Computational Personalization:
Data science methods for personalized health

prof. dr. Maurits C. Kaptein

Computational Personalization:

Data science methods for personalized health

1

Introduction

In the last decade authoritative scientific journals such as Science (Ng et al., 2009) and the New England journal of Medicine (Hamburg and Collins, 2010), as well as legislative bodies such as the American Food and Drug Administration (FDA) and the European Union (EU), have stressed the importance of *personalized healthcare.* By personalizing medical treatments, where the term *treatment* covers a broad range of interventions, from medication to education to eHealth, we can improve their effectiveness, decrease costs, and provide better care.

The idea that personalization is effective is based on what I would call—in modern methodological jargon—the existence of treatment effect *heterogeneity:* we believe that the effect of a specific treatment is different for different patients. In the last decades, driven by advances in a wide range of fields from genomics to medical imaging, the existence of treatment effect heterogeneity has been firmly established. To give a concrete example, in August 2011, the FDA approved the drug Zelboraf to treat metastatic melanoma (see, for example Chapman et al., 2011). Metastatic melanoma is a highly aggressive form of skin cancer with a low 5-year survival rate. Zelboraf is a drug that works by inhibiting a gene mutation, however, this mutation is only found in approximately half of the patients. Zelboraf is ineffective for those without the mutation. Luckily we can find the mutation, and we can accurately predict for which patients the treatment will be effective.

Examples such as Zelboraf that show that personalizing treatments can significantly improve their effectiveness. Indeed, the Zelboraf case demonstrates the benefits of *providing the right treatment to the right patient, at the right dose at the right time;* the very definition of personalized healthcare as used by the EU (Scholz, 2015). Regretfully, this definition is, in my opinion, hardly informative: it does not provide any guidance on how to make personalized healthcare a reality. Also, the definition does not clarify the meaning of the term "right"; this despite the excessive use of the word in a single sentence.

Inspired by the EU definition I will provide an alternative, and more constructive, definition of personalized healthcare. This alternative definition is useful since it allows us to pinpoint the key methodological and statistical challenges we face when trying to make personalized healthcare a reality. After providing this alternative definition, I will use it to analyze our current approach to per-

sonalizing treatments which is based on the randomized controlled trial (RCT). I will highlight both the advantages and disadvantages of this approach. Subsequently, I will formulate an alternative approach to choose the right treatment for the right patient. I will highlight why this novel method is both challenging and promising, and I will use it to formulate a research agenda for the field of "computational personalization".

2

A definition of personalized healthcare

In an attempt to redefine personalized healthcare, let us revisit the EU's definition: "Providing the right treatment to the right patient, at the right dose at the right time". Apparently both the patient and the time are important, as well as the choice of *treatment* and the associated dose. Thus, we can start setting up our problem by noticing that we are looking for some relationship, some *mapping,* between these and some associated health outcome. We can more formally denote this as

$$\{\text{patient}, \text{time}, \text{treatment}, \text{dose}\} \xrightarrow{f} \text{outcome}.$$

Here, admittedly, I am already inserting the idea that we need to interpret "right" with respect to some *outcome* of interest.

Based on my own experience personalizing e-Health applications, under the supervision of prof. Emile Aarts, prof. Panos Markopoulos, and prof. Boris de Ruyter, I will change this notation to read:

$$\text{outcome} \xleftarrow{f} \{\text{patient}, \text{time}, \text{treatment}, \text{dose}\}$$
$$r \xleftarrow{f} \{\text{patient}, \text{time}, \text{treatment}, \text{dose}\}$$
$$r \xleftarrow{f} \{x, a\}$$
$$r = f(x, a; \theta),$$

where, in the first line, I merely reorder the left and right hand side terms. In the second line I substitute "outcome" for the more convenient—because it is shorter—letter r which stands for *reward*. Next, I add some structure: the inputs of the mapping can be partitioned into two sets that are of separate interest:

1. The first set contains all the elements that we cannot control, often called the *context,* which I denote using the letter x. This set includes a description of the current patient and the state of the world at this point in time.

2. The second set contains all the elements that we can control, denoted using the letter a. These are the *actions* we can take, and this set contains of the treatment, the dose, and the timing.

I don't know whether EU agrees with this partitioning: by stressing the importance of the *right patient* they seem to imply that we can achieve personalized healthcare by selecting patients. I think we should treat every patient who arrives at our doorstep.

