decision time): 1 time a day
- Distal outcome: self-reported activeness
- Proximal outcomes: daily total step counts

Proximal outcome is on the causal pathway to the distal outcome



w: week

R: distal outcome (self-reported activeness)

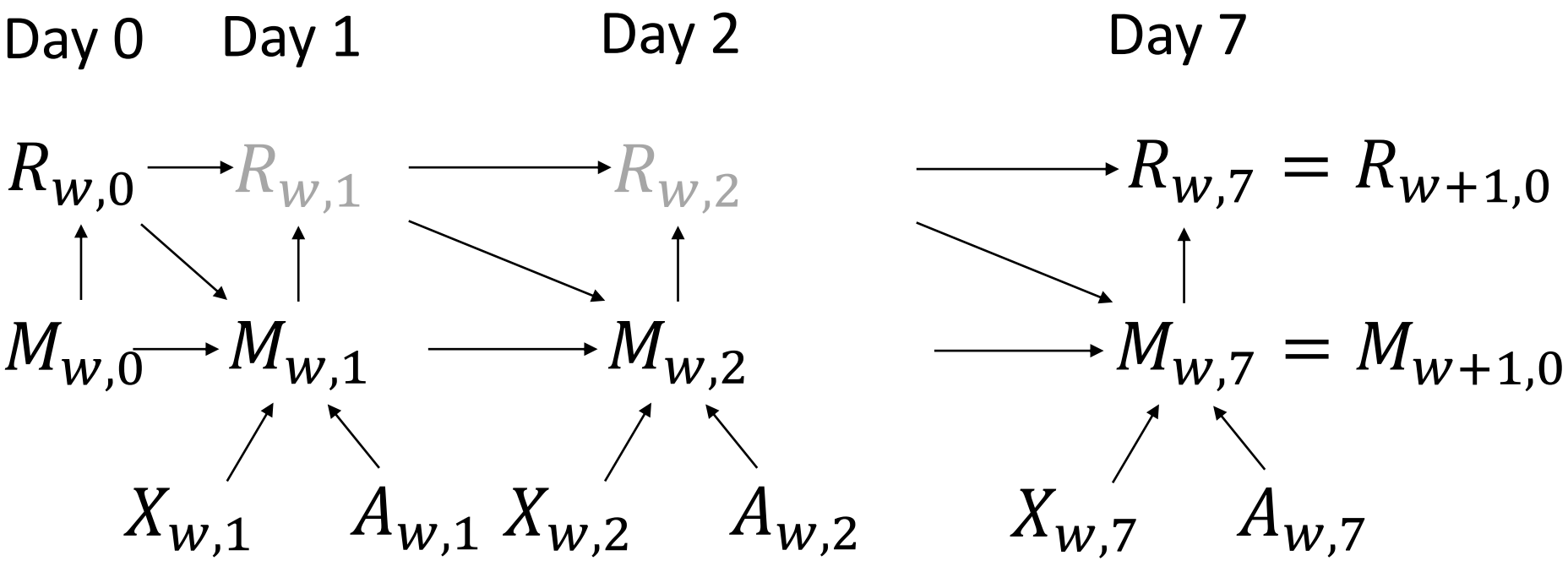
M: proximal outcome (total step counts per day)

A: action (morning motivational message)

X: weather, location, etc.

However, it is burdensome to measure distal outcome.

Distal outcome is sparsely observed

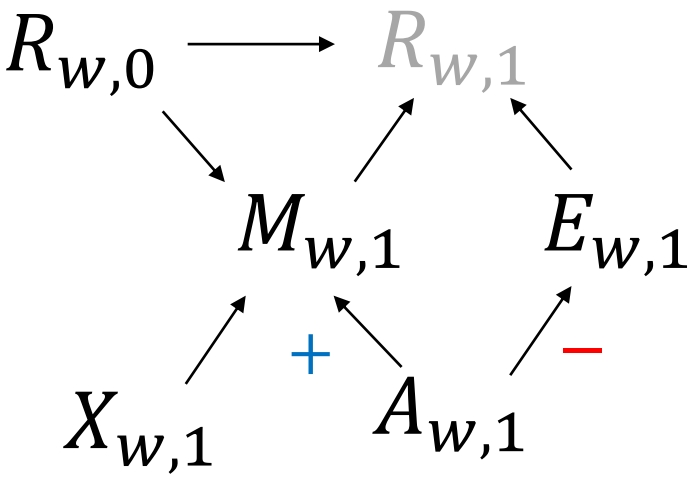


Most personalizing JITAIs only optimize the proximal outcome M. If there is only one mediator and $M \uparrow$ implies $R \uparrow$, then optimizing M also optimizes R.

