Transfer Learning from Well-Curated to Less-Resourced Populations with HIV

Sonali Parbhoo SPARBHOO@SEAS.HARVARD.EDU
John A. Paulson School of Engineering and Applied Sciences
Harvard University
Cambridge, MA, USA.

Mario Wieser MARIO.WIESER@UNIBAS.CH
Department of Mathematics and Informatics
University of Basel
Basel, Switzerland.

Volker Roth VOLKER.ROTH@UNIBAS.CH
Department of Mathematics and Informatics
University of Basel
Basel, Switzerland.

Finale Doshi-Velez FINALE@SEAS.HARVARD.EDU
John A. Paulson School of Engineering and Applied Sciences
Harvard University
Cambridge, MA, USA.

Abstract

In Europe and North America, more homogeneous virus types and the relatively high availability of sequencing technologies have helped transform HIV from a life-threatening disease to a manageable chronic condition. However, modern therapies have been less successful in managing HIV in Africa, where there is more viral heterogeneity and access to sequencing is much less available. In this work, we present a novel mixture based approach that uses a deep information bottleneck to transfer patterns learned from European HIV cohorts—where genomic data is readily available—to African patients where no such data is available. We demonstrate its utility for optimising treatments for the first time in a set of HIV patients in Africa, and note how this approach may be applicable to many other scenarios where a variable is measured in some population but is missing from the target population.

1. Introduction

Human Immunodeficiency Virus (HIV) affects more than 36 million people worldwide. Fortunately, administration of combinations of drugs known as Antiretroviral Therapy (ART), targeting different phases of the viral replication cycle, can transform this life-threatening virus into a manageable chronic condition. However, this requires choosing the right combinations at the right time. In Europe and North America, monitoring the viral variants via genotype resistance testing has enabled clinicians to track the rapidly mutating virus and

apply the appropriate ART (Bogojeska et al., 2012; Deeks et al., 2013) [1]. Unfortunately, these therapies have seen lower successes in Africa. For instance, in 2018 two-thirds of the world's 770,000 HIV-associated deaths came from Africa (World Health Organisation, 2018). While some of these deaths are a result of a lack of treatment, 81% of people living in Africa know their HIV status and 68% have access to ART therapy. Lack of regular access to sequencing technologies makes it difficult to monitor the evolution of the virus and thus appropriately target therapies; the virus also exhibits more heterogeneity in Africa.

In this work, we consider the following question: to what extent is it possible to transfer treatment insights from a relatively homogeneous, well-measured population (e.g. cohorts of patients in Europe or North America) to a more heterogeneous, less-measured target population (e.g. African patients)? Addressing this question requires both understanding *when* to transfer—if an African patient has a viral subtype not present in a European population, copying a strategy that worked in Europe may not help—as well as *how* to transfer in the face of very few measurements in the target population.

We answer this question using the deep Information Bottleneck principle (IB) (Tishby and Zaslavsky, 2015; Parbhoo et al., 2020) to perform a sufficient reduction of the covariates for inferring treatment outcomes. The IB enables us to build a discrete reference class over patients with "complete data" during training, to which we can map patients with missing genetic information at test time, and subsequently estimate treatment effects on the basis of these groups. We next apply a mixture-of-experts approach to learn when the IB transfer will be useful, and when it is better to rely on similar patients in the target population.

Many works have studied the problem of HIV therapy selection in the context of machine learning (e.g. Bickel et al. (2008); Bogojeska (2011); Parbhoo et al. (2017)). However, to date these have all been applied to European and North American cohorts, and rely heavily on the availability of genotype data both at training time and at test time. In this paper, we present for the first time an approach for reliably reasoning about the effects of a series of interventions for HIV therapy selection in populations where genetic sequence data is unavailable at test time—specifically, on an African population.

Generalisable Insights about Machine Learning in the Context of Healthcare

While the approach we present here is particularly important for HIV therapy selection in the absence of genotype information at test time, there are several takeaways that may be generalisable to a broader healthcare context. Specifically, our method may be applied to several other contexts where there is a systematic missingness in data at test time, as a result of the costs of performing a particular test or other restrictions. For instance, different hospitals may have different protocols for collecting the measurements of patients with a particular disease which may result in a systematic missingness of some variables for a group of individuals associated with a particular hospital; alternatively, patients with health insurance may be more likely to have certain tests performed in hospital for diagnosing and treating their condition in comparison to patients without health insurance. In these instances, the IB approach may be useful to compress the relevant statistics from a group of

1. Genetic sequencing has enabled at least 55 recombinant forms of HIV to be established for these populations to date (Alamos, 2020)

individuals with complete measurements to make inferences about those patients without these. We also see our approach as being particularly useful for evaluating the performance of existing HIV therapy guidelines or as basis for ranking hypothetical clinical policies in different scenarios. Such insights could lead to improvements in existing therapy guidelines and offer better opportunities for intervention.