Finally, in the last line, I emphasize that the mapping that we are interested in can often be parameterized in some way; I am using θ to denote these parameters. My interpretation of this notation is broad: $f()$ can be an extremely flexible mapping, θ can have a very large dimension, and the inputs can be extremely diverse.

Let's assume that we know the mapping $f()$, and thus we know the exact outcome of every treatment for every person. In this case personalized healthcare simply boils down to doing the following:

$$\arg\max_a f(x, a)$$

which means nothing more than selecting the treatment that maximizes the outcome for a given patient.

This statement is a bit over-simplified: in actuality we interact with multiple people, often multiple times, and at each *interaction* we select the best action. Hence, the notation

$$\sum_{t=1}^{T} \arg\max_a f(x_t, a_t), \qquad (1)$$

where T is our total number of interactions, indicates that we aim to maximize the outcome over the whole population.

In the remainder of this talk I will take Equation 1 as very definition of personalized healthcare. Personalized healthcare is a simple, albeit possibly very high dimensional, maximization problem.

Admittedly, this is slightly abstract, so perhaps this is easier to follow when visualized. If we focus on a low-dimensional example we can visualize the relationship between the *context*, the *actions*, and the *rewards*. Figure 1 shows a possible relationship between the weight of a patient (the context in this example), the dosage of medication (the action), and the probability of survival (the reward). The Figure indicates that low weight patients require a low dosage of medication to be effective, and that too high a dose can lead to adverse effects.
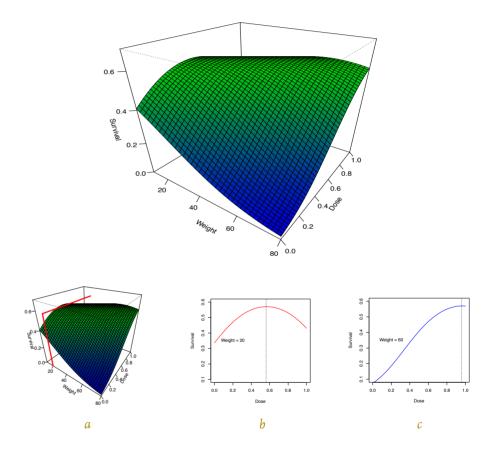

Figure 1: Simple visualization of a possible relationship between context, action, and rewards. This example displays the (hypothetical) relationship between the weight of a patient, the dosage (in Mg) of some medication, and the survival rates; clearly, for light patients the optimal dose is different than for heavy patients.

The panels on the bottom of Figure 1 illustrate the personalization challenge: when a child weighting 20 kilograms presents herself, effectively the context is fixed and hence we are looking at a 2d slice of the 3d plot. Subsequently, we can look at the possible dosages for this specific child, and we find that the optimal dosage choice is a bit over $^1/_2$Mg. If, at the next interaction, we are presented with an adult weighting in at 60 kilograms, we look at another slice of our plot and see that the optimal dose is close to 1Mg. Although this is a very simplified situation, this example illustrates that as long as we know the function that relates the context and the actions to the rewards, we can simply pick the action that leads to the highest outcome for every patient we encounter.

Now, what did we gain by our formalism? For one, the fancy mathematical notation to denote "the right treatment and the right time" makes us look scientific. However, we have also made some actual progress: we defined "right" in terms of maximizing some outcome and we split up our set of variables into those that we have under our control (the dose) and those that we do not have under our control (the weight). This is a methodologically important distinction. Also, we highlighted the sequential nature of personalization; we select treatments at each interaction. These notions allow us to better understand the problem that we are facing.

2.1 The challenges of treatment personalization.
In the stylized example we just looked at personalizing treatments seemed easy: we just compute which action has the highest reward. In reality however, personalizing is not easy. The most important reason that today most treatments are still "one- size-fits all" (Hamburg and Collins, 2010; Ng et al., 2009) is the simple fact that we, in actuality, do not know the relationship between between the context, our actions, and the resulting rewards. In short, $f()$ in Equation 1 is not known to us. This greatly complicates computing which treatment has the highest reward.