But there are often multiple mediators.

Engagement is a crucial mediator

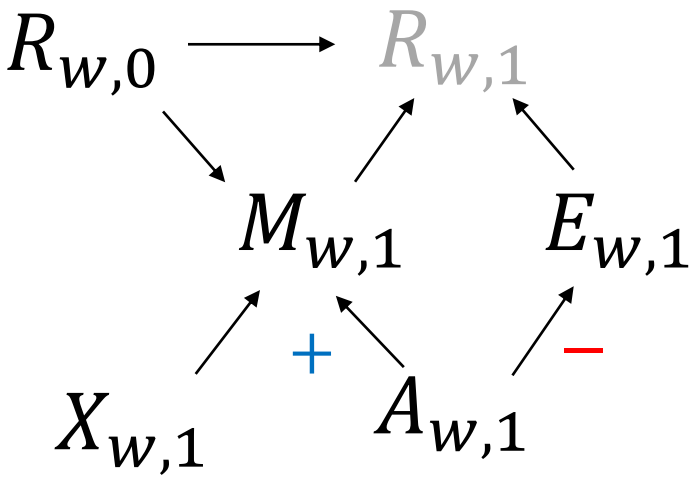
Too many interventions can lead to burden and reduce a participant's engagement.



E: engagement

We can use the DAG to learn the individualized weighting of the proximal outcome and the engagement.

The causal relationship between proximal and distal outcomes can be learned, which improves personalization



$$R_{w,1}(S_{w,1}, A_{w,1}) = \eta M_{w,1}(S_{w,1}, A_{w,1}) + \xi E_{w,1}(S_{w,1}, A_{w,1}) + f(S_{w,1}) + \epsilon$$

$$S_{w,1} = \{R_{w,0}, M_{w,0}, E_{w,0}, X_{w,1}\}$$

$$\Theta = \{\eta, \xi, \theta_M, \theta_E, \theta_f\}$$

- We can learn the posterior distributions of η and ξ per person.

$\Rightarrow$ More personalized and scientific.

- We use mediators to reduce variance in a sparse reward problem.

$\Rightarrow$ More efficient.

Hard to automate the learning of the posterior distributions of Θ

High-level pseudo code

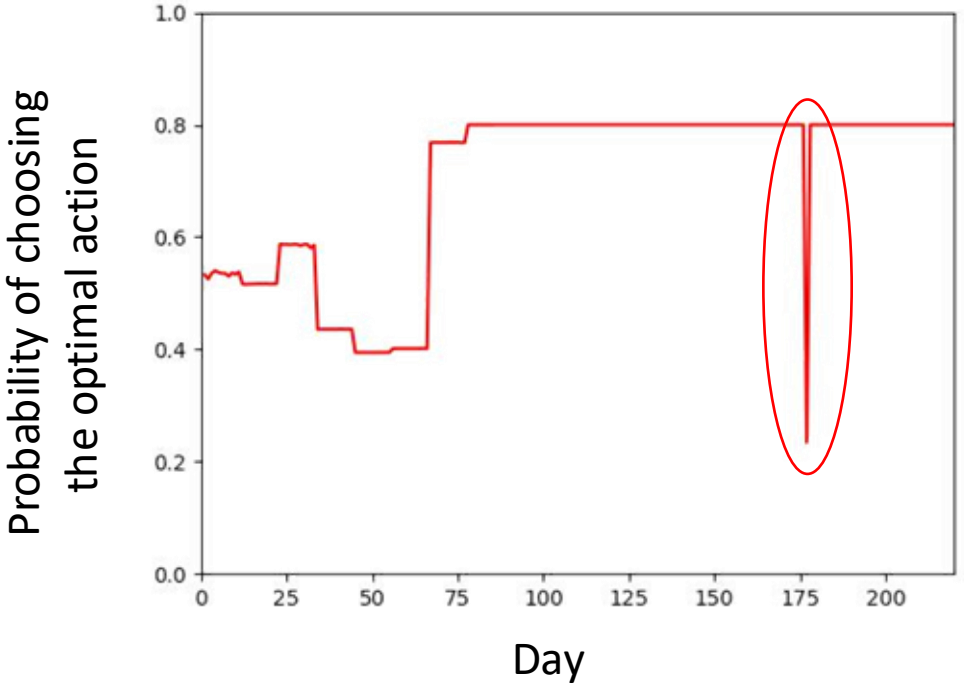
Assign prior distributions to the DAG parameters.