2. Related Work

Disease progression models are important for understanding the critical steps during the development of diseases. In the medical and machine learning literature, several approaches have been developed for predictive modelling of progression of HIV and other diseases. Some of these methods (e.g. Bogojeska et al. (2012); Bickel et al. (2008); Saigo et al. (2011)) try to explain changes in a patient's outcomes in terms of the outcomes of patients with similar treatment histories. Overall, these approaches work well when patients exhibit significant overlap in their treatment histories. In contrast to non-parametric approaches, several model-based approaches have also been developed to reason about treatment effects. The vast majority of these methods try to explain changes in a patient's outcomes by learning a representation of some baseline covariates, based on which dynamic predictions may be generated, e.g HMMs, state space models Ernst et al. (2006), autoregressive models and Gaussian Processes (Schulam and Saria, 2017). In general, these approaches can capture model uncertainty more effectively and can perform well on outlier cases where there are few neighbouring patient histories. Previously, approaches have been proposed for HIV therapy selection that combine parametric and non-parametric models using variations of mixture models e.g. Parbhoo et al. (2017, 2018). Both of these methods account have been shown to be performant on multiple cohorts across Europe and North America, however have only been applied to cases where complete clinical and genetic information is available.

More closely related to the work we present here, are several methods that have been proposed for *transfer learning*. Among these, Künzel et al. (2019) develop a framework for estimating heterogeneous treatment effects by jointly training neural networks with shared weights such that important features may be re-used where necessary. Based on the treatment effects, one may subsequently infer the optimal treatment policy. While Künzel et al. (2019) use shared weights to transfer knowledge and reason about treatment effects, methods such as Hidden Parameter Markov Decision Processes (HiP-MDPs) first learn a policy directly, and subsequently adapt the policy in a robust way such that it may be applied to a related task (Doshi-Velez and Konidaris, 2016; Killian et al., 2017). Unfortunately however, none of these methods are applicable where a subset of patient covariates is missing at test time. Moreover, these methods also do not explicitly optimise for retaining the information that is relevant for predicting treatment outcomes as we require in this context.

Recently, Parbhoo et al. (2020) proposed a Cause-Effect Information Bottleneck (CEIB) to reason about treatment effects in high-dimensional settings where a subset of covariates is unavailable at test time. Unlike previous approaches, this method explicitly considers the relevance of information for predicting treatment outcomes. Here, the authors identify a set of meaningful information using the Information Criterion (Tishby et al., 2000; Tishby and Zaslavsky, 2015; Parbhoo et al., 2020) and transfer only the relevant information to

cases where complete measurements are unavailable during testing such that the treatment outcomes can still be inferred. In this paper, we adapt this framework to a multi-treatment setting, and incorporate this knowledge into a mixture-of-experts model for reasoning about treatment effects over heterogenous patient groups with incomplete measurements.

3. Information Transfer with Information Bottleneck

Our goal is to use a well-curated set of patients (e.g. European and North American cohort) to inform treatments for a target patient population that (a) may have different disease subtypes than the source population, and (b) is missing many of the measurements available in the source population (e.g genotype information). For this, we propose a mixture-of-experts model. Mixture models are a popular tool for representing heterogeneity among populations to reason about treatment effects in precision medicine. Not only are they flexible and scalable, but are especially useful for discovering latent subclasses of individuals with particular response patterns. In doing so, policy makers or individuals seeking to improve care could design interventions tailored explicitly to these groups.

Our core idea is the following: when trying to predict how a patient will respond to a treatment, using a mixture-of-experts network, we first determine whether to attempt to transfer the relevant information from the European population or to use the less-well-curated African cohort. If our mixture-of-experts model learns that transfer is not possible, then we follow a local strategy of, for instance, Bogojeska et al. (2012, 2010); Saigo et al. (2011); Bickel et al. (2008) to attempt a prediction. If we choose to transfer, we apply a variant of the Information Bottleneck (Tishby et al., 2000) that allows us to map both cohorts into a low-dimensional space. Our overall strategy is summarised in Figure 1. In the following, we first describe each of these prediction mechanisms and then the mixture-of-experts that chooses between them.