Since we do not know $f()$, we have to learn $f()$ using the inherently limited and often noisy data that we have at our disposal. Thus, we are not faced with a seemingly doable maximization problem, but rather we are faced with a challenging sequential learning problem: as we go along and treat patients we need to gradually learn which treatment is right for whom. This sequential learning is challenging for three reasons:

1. **High dimensional learning from noisy data:** The first challenge we face in developing personalized healthcare is that we need to learn $f()$ using limited and often noisy data. This learning problem is complicated by the fact that the *space of the problem is tremendous:* in practical terms this means that the relevant background characteristics of a patient are not just the weight, but rather the weight, age, genetic make-up, their culture, etc. etc. Similarly for the possible treatment; we do not just choose a dose, but we choose a combination of interventions, medicines, and treatments. Thus, any method to develop personalized healthcare needs to a) deal with the inherent uncertainty that arises from the limited number of observations that are available, and b) find an effective way to deal with the extremely large space of the learning problem.

2. **Learning causal relationships:** The second challenge is presented by the fact that what is learned from observational data might not properly reflect the knowledge we seek, namely, the effect of changing our treatments. To illustrate, suppose we currently, and naively, set out to model the relationship between chemotherapy (the action) and survival rates (the outcome) for breast cancer patients (the context) using existing registry data. In the observational data we will find that those who do not receive chemotherapy have a higher survival rate than those who do. However, this higher survival rate is not *caused* by refraining from chemotherapy; actually, patients with a mild tumor are both less likely to receive chemotherapy and more likely to survive. The relation present in the observational data is thus explained by a common cause and does not quantify the causal effect of the treatment. Since we need to learn a function that explicitly contains the effect of the "things that we can control", we need to be very careful about this distinction.

3. **Balancing learning and earning:** Thirdly, compared to so-called supervised learning—a fairly well understood machine learning task in which a computer learns a function between some observed input and some desired output (Hastie et al., 2013)—our problem is complex since we do not have any data regarding *the outcomes of actions that we have never actually tried out*. Hence, anytime we select a treatment, we need to balance choosing the best treatment as dictated by our current knowledge with the value of trying out new treatments that allow us to learn more about $f()$. This problem is

known as the "exploration-exploitation trade-off" or simply the "earning vs. learning problem". The problem arises because we have to learn $f()$ based on data with so-called "bandit feedback"; we do not observe what would have happened if we had administered another treatment (see, e.g., Ortega and Braun, 2013; Agrawal, 2012; Osband, 2015; Bastani and Bayati, 2015; Eckles and Kaptein, 2014).

These three problems, learning complex functions that properly model the causal effects of interest, based on bandit feedback, comprise the core data science challenges involved in personalized healthcare.

3

Our current approach to personalized healthcare: the RCT

Since we already have successful instances of personalized healthcare—such as the Zelboraf treatment for melanomas that I introduced earlier—we must have solved, or at least *addressed,* the challenges involved. Let us have a good look at how we currently address these problems.

In evidence based medicine today we find that the RCT constitutes our highest level of evidence (Evans, 2003; Grol and Grimshaw, 2003). The RCT is conceptually simple: randomly, for example by flipping a coin, we administer treatment A to half of our patients, and treatment B to remaining half. Next, after treating a pre-determined number of patients n in this way with either treatment A or B, we examine the outcome of interest in both groups. If, on average, in group A the outcome is higher then in group B, we select treatment A. For the dose finding example this would boil down to treating 100 patients with a low dose, say $^1/_2$Mg, while another 100 patients would receive a high dose, say 1Mg. Based on our example function given previously (see Figure 1), a naive RCT would conclude that the 1Mg dosage outperforms the $^1/_2$Mg dose, despite the adverse effects for children.

Let us examine in detail how the RCT addresses the three problems highlighted above:

1. **High dimensional learning from noisy data:** The RCT tackles the problem of high-dimensional learning from noisy data in two ways; first, the RCT heavily limits the problem space by pre-selecting a very small number of actions and contexts. The RCT compares only two treatments, and, only when the focus is on personalized healthcare, includes a very small number of descriptions of the context. When there is no focus on personalization the context is fully ignored. Exactly which treatments and which contexts to focus on is determined by our *theoretical* understanding of the process involved. Second, after limiting the problem space based on our existing theories, RCTs use a fairly simple method of dealing with noise; if, assuming that the two treatments have the exact same outcome, the actually observed, or a more extreme outcome is unlikely—quantified using the magical p-value that some of you might be familiar with—we reject the null hypothesis that the treatments are equally effective, and adopt whichever treatment had the highest average outcome in the trial.

2. **Learning causal relationships:** The RCT tackles the problem of learning the causal effect of the actions by virtue of its use of randomization. By "flipping the coin" we determine who receives which treatment, and we make sure that this treatment assignment is not confounded by patient characteristics such as the "severity of the tumor" as in the breast-cancer example.

3. **Balancing learning with earning:** To appreciate how the RCT solves the last challenge, we have to view the RCT not just on its own, but we have to include the treatments that are administered *after* the RCT has been carried out. For example, after the Zelboraf trail, we now routinely treat melanoma's using Zelboraf. Approached in this way we can see that the RCT balances learning and earning by first spending a pre-determined number of interactions on learning (the trial itself), and subsequently moving to earning: after the trial, the results are accepted with full certainty, and future patients will receive the treatment that performed best during the trial.

At this point I have to note that the RCT is not inherently a method for personalization; rather, it is a method for selecting one out of two competing treatments. However, by doings RCTs within subgroups of patients—for example within all children with a low weight—this method is now the gold standard to select treatments for specific subgroups of patients.

3.1 Advantages of the RCT

Now that we understand how the RCT addresses our three challenges, we can evaluate the quality of this approach. Let me start by discussing the strengths of the RCT.

The RCTs approach to high dimensional learning is appealing since by severely restricting the space of actions and context the outcomes of the trial become *transparent* and human-understandable. While obviously the quality of our restrictions of space depend heavily on the quality of the theories that we use—something that I fear is hard to assess—the outcomes of the an RCT are at the very least easily interpretable: the survival rate in the patient group that received 1Mg was higher than in the group that received $^1/_2$Mg, and hence you get 1Mg. Next, the RCTs approach to the problem of learning causal relationships is extremely solid (Rubin, 1978; Imbens and Rubin, 2015). There is no better method to assess causal effects than randomization, which is exactly what the

RCT excels at. Finally, the RCT's approach to balancing earning vs. learning is practically appealing: by moving all the learning to the beginning, into the trail, and all the earning to the resulting guidelines, we make a nice and convenient deterministic choice.

3.2 Disadvantages of the RCT

Our analysis also allows us to identify drawbacks of the ICT. First of all, the singling out of very small subsets off all possible actions and context in *sequential* RCTs—since in actuality we build our knowledge one RCT at a time—basically constitutes a limited and naive strategy for learning $f()$. We effectively assume that only very small parts of the context and treatment are important and we ignore all others. Already in our simple weight-dose example introduced earlier, the RCT would only examine a small number of specific points in the 3d space, as opposed to examining or modeling the whole plane of outcomes. Furthermore, perhaps implicitly, we assume that the relationship between context and actions is only as complex as our theories allow us to understand.

Another disadvantage of the RCT originates from our insistence on a hard cut-off between learning and earning. The RCT—and the deterministic decision strategy inspired by the null hypothesis significance test—leads us to either adopt or ignore a new treatment, possibly for some subgroup of people, with *certainty*. However, these certain decisions are made based on noisy data, and hence full certainty is too much to ask. Given limited and noisy data there is always a non-zero probability of making the wrong choice. And, the more we try to personalize treatments, the more severe this problem becomes since at the level of small groups of patients we have very limited data at our disposal. If we truly believe in treatment heterogeneity, then we have to accept that each patient is unique and hence we will never have a large homogenous sample available to make deterministic decisions.

Regretfully, this not the last disadvantage of the RCT as a method of solving Equation 1; because of our determinism, the data that we collect after a trial also turn out to be very hard to re-use: once the probability of receiving chemo-therapy for breast cancer patients with a severe tumor is 1, and for those with a mild tumor is 0, we cannot use the future data to evaluate alternatives simply because no such data is collected. Our deterministic decisions prohibit our future learning.

4

A sketch of an alternative: a computational approach to personalization

I would like to sketch a possible alternative method to the RCT. Note that I will only provide an intuition for this alternative method; some technical details are provided in the footnotes of the transcript of this talk.

I propose to do the following: First of all, I propose to use a modern and flexible machine learning model to learn the relationships between the actions, context, and rewards. In recent years we have seen a revolution in our abilities to learn flexible, extremely high-dimensional functions (Hastie et al., 2013; Pratola et al., 2016; Mohammadi and Kaptein, 2016; Bishop, 2006), and hence there no need to artificially reduce the model space by focusing on very small numbers of patient or treatment characteristics.

Second, we can utilize novel breakthroughs in our understanding of causality; as it turns out, it is strictly not necessary to resort to uniform random allocation as is done in the clinical trial. Rather, as long as we can compute and store the probability of receiving a treatment conditional on the patient characteristics, we can use the collected data to estimate causal effects (Bang and Robins, 2005; Funk et al., 2011).

Finally, we can use novel methods of balancing earning and learning: as opposed to going instantly from pure learning to a deterministic choice as in the RCT, we can gradually balance the two. An allocation scheme called Thompson sampling allows us to, over time, gradually change the probabilities of receiving different treatments. Thompson sampling selects treatments with a probability that is proportional to its effectiveness. Thus, as we gain more evidence that an action is effective, we will increase the probability of selecting it.

This way we can optimally balance exploration and exploitation (Ortega and Braun, 2013; Osband, 2015; Eckles and Kaptein, 2014).[1]

This computational approach to treatment personalization can be realized by, every time we visit a doctor (or go to a website for health information, or use an motivational eHealth application), sending our data—the *context*—to a central server.

Next, this central server estimates a model that relates the context, the actions, and the rewards. This model is our estimate of the illustrious function $f()$ in our definition of personalization. Finally, the central server selects an action

based on this model while balancing learning and earning. Note that as a result of this method we never make a definite choice between different treatments. However, we do make the best choice we can given all the information available.

Admittedly, this computational method to personalization might look a bit distant from reality, but the models I propose, and the methods by which earning and learning can be balanced, are already, at least conceptually, developed. Also, we can already transmit large amounts of data around the world in a split second; large web companies like Facebook and Google do this constantly. Hence I believe that, in the near future, my suggestion is *technically* feasible.

4.1 Disadvantages of computational personalization
Contrary to the RCT, I will start by discussing the disadvantages of my computational approach to personalization. Two disadvantages easily come to mind, the first being "which variables, thus which contexts and which actions, should we include in such a gigantic machine learning model?", And the second "which outcomes should we actually care about?" I believe these are genuine questions, but they are not disadvantages of the method: these questions equally need answers when designing an RCT. Actually, my proposed approach allows for much greater flexibility than the RCT: we can include a larger number of contextual variables and we can potentially collect data regarding multiple outcomes. Thus, if anything, my proposal makes answering these questions easier as opposed to harder. However, there are more serious concerns: First of all, my proposed approach looses, at least superficially, all notions of *transparency*. It is not at all clear anymore why a specific patient, at some specific point in time, receives a specific treatment. This will be hidden away in some "black-box" learning model. While the underlying logic can theoretically still be distilled from the model parameters, such distilling is not easy. And, by loosing transparency, we probably also loose *accountability;* if we don't know why we are subscribing some treatment, than who should we hold responsible in case of a calamity?

Next, the proposed method, at least in theory, never leads to a definite, deterministic, choice. Hence, there will always be a non-zero probability of receiving a specific treatment. This might be fine for things like eHealth coaching and health education, but we will be presented with a logistic nightmare if we intend to keep all possible pills available at all pharmacies all around the world for the unlikely event that we should administer one of them.

By abandoning the RCT assessing causality becomes more challenging. How can we still be sure that the model we learn is actually learning the effects of our treatments, and not learning some spurious, non-causal, relationship? In recent decades this problem has however largely been solved (Bang and Robins, 2005; Funk et al., 2011; Pearl, 2009; Imbens and Rubin, 2015): we have recently come to realize that as long as we know the probability of receiving a treatment, we can validly estimate causal effects even when treatments are not uniformly randomized.

Finally, implementing computational personalization at the scale that I am suggesting will not be easy; the underlying models and methods are still being developed, and many details are not yet finished. For example, we need to be able to deal large volumes of dependent data that are collected continuously; a technical topic my recently graduated PhD student Lianne Ippel has made a large contribution to (Ippel et al., 2016b,a). Furthermore, we need the infrastructure to make all of this technically work; recently Jules Kruijswijk and Robin van Emden have gone through great lengths to build an open source platform that allows us to do exactly this, but it needs further development (Kaptein and Kruijswijk, 2016; Kaptein et al., 2016). Next, we need to develop methods to fit these models faster, on large datasets; work that is currently being done by my colleagues and collaborators Matthew Pratola and Reza Mohammadi (Pratola et al., 2016; Mohammadi et al., 2015). We also need to understand much better how we can combine multiple outcomes measures into a single reward; a problem Xynthia Kavelaars will be contributing to in her PhD project. This is a promising project that I am honored to supervise together with dr. Joris Mulder.

4.2 Advantages of computational personalization
By now you might wonder why I have bothered to propose this new method. My proposal seems plagued with challenges and needs lots of work; probably enough work to keep me and my PhD students busy for the next few years. Cynically, you could imagine that I propose this method precisely because I want to keep myself and my PhD students busy, but this is not the core motivation. My actual motivation comes from the advantages of the method. Or, to be more precise, *its single advantage:* with this method we will have a better outcome. Now that's a bold statement, and one that I cannot quantify for the scale at which I am suggesting the method to be used. The number of future interactions, the number of possible actions, and the number of meaningful contextual factors is

simply too large to say anything precise. However, at smaller scales, for simple versions of the personalization problem, we can quantify the benefits.

The performance of a personalization method can be measured in terms of its regret: the realized outcome of a method compared to the outcome we could have achieved with full information. Suppose we compare the RCT to my proposal in a simple case in which we choose one of two possible treatments for 1000 (homogenous) patients, and where the true probabilities of success are .4 and .5. In the worst case we would obtain an expected 400 successes, while in the best case we expect to obtain 500 successes. Thus, a strategy that always selects the poorest treatment obtains a regret of 100, while randomly picking treatments results in an expected regret of 50. In this setting, the RCT has an expected regret of about 36, while my proposal weighs in at about 12; a difference of 25 successes as shown in Figure 2a. This difference results from a better balancing of earning and learning. Furthermore, the difference is magnified when we include a context and focus on smaller and smaller groups of patients; this is exactly what we want to do when personalizing our treatments.
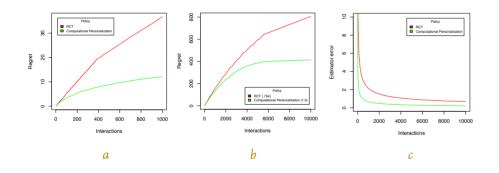

Figure 2: Three panels showing the superior performance of computational methods for personalization and the RCT. See main text for explanations of each panel.

Scaling the problem to 10.000 decisions and 10 possible treatments (with success probabilities .5 and .4 for the best two, and .3 for those remaining), the superior performance of computational personalization is even more striking: the regret of the RCT is 800, while that of computational personalization is only 400, as displayed in Figure 2b. Even more interestingly, the practice of sequential, binary RCTs identifies the best treatment in only 3/4 of the cases while for my proposed method the probability of finding the best treatment converges to 1. This latter difference is caused by stepping away from simple binary tests to learn a complex relationship, as is the case with the RCT, towards examining and comparing multiple treatments in one go. Also this difference is magnified when we consider personalized treatments since the more we expand the context-action space, thus, the more characteristics of the patient or the treatment we consider, the poorer the performance of the RCT will be.

Finally, as long as we store the probabilities of receiving a specific treatment conditional on the context, we can effectively re-use the data that we collect; something that is almost impossible when using RCTs. A recent theoretical analysis by Agarwal et al. (2016) shows that such re-use of the data reduces estimation errors of our models by orders of magnitude. Figure 2c shows the estimated standard errors as a function of the number of datapoints collected using the different methods. Simply put, using a computational approach to personalization allows us to learn more efficiently than using repeated RCTs. These simple computations show that the RCT is grossly outperformed by my suggested alternative. Furthermore, it is reasonable to expect that the RCT will comparatively suffer more from making the problem more realistic than the method I propose. Thus, if anything, the presented differences in expected outcomes are underestimates of the actual outcomes rather than overestimates.

5
Conclusion

I am well aware that I have just introduced a fairly abstract alternative to the RCT as a means of advancing knowledge and making decisions in the health and life sciences. Obviously, I understand that making a change to the fundamental way in which we develop our knowledge is tremendously scary. However, I hope that by now you are convinced that recent advances in research methods, statistical learning, and data science, have provided us with methods for personalizing treatments that will undoubtedly safe lives compared to our current practice, solely at the costs of transparency and accountability. For many, these costs are too big to bear. These researchers cling to the RCT as the only valid and understandable way of advancing our knowledge. I disagree; I think we should actively examine radically different alternatives. We should not refrain from using new computational methods because they have challenges, but rather we should try to address these challenges. Bluntly put, sticking to the RCT as the only means of realizing personalized healthcare in a day and age in which we have the technical and methodological tools at our disposal to grossly outperform the RCT, is unethical.

In this talk I have tried to honestly display that the alternative, computational, approach to personalization is still in its infancy. We need to develop it further. Because of this, I am tremendously honored that with the generous help of CZ health-insurance, and with talented and motivated PhD students such as Bas Willemse, Ylva Hendriks, Jules Kruijswijk, Robin van Emden, and Xynthia Kavelaars, we will be working on making computational personalization a reality in the context of eHealth interventions. We will focus specifically on eHealth since in this application area the outcomes of interest are relatively easily measured, and the treatments, consisting of the feedback provided on the screens of users, is often easily and cheaply experimented with. I truly hope that in the next 4 to 5 years we will be able to provide a convincing proof of concept that, using computational personalization, we can indeed be more effective than using our current standards. And, if we are, I hope that our novel data science methods to personalize eHealth ultimately allow us to improve healthcare in general.

6
Thanks

It is a custom to take some time at the end of a speech like this to thank all of those who have provided support throughout the years. However, so many people have supported my work and inspired my thinking that these are too many to mention by name; any attempt to do so in a reasonable time would undoubtedly leave out mentors and colleagues that have been of great importance to me.

So, I am going to keep this very brief, next to my PhD students and my former supervisors, whose names I have already mentioned during this talk, I want to explicitly thank prof. Jeroen Vermunt for his continued support and for providing myself, and the whole research methods and statistics group at Tilburg University with a great place to work. I also would like to thank prof. Arjan van der Born for providing me with the amazing opportunity to start a research lab at JADS, and for his continuous support of my research line. I want to thank my inspiring collaborators, I guess you know who you are, and I want to thank my parents for the amazing job they did raising my brother and myself, and for instilling the continuous urge to be critical, skeptical, and weary of authority, traits that would likely get me in trouble everywhere but in a challenging academic environment. Finally, Rosa, thanks.

I have spoken.

References

Agarwal, A., Bird, S., Cozowicz, M., Hoang, L., Langford, J., Lee, S., Li, J., Melamed, D., Oshri, G., Ribas, O., Sen, S., and Slivkins, A. (2016). A multi-world testing decision service. *CoRR,* abs/1606.03966.

Agrawal, S. (2012). Further Optimal Regret Bounds for Thompson Sampling. *CoRR,* none:1–14.

Bang, H. and Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics,* 61(4):962–973.

Bastani, H. and Bayati, M. (2015). Online Decision-Making with High-Dimensional Covariates. (i):1–47.

Bishop, C. M. (2006). *Pattern recognition and machine learning.* springer.

Chapman, P., Hauschild, A., Robert, C., Larkin, J., Haanen, J., Ribas, A., Hogg, D., O'Day, S., Ascierto, P., Testori, A., et al. (2011). Phase iii randomized, open-label, multicenter trial (brim3) comparing braf inhibitor vemurafenib with dacarbazine (dtic) in patients with v600ebraf-mutated melanoma. *Journal of Clinical Oncology,* 29(18 suppl):LBA4–LBA4.

Chipman, H. A., George, E. I., McCulloch, R. E., et al. (2010). Bart: Bayesian additive regression trees. *The Annals of Applied Statistics,* 4(1):266–298.

Eckles, D. and Kaptein, M. C. (2014). Thompson sampling with the online bootstrap. *arXiv preprint arXiv:1410.4009.*

Evans, D. (2003). Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of clinical nursing,* 12(1):77–84.

Funk, M. J., Westreich, D., Wiesen, C., Stu rmer, T., Brookhart, M. A., and David- ian, M. (2011). Doubly robust estimation of causal effects. *American journal of epidemiology,* 173(7):761–767.

Grol, R. and Grimshaw, J. (2003). From best evidence to best practice: effective implementation of change in patients' care. *The lancet,* 362(9391):1225–1230.

Hamburg, M. A. and Collins, F. S. (2010). The path to personalized medicine. *N Engl J Med,* 2010(363):301–304.

Hastie, T., Tibshirani, R., and Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* volume 11. Springer Science & Business Media.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

Ippel, L., Kaptein, M. C., and Vermunt, J. (2016a). Dealing with data streams. *Methodology.*

Ippel, L., Kaptein, M. C., and Vermunt, J. K. (2016b). Estimating random-intercept models on data streams. *Computational Statistics & Data Analysis,* 104:169–182.

Kaptein, M. C., Iannuzzi, D., et al. (2016). Tracking the decoy: maximizing the decoy effect through sequential experimentation. *Palgrave Communications,* 2.

Kaptein, M. C., and Kruijswijk, J. (2016). Streamingbandit: Developing adaptive per- suasive systems. *arXiv preprint arXiv:1602.06700.*

Kaufmann, E., Korda, N., and Munos, R. (2012). Thompson Sampling: An Asymp- totically Optimal Finite Time Analysis. *Algorithmic Learning Theory,* (1):15.

Mohammadi, A. and Kaptein, M. C. (2016). Contributed discussion on article by pratola [comment on" mt pratola, efficient metropolis-hastings proposal mechanisms for bayesian regression tree models"].

Mohammadi, A., Wit, E. C., et al. (2015). Bayesian structure learning in sparse gaussian graphical models. *Bayesian Analysis,* 10(1):109–138.

Ng, P. C., Murray, S. S., Levy, S., and Venter, J. C. (2009). An agenda for personalized medicine. *Nature,* 461(7265):724–726.

Ortega, P. and Braun, D. (2013). Generalized Thompson Sampling for Sequential Decision-Making and Causal Inference. *arXiv preprint arXiv:1303.4431.*

Osband, I. (2015). Bootstrapped Thompson Sampling and Deep Exploration. pages 1–10.

Pearl, J. (2009). *Causality.* Cambridge university press.

Pratola, M. T. et al. (2016). Efficient metropolis–hastings proposal mechanisms for bayesian regression tree models. *Bayesian analysis,* 11(3):885–911.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of statistics,* pages 34–58.

Scholz, N. (2015). Personalised medicine; The right treatment for the right person at the right time. Technical report.

Thompson, W. R. (1933). On the Likelihood that one Unknown Probability Exceeds Another in view of the Evidence of two Samples. *Biometrika,* 25(3-4):285–294.

7
End notes

[1] Let me briefly give some more details to those who have a background in data science; in practice, I propose to use a flexible machine learning model such as the Bayesian additive regression tree model, or shorthand BART, which can be denoted as follows:

$$r = \sum_{j=1}^{m} g(x, a; \theta) + \epsilon.$$

In this model the rewards are modeled as a function of both the context and the actions using a sum over m binary decision-tree models (see Chipman et al., 2010, for details on BART). Trees provide a flexible modeling approach that can handle a wide variety nonlinear relationships and a large number of inputs.

While admittedly at this point in time neural network models seem to be more popular, the Bayesian specification of the BART model allows one to fully quantify ones posterior belief regarding the parameters of this model based on our prior beliefs and the likelihood of the data using Bayes theorem

$$P(\theta_{t+1}|D_{t+1}) \propto P(x_{t+1}, a_{t+1}, r_{t+1}|\theta_t) P(\theta_t|D_t),$$

where using Dt I denote the full dataset up to time t and hence theoretically provides a principled way for sequential learning. Furthermore, this model effectively guards against overfitting by virtue of a prior restricting trees of large depth, and allows us to balance earning and learning using a method known as Thompson sampling (Thompson, 1933; Eckles and Kaptein, 2014; Kaufmann et al., 2012; Ortega and Braun, 2013); When applying Thompson sampling we choose our actions with a probability proportional to our posterior belief that the action is the best action which can be formalized as:

$$\int \mathbb{1}\left[\mathbb{E}(r|a, \theta) = \max_{a'} \mathbb{E}(r|a', \theta)\right] P(\theta|\mathcal{D}) d\theta$$

Thompson sampling provides an easy to implement, and asymptotically optimal method of balanc- ing earning and learning. As long as we apply this scheme, and store the probability of receiving an action at each point in time for each patient (know as the propensity-score), we can not only use this scheme to select actions, but also to create a dataset that we can re-use in offline evaluations of alternative decision policies.

Computational Personalization: data science methods for personalized health