At decision time t = 1,…T,

- Observe the state.
- Use the posterior distributions of the DAG parameters to choose actions. (Use the prior distributions if t=1.)
- Observe mediators (and the reward)
- Update the posterior distributions of the DAG parameters.

Hard to automate the learning of the posterior distributions of Θ

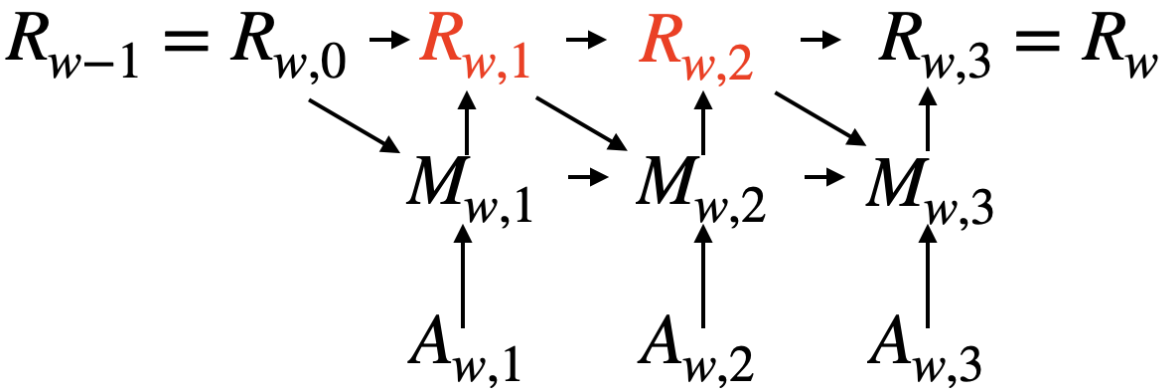
MCMC failed to run robustly without human inspection



We are currently exploring using Bayesian Bootstrap to learn the posterior distributions.

We are also exploring a joint model-based and model-free method to reduce the impact of model misspecification.

A joint model-based and model-free method may be more robust to model misspecification



- Model-based: At d=2, given a posterior sample θ, estimate $E\left[R_{w,2}\middle|R_{w,0}, M_{w,1}, A_{w,2} = 1; \theta\right] - E\left[R_{w,2}\middle|R_{w,0}, M_{w,1}, A_{w,2} = 0; \theta\right]$.

  Choose $A_{w,2} = 1$ if the estimate is greater than 0.

- Joint model-based and model-free:

  At d=2, sample a posterior sample of $R_{w,1}$ using the model.

  To choose $A_{w,2}$, consider a linear model $\Phi\left(r_{w,1}, M_{w,1}, A_{w,2}\right)\tau$.

=> Model is only used for imputation.

Bootstrap TS procedures

We are only considering this approach because we want stability over time.

Say $\theta \sim P$, $y|x \sim Q_\theta$.

Say at time t, we have observed $D_t = \{x_1, y_1, \ldots, x_t, y_t\}$.

The goal is to get N samples from the posterior distribution $\theta \mid D_t$.

Say there is a hyperparameters J.

This parameter determines the strength of the prior. However, this creates an extra hyperparameter to choose.

For n=1,…,N:

- Sample $\theta_n \sim P$
- Generate imaginary data $\tilde{D}_J = \{\tilde{x}_1, \tilde{y}_1, \ldots, \tilde{x}_J, \tilde{y}_J\}$ using θ_n.
- Bayesian bootstrap $D_t \cup \tilde{D}_J$ once.
- Get point estimate $\hat{\theta}_n$ using the Bayesian bootstrap sample

Approximate the posterior samples with $\{\hat{\theta}_n\}$

This posterior is inaccurate.

Summary

By considering both proximal outcome and distal outcome in the RL framework, we enable behavioral scientists to:

- incorporate domain knowledge (both in terms of the model and the prior for the causal DAG)
- find a decision rule that optimizes the distal outcome

Close the loop in healthcare: Bringing the benefits of machine learning back to the people