3.1. Transfer in the Presence of Missing Measurements and Multiple Treatments

Our transfer expert takes the form of a Cause-Effect Deep Information Bottleneck (CEIB) (Parbhoo et al., 2020) to reason about the effects of treatments where data are missing at test time. The original CEIB learns a sufficient covariate representation to approximate the Average Treatment Effect (ATE) (Dawid, 2007; Guo and Dawid, 2010) during testing. However, in general the ATE cannot be computed for multiple treatment groups as in the HIV therapy selection task. Here, we adapt CEIB to handle *missing data at test time in a multi-treatment setting.*

Problem Setup Suppose $X = (X_1, X_2) := \{h_n\}$ denotes a set of patient histories of covariates and prior treatments for $n = 1 \dots N$. Here, each history h_n consists of a complete list of clinical and genetic features for a patient at a certain time. During training we assume all the covariates $X \in \mathbb{R}^d$ can be observed. These correspond to, for instance, the complete measurements of a set of patients participating in a medical study, where dimension d is

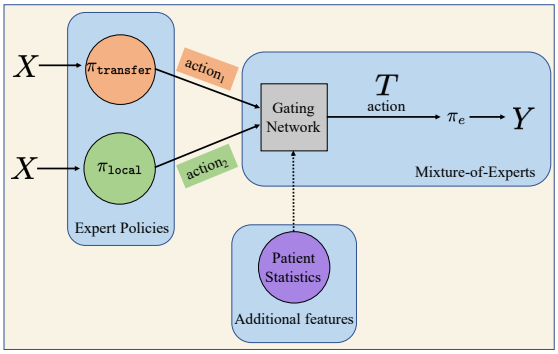

Figure 1: Overview of our approach. Given patient data X from two cohorts, the transfer policy $\pi_{\texttt{transfer}}$ suggests a treatment on the basis of a compressed representation of information from the European patients. The local policy $\pi_{\texttt{local}}$ finds the most similar patients from the African cohort and suggests a treatment on the basis of these. The mixture-of-experts determines when to follow each expert, thereby producing a mixed policy π_e of treatments T that we can evaluate to estimate patient outcomes Y.

large [2]. During testing however, we assume a portion of covariates X_1 are not observed, as a result of for instance, resource limitations or the costs associated with genotyping. Let $Y \in \mathbb{R}$ denote the outcomes for a patient following treatment T. Our goal is to learn a low-dimensional compressed representation Z of the relevant information in X to estimate long-term outcomes Y for a given treatment T. Based on this, we infer a suitable treatment policy for patients with incomplete covariate data at test time. The causal graph we consider for our first expert is shown in Figure 2 .

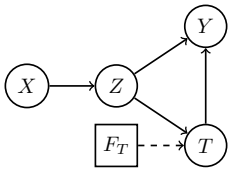

Figure 2: Causal graph corresponding to the proposed problem. F_T denotes the interventional distribution on T. The transfer expert compresses X into Z and conditions on this with T to reason about Y.

Assumptions Throughout this section, we make the simplifying assumption that all confounders are measured. This is also referred to as strong ignorability. This is a fairly strong assumption which, though untestable, is common in existing literature; in high-dimensional

2. Hereafter, we refer to this as well-curated. By well-curated we mean we have complete information with no imputation.

tasks though, it may be natural to assume one can measure everything that is relevant for estimating treatment effects for a subset of the patients, and attempt to transfer this distribution of information to a potentially larger set of test patients, as opposed to making more stringent assumptions e.g about the availability of proxy confounders. We also assume that a treatment T has not previously been applied to X. This assumption is necessary to ensure that compression Z does not capture post-treatment variables, which may otherwise bias predictions. In the context of HIV, for a patient this means that a particular therapy combination has not been encountered before. That is, while patients can occasionally switch back to old combinations, we make the simplifying assumption that this does not occur. Finally, we assume that our compressed representation Z serves as a sufficient statistic for the history for a patient, based on which we can evaluate their treatment policy.

Method We consider an extended formulation of the Information Bottleneck criterion as in Parbhoo et al. (2020) that enables us to learn a sufficient statistic of relevant information Z. Based on this, we can infer treatment effects Y given T for cases where some covariates are missing at test time. The adapted IB criterion is given by

$$\max_{\phi,\theta,\psi,\eta} -I_\phi(V_1;X_1) - I_\eta(V_2;X_2) + \lambda I_{\phi,\theta,\psi,\eta}(Z;(Y,T)), \tag{1}$$

where V_1 and V_2 are compressed discrete representations of the covariates, $Z = (V_1, V_2)$ is a concatenation of V_1 and V_2 and I represents the mutual information parameterised by ϕ, ψ, θ, and η respectively. We assume a parametric form of the conditionals $q_\phi(v_1|x)$, $q_\eta(v_2|x)$, $p_\theta(y|t,z)$, $p_\psi(t|z)$. Here $q_\phi(v_1|x)$ and $q_\eta(v_2|x)$ serve as variational approximations of $p(v_1)$ and $p(v_2)$ respectively. Then we have, $I_\phi(V_1;X_1) = \mathbb{E}_{p(x_1)} D_{KL}(q_\phi(v_1|x_1)||p(v_1))$ and $I_\eta(V_2;X_2) = \mathbb{E}_{p(x_2)} D_{KL}(q_\eta(v_2|x_2)||p(v_2))$. The first two terms of the criterion are optimised using two encoder networks that compress X into discrete latent representations V_1 and V_2 using the Gumbel-Softmax reparameterisation trick (Jang et al., 2016) to draw samples Z from a categorical distribution with certain probability. The last term of the criterion is optimised by a decoder architecture. Here, we introduce a different decoder model

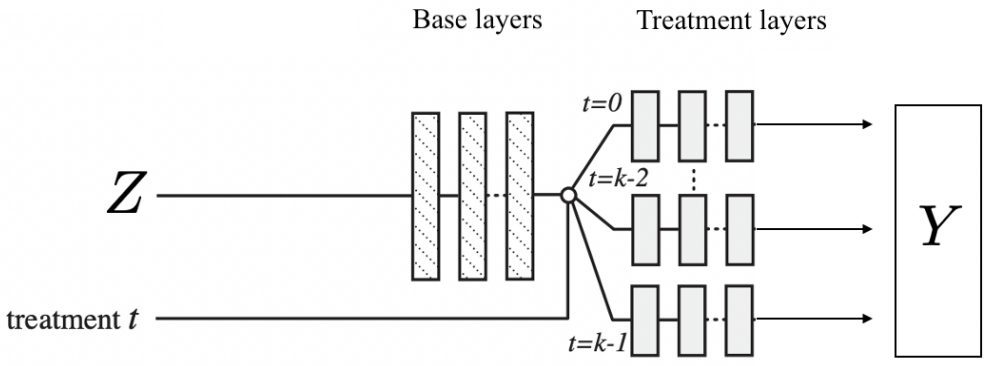

Figure 3: Proposed decoder architecture for our first expert.

for reasoning about treatment effects specifically in a multi-treatment setting. Specifically, we modify the TARNet architecture (Johansson et al., 2016) to condition on reduced covariate Z instead of high-dimensional X. For our decoder architecture, we use k networks corresponding to each treatment over a set of L_1 shared base layers as shown in Figure 3. Based on sufficient covariate Z, we can then formulate the conditionals as,

$$p_\psi(t|z) = \text{Cat}(f_1(z))$$
$$p_\theta(y|t,z) = \mathcal{N}(\mu = \hat{\mu}, \varsigma^2 = \hat{s}), \tag{2}$$

where Cat refers to the categorical distribution of treatments. As a result, we can account for multiple therapy combinations or treatment options for estimating treatment effects for HIV.

Overall, our transfer expert enables learning equivalence classes among patients with similar sufficient statistics Z. During testing where a subset of features is unavailable, we can thus cluster those patients with incomplete measurements based on their discretised representation Z. Adjusting λ in Eqn. 1 controls the degree of compression and allows us to interpret our resulting predictions in terms of learnt representation Z. Once we learn the latent representation, we build a set of statistics based on the discretisation for each patient, $\{(z_i, t_i)\}$, where t_i is given by the distribution of treatments for the particular group of patients with reduced z. These define the transfer policy $\pi_{\texttt{transfer}}$. Based on these statistics, we can evaluate the transfer policy using off-policy evaluation (see Section 4).

3.2. Predictions without Transfer: Kernel-based Learning

Our approach above enables transfer even when the target population has less measurements than the source population, as long as there is some way to infer enough of the bottleneck Z in the target population. However, we may not always be able to transfer: the patient may be missing too many measurements, or the patient's disease may be so different from the source population that no mapping exists. In this case, we need a fallback for making predictions. Earlier work such as Bogojeska et al. (2012) has found that kernel-based predictors can often do well in the setting of HIV drug response prediction.[3] The predictor has the form:

$$\hat{Y_n} = \sum_{h'_n} k(h_n, h'_n) Y_n, \qquad \forall h'_n \in X. \tag{3}$$

Here, $k(h_n, h'_n) \geq 0$ is a kernel function satisfying $\sum_{h'_n} k(h_n, h'_n) = 1$, $\forall h_n \in X$. During training where complete measurements are available, we compare the histories of patients within the European cohort to compute the similarity score. However at test time, where such measurements are unavailable, the best we can do is compare the histories of patients *within the African cohort* