Personalization in biomedical-informatics: methodological considerations and recommendations

Prof. dr. Maurits Kaptein

Jheronimus Academy of Data Science & Tilburg University
Statistics and Research Methods
Sint Janssingel 92
5211 DA 's Hertogenbosch, the Netherlands
m.c.kaptein@uvt.nl

Abstract

Over the last decades there has been an increasing interest in personalization: can we make sure that treatments are effective for individual patients? The quest for personalization affects biomedical informatics in two ways: first, we design systems—for example eHealth applications—that directly interact with patients and these systems might themselves one day be personalized. Hence, we seek effective methods to do so. Second, we design systems that collect the data which will one day be used to personalize treatments: hence, we need to critically consider design requirements that improve the utility of (e.g.,) personal health records for future treatment personalization. By clearly defining personalization and analyzing the effectiveness of different personalization methods this discussion highlights how we should embrace sequential experimentation—as opposed to the traditional randomized trial—if we want to personalize our informatics systems efficiently. Furthermore, we need to make sure that we capture the treatment assignment process in our health records: doing so will greatly increase the utility of the collected data for future personalization attempts.

Keywords: Personalization, Health Information Systems, Research Methods, Sequential Experimentation, Personal Health Records

1. Introduction

In the last decade authoritative scientific journals such as Science [1] and the New England journal of Medicine [2], 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 the existence of treatment effect *heterogeneity*: we generally 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 3]. 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 [4].

Regretfully, this definition is not very informative: it does not provide any guidance on how to make personalized healthcare a reality. This discussion examines an alternative, and more constructive, definition of personalization in the health and life sciences. 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. The subsequent analysis highlights the two ways in which future personalization attempts will affect research in biomedical informatics:

1. We design and evaluate systems that directly interact with patients [see, e.g., 5, 6, 7]. These systems are increasingly personalized. Hence, we seek effective (statistical) methods to personalize treatments — i.e., to select a treatment for an individual given the available data — and we seek methods to evaluate the effectiveness of personalize systems. This discussion argues that our current reliance on randomized controlled trials (RCTs) for these purposes is ineffective and we should explore sequential and adaptive experimentation methods.
2. We design and evaluate systems that collect, contain, and combine health data [see, e.g., 8, 9]. Whether or not the data in these systems can be used for effective personalization [10] depends heavily on the data we store and the ways in which we make this data accessible. In this discussion we argue that we need to store all treatment *actions* and their respective probabilities, as well as the associated outcomes, to ensure that our collected data will be useful for future personalization attempts.

The next section provides a definition of treatment personalization and uses it to analyze the current approach to personalization which is largely based on the randomized controlled trial (RCT). Subsequently, an alternative, sequential, approach to choose the right action for the right patient is formulated and evaluated. Finally the impact of sequential methods for treatment personalization on biomedical information systems and personal health data is discussed.

2. A general definition of personalized healthcare

In an attempt to better understand personalized healthcare and the possible challenges for biomedical informatics, 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, personalization can be setup more formally by noticing that we are looking for some relationship, some *mapping*, between the patient and the treatment on one hand, and some associated health outcome on the other hand. This can denoted this as

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

Based on extensive experience personalizing e-Health applications [see, e.g., 11, 12], this notation can be changed to read:

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

where, the first line merely reorders the left and right hand side terms. In the second line substitutes "outcome" for the letter r which stands for *reward*. Next, some structure is added: 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 will 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.

Finally, the last line emphasizes that the mapping that we are interested in

can often be parameterized in some way; θ is used to denote these parameters. This notation should be interpreted broadly: $f()$ can be an extremely flexible mapping, θ can have a very large dimension, and the inputs can be extremely diverse.

If we assume that we know the mapping $f()$, and thus know the exact outcome of every treatment for every person, personalized healthcare boils down to doing the following:

$$\arg\max_{a} f(x, a).$$

This 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

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

where T is our total number of interactions, indicates that we aim to maximize the outcome over the whole population. Let us take Equation 1 as the *definition of personalized healthcare*. Hence, abstractly, personalized healthcare is a simple, albeit possibly very high dimensional, maximization problem.

Focussing on a low-dimensional example, the relationship between the context, the actions, and the rewards can be visualized. 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. The three panels on the right of Figure 1 illustrate the per-

sonalization 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 $\frac{1}{2}$. 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 1. 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.

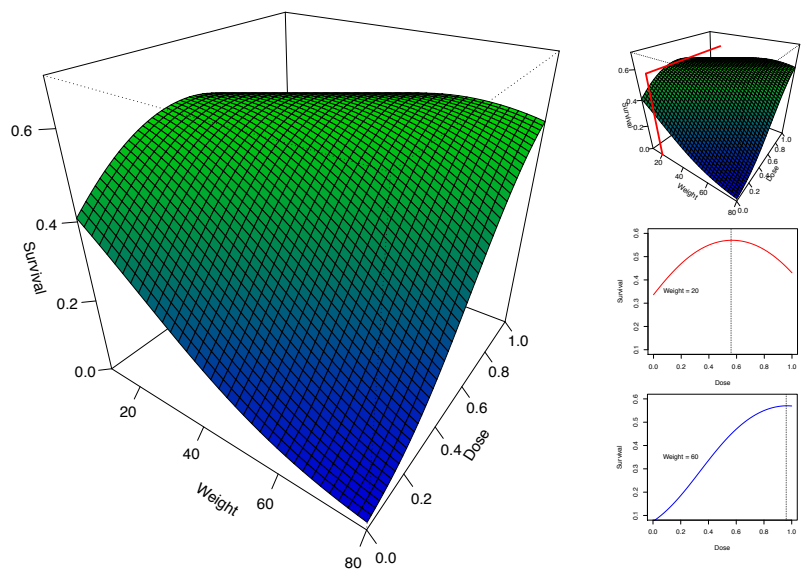


Figure 1: Simple visualization of the possible relationship between context, action, and rewards. This example display the (hypothetical) relationship between the weight of a patient, the dosage (in Mg) of some medication, and the survival rates; clearly, for light patients the optimal dose is different than for heavy patients. The smaller Figures on the righthand-side of the Figure illustrate different slices associated with different contexts.

The above definition formalized "right" in terms of maximizing some outcome and it 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), which is a methodologically important distinction. Also, the definition highlights the sequential nature of personalization: we select treatments at each interaction that we have with patients. These notions jointly allow us to better understand the problem that we are facing when personalizing biomedical informatics systems and collecting the data necessary for future personalization.

2.1. The challenges of treatment personalization.

In the stylized example above 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" [2, 1] 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.

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 in practice 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 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—which much of the data we have at our disposal in, for example, electronic patient records actually is—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) on 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 tumors are both less likely to receive chemotherapy and are 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 well understood machine learning task in which a computer learns a function between some observed input and some desired output [13]—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., 14, 15, 16, 17, 18].

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

3. Our current method: the RCT

Successful instances of personalized healthcare—such as the Zelboraf treatment for melanomas—do exist. Hence, 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 the RCT constitutes our highest level of evidence [19, 20]. 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 $\frac{1}{2}$ Mg, while another 100 patients would receive a high dose, say 1 Mg. Based on our example function given previously (see Figure 1), a naive RCT would conclude that the 1Mg dosage outperforms the $\frac{1}{2}$ Mg dose, despite the adverse effects for children. Thus, the canonical RCT addresses the three problems highlighted above as follows:

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 p-value—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.

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.

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 $\frac{1}{2}$ mg, and hence you get 1 mg. Next, the RCTs approach to the problem of learning causal relationships is extremely solid [21, 22]. 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.

The analysis above allows us to identify several drawbacks of the ICT. First of all, the singling out of very small subsets off all possible actions and context in the repeated execution of 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, than 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 chemotherapy 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. Alternative, sequential, methods for personalization

This section provides a sketch of a possible alternative method to the RCT; technical details for specific implementations can be found in the references. As an alternative to the RCT we could start by using 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 [13, 23, 24, 25], and hence there is 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 to obtain unbiased estimates of causal effects. Rather, as long as we can compute and store the probability of receiving a treatment conditional on the patient characteristics (the *propensity*), we can use the collected data to estimate causal effects [26, 27]. 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 our belief that the treatment has the highest reward [see, e.g., 28, 29]. 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 [14, 16, 18].

To be specific, we suggest to use a flexible machine learning model such as the Bayesian additive regression tree model, or shorthand BART [30, 23, 24], which can be denoted as follows:

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

Using the BART 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 31, 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 (deep) neural network models seem to be more popular [see 32, for an introduction], the Bayesian specification of the BART model conveniently allows for the direct quantification of the associated uncertainty [33]. Furthermore, this model effectively guards against

overfitting by virtue of a prior restricting trees of large depth.

The uncertainty quantification originating from the Bayesian specification of the BART model allows for the direct implementation of Thompson sampling: we choose our actions with a probability that is proportional to our posterior belief that the action. This probablity is given by:

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

Although the above integral might be hard to evaluate in practice, Thompson sampling is often easily implemented by simply taking a single MCMC draw from the posterior and selecting the action that maximizes the expected reward given that draw [see 18, for more details and alternative sampling methods]. As long as we apply this scheme, and store the probability of receiving an action at each point in time for each patient (i.e., the propensity-score), we can not only use BART and Thompson sampling to select actions, but also to create a dataset that we can re-use in offline evaluations of alternative decision policies [see 34, 35, for examples].

This alternative approach to treatment personalization can practically 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 might look a bit distant from reality. However the models we 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, in the near future, this suggestion is at least *technically* feasible.

4.1. Disadvantages of sequential methods for personalization

Contrary to the RCT, let us start by discussing the disadvantages of a sequential and adaptive 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?" These are genuine questions, but they are not disadvantages of the method: these questions equally need answers when designing an RCT. Actually, the proposed sequential 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, the proposed method makes answering these questions easier as opposed to harder.

However, there are more serious concerns: First of all, the 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 [26, 27, 36, 22]: 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.

4.2. Advantages of sequential methods for personalization

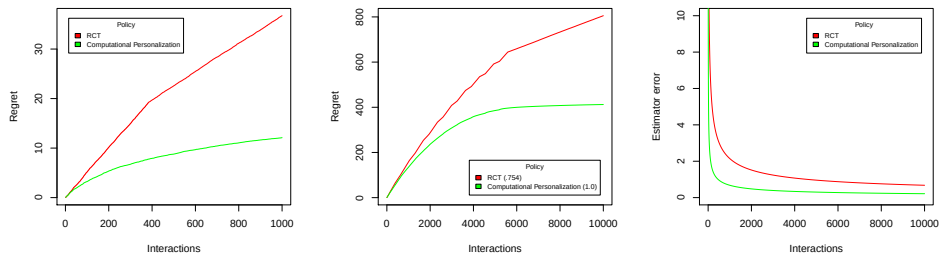
Why should we consider these novel, sequential, methods? They seem plagued with challenges. However, these novel methods should be adopted for one simple reason: with these methods we will have a better outcome. Now that's a bold statement, and one that is hard to quantify for healthcare in general. The number of future interactions, the number of possible actions, and the number of meaningful contextual factors is simply too large to say anything remotely precise for the healthcare system as a whole. However, at smaller scales, for simple versions of the personalization problem, we can actually quantify the benefits.

The performance of a personalization method can be evaluated. We often measure the quality of a *policy*—a decision-making strategy that dictates which treatment to select for whom given our historical data—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—which itself is, combined with the resulting guidelines, just another allocation policy—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—

coined "Computational Personalization" in the Figure—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 do when we personalize our treatments.

Scaling the problem to 10.000 decisions and 10 possible treatments (with success probabilities .5, .4 for the best two, and .3 for those remaining), the superior performance of sequential allocation using Thompson sampling is even more striking: the regret of the RCT is 800, while that of the sequential policy is only 400, as displayed in Figure 2b. Even more interestingly, the practice of sequential, binary RCTs identifies the best treatment in only $\frac{3}{4}$ 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. Again, 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. [35] 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



(a) (b) (c)

Figure 2: Three panels showing the superior performance of a sequential method for personalization and the RCT. The first two panels show the effectiveness and efficiency of sequential treatment selection over the RCT in terms of the regret of each policy: a computational approach that implements Thompson sampling outperforms the RCT when making a binary choice between treatments for a population of $N = 1000$ patients (panel A), and when (sequentially) making a choice between 10 competing treatments for a population of $N = 10000$ patients (panel B). Panel C demonstrates the standard error of an estimator as a function of the number of datapoints when comparing the two procedures Agarwal et al. [the latter result is adapted from 35].

using repeated RCTs.[1]

These simple computations show that the RCT is outperformed by sequential experimentation methods. Furthermore, it is reasonable to expect that the RCT will comparatively suffer more from making the problem more realistic, and therefore more complex. Thus, if anything, the presented differences in expected outcomes are underestimates of the actual outcomes rather than overestimates.

5. Conclusion

This discussion tried to analyze formally what we mean by *personalization* in healthcare and biomedical-informatics. By providing a formal definition of personalization we were able to critically evaluate the current approach to

[1]The [R] scripts to replicate the results presented here can be downloaded at `http://www.nth-iteration.com/downloads/`.

personalization—performing RCTs amongst smaller and smaller subgroups of patients—to alternative sequential methods. Admittedly, the alternative can be operationalized in a number of ways; effective implementations will depend on the context and on practical, technical and ethical constraints. However it remains that if we effectively and efficiently want to personalize treatments, we need to step away from the RCT and embrace adaptive sequential methods. Furthermore, we need to make sure that our data is stored such that future, offline, evaluations of alternative treatment allocation policies are possible.

The arguments provided in this discussion have consequences for our current practice. First of all, if we design systems that directly interact with patients [see, e.g., 5, 6, 7], we should consider effective methods to personalize the information presented in these systems. Three recommendations follow from our analysis:

1. We should move away from the traditional idea implicit in clinical trials that we can select, based on a limited time trial, the most effective intervention for all patients. Rather, we should embrace the fact that we live in a dynamical world in which interventions come and go, estimates of intervention effects always contain uncertainty, and the effects of interventions are heterogeneous. Hence, we need to embrace sequential and adaptive approaches. Admittedly, the analysis of sequential experimentation methods is more involved than that of our current, static, RCTs. However, many methods to do so exist, and the implications of sequential allocation on frequentist errors (i.e., Type I and Type II) errors are properly understood for many sequential designs [see 37, for a thorough introduction] and recently developed ability to fit complex supervised learning models on extremely large datasets enable us to implement the analysis suggested in this discussion. Furthermore, sequential designs are administratively and organizationally more challenging than traditional designs since treatment allocation

probabilities change over time: we would argue however that the technological means necessary to carry out sequential allocation at large scales, over multiple locations, are, by virtue of high-speed computer networks, currently already in place.

2. We should focus not on evaluating a specific intervention, but rather on evaluating treatment allocation policies that choose interventions for patients. The distinction between these two concepts is meaningless when selecting one-size-fits-all interventions: in these cases the policy assigns the same intervention to each unit and hence evaluating the effect of the intervention equates to evaluating the effect of the treatment allocation policy. However, as we move to sequential adaptive policies this is no longer the case: we should evaluate the outcomes of allocation policies, and their evolution as evidence accumulates, as opposed to the effects of interventions.
3. We should collect evidence on the progression of the effect of a set of policies as opposed to estimating the effect of a single treatment allocation policy at a single point in time. As sequential, adaptive, policies trade of exploration and exploitation their effectiveness should improve over time; it is this improvement, and the robustness to the introduction of new interventions, that should drive our decision to embrace a treatment allocation policy.

It might seem as though the methods proposed in this discussion inherently "need" much more data than the RCT; this is however not true. Even within current RCTs we can often benefit from adaptive, sequential allocation schemes [see, e.g., 38], and the data collected in our current RCTs can directly be used to estimate individual level effects of treatments using the proposed BART model [see 33]. I have tried to argue that additional benefit can also be gained from continuous data collection after the RCT. Interestingly, the proposed method could even effectively deal with much smaller datasets: while current $n = 1$ trials rely solely on within subject repetitions,

the non-parametric BART model could be used to generalize findings between $n = 1$ trials be "borrowing strength" from results attained for similar—but not the same—patients.

Next to designing patient facing systems ourselves, our field is also concerned with the development and evaluation of systems that store and distribute health data [see, e.g., 8, 9]. The analysis of personalization methods presented above gives rise to the following recommendations:

1. We should clearly separate the data we collect into context—the things we do not control—and actions—the things we do control. Furthermore, we should store each of these accompanied by their respective time-points and as much as possible we should associate outcomes directly to the actions—or set of actions—the most likely causes the outcome.
2. We should make an effort to store propensities—the probabilities of different actions given the current context—in our databases. For randomized trials this is trivial. For observational data this is more challenging. Clinical guidelines should, in theory, decide which action to choose for whom and lead to propensities of 0 or 1 which effectively renders the data useless to evaluate alternative allocation policies offline. However, we often find that guidelines are not followed closely [39], and variation in action selection does exist. We should capture this variation, and if its truly random we should store the respective probabilities. We could even consider explicitly introducing random variation when guidelines are indecisive.

Following the recommendations above would simplify the retrospective analysis of health data: by adhering to methodologically driven standards it is easy for an analysts to (re-)analyse existing health data. Furthermore, adhering to these methodologically driven recommendations would increase the value of our collected data: the inclusion of the causal mechanisms involved and the relevant propensity scores will allow for the effective offline evaluation

of treatment allocation policies. Effectively, incorporating methodological concerns for personalization into the design of biomedical informatics systems that store personal health data has the potential to greatly increase the utility of our health data.

References

[1] P. C. Ng, S. S. Murray, S. Levy, J. C. Venter, An agenda for personalized medicine, Nature 461 (2009) 724–726.

[2] M. A. Hamburg, F. S. Collins, The path to personalized medicine, N Engl J Med 2010 (2010) 301–304.

[3] P. Chapman, A. Hauschild, C. Robert, J. Larkin, J. Haanen, A. Ribas, D. Hogg, S. O'Day, P. Ascierto, A. Testori, et al., 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 (2011) LBA4–LBA4.

[4] N. Scholz, Personalised medicine; The right treatment for the right person at the right time, Technical Report, 2015. URL: `http://www.europarl.europa.eu/RegData/etudes/BRIE/2015/569009/EPRS_BRI(2015)569009_EN.pdf`.

[5] S. A. Collins, L. M. Currie, S. Bakken, D. K. Vawdrey, P. W. Stone, Health literacy screening instruments for ehealth applications: a systematic review, Journal of biomedical informatics 45 (2012) 598–607.

[6] S. Triberti, S. Barello, The quest for engaging ami: Patient engagement and experience design tools to promote effective assisted living, Journal of biomedical informatics 63 (2016) 150–156.

[7] H. op den Akker, M. Cabrita, R. op den Akker, V. M. Jones, H. J. Hermens, Tailored motivational message generation: A model and practical

framework for real-time physical activity coaching, Journal of biomedical informatics 55 (2015) 104–115.

[8] A. Roehrs, C. A. da Costa, R. da Rosa Righi, Omniphr: A distributed architecture model to integrate personal health records, Journal of biomedical informatics 71 (2017) 70–81.

[9] G. Namulanda, J. Qualters, A. Vaidyanathan, E. Roberts, M. Richardson, A. Fraser, K. H. McVeigh, S. Patterson, Electronic health record case studies to advance environmental public health tracking, Journal of biomedical informatics 79 (2018) 98–104.

[10] Z. Huang, Z. Ge, W. Dong, K. He, H. Duan, Probabilistic modeling personalized treatment pathways using electronic health records, Journal of biomedical informatics (2018).

[11] M. Kaptein, B. De Ruyter, P. Markopoulos, E. Aarts, Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking, ACM Transactions on Interactive Intelligent Systems (TiiS) 2 (2012) 10.

[12] M. Kaptein, P. Markopoulos, B. De Ruyter, E. Aarts, Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles, International Journal of Human-Computer Studies 77 (2015) 38–51.

[13] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, volume 11, Springer Science & Business Media, 2013. URL: `https://books.google.com/books?id=yPfZBwAAQBAJ{&}pgis=1`.

[14] P. Ortega, D. Braun, Generalized Thompson Sampling for Sequential Decision-Making and Causal Inference, arXiv preprint

arXiv:1303.4431 (2013). URL: http://arxiv.org/abs/1303.4431. arXiv:arXiv:1303.4431v1.

[15] S. Agrawal, Further Optimal Regret Bounds for Thompson Sampling, CoRR none (2012) 1–14. arXiv:arXiv:1209.3353v1.

[16] I. Osband, Bootstrapped Thompson Sampling and Deep Exploration (2015) 1–10. arXiv:arXiv:1507.00300v1.

[17] H. Bastani, M. Bayati, Online Decision-Making with High-Dimensional Covariates (2015) 1–47.

[18] D. Eckles, M. Kaptein, Thompson sampling with the online bootstrap, arXiv preprint arXiv:1410.4009 (2014).

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

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

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

[22] G. W. Imbens, D. B. Rubin, Causal inference in statistics, social, and biomedical sciences, Cambridge University Press, 2015.

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

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

[25] C. M. Bishop, Pattern recognition and machine learning, springer, 2006.

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

[27] M. J. Funk, D. Westreich, C. Wiesen, T. Stürmer, M. A. Brookhart, M. Davidian, Doubly robust estimation of causal effects, American journal of epidemiology 173 (2011) 761–767.

[28] W. R. Thompson, On the Likelihood that one Unknown Probability Exceeds Another in view of the Evidence of two Samples, Biometrika 25 (1933) 285–294. URL: `http://biomet.oxfordjournals.org/content/25/3-4/285.extract`. doi:`10.1093/biomet/25.3-4.285`.

[29] E. Kaufmann, N. Korda, R. Munos, Thompson Sampling: An Asymptotically Optimal Finite Time Analysis, Algorithmic Learning Theory (2012) 15. URL: `http://link.springer.com/chapter/10.1007/978-3-642-34106-9{_}18{%}5Cnhttp://arxiv.org/abs/1205.4217`. `arXiv:1205.4217`.

[30] M. T. Pratola, H. A. Chipman, J. R. Gattiker, D. M. Higdon, R. McCulloch, W. N. Rust, Parallel bayesian additive regression trees, Journal of Computational and Graphical Statistics 23 (2014) 830–852.

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

[32] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, volume 1, MIT press Cambridge, 2016.

[33] B. R. Logan, R. Sparapani, R. E. McCulloch, P. W. Laud, Decision making and uncertainty quantification for individualized treatments, arXiv preprint arXiv:1709.07498 (2017).

[34] L. Li, W. Chu, J. Langford, X. Wang, Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms, in: Proceedings of the fourth ACM international conference on Web search and data mining, ACM, 2011, pp. 297–306.

[35] A. Agarwal, S. Bird, M. Cozowicz, L. Hoang, J. Langford, S. Lee, J. Li, D. Melamed, G. Oshri, O. Ribas, S. Sen, A. Slivkins, A multiworld testing decision service, CoRR abs/1606.03966 (2016). URL: `http://arxiv.org/abs/1606.03966`.

[36] J. Pearl, Causality, Cambridge university press, 2009.

[37] D. A. Berry, B. Fristedt, Bandit problems: sequential allocation of experiments (monographs on statistics and applied probability), London: Chapman and Hall 5 (1985) 71–87.

[38] B. P. Carlin, S. M. Berry, J. J. Lee, P. Muller, Bayesian adaptive methods for clinical trials, CRC press, 2010.

[39] V. K. Ho, R. A. Damhuis, H. H. Hartgrink, Adherence to national guidelines for gastric cancer in the netherlands: A retrospective population-based audit, International journal of cancer 132 (2013) 1156–1161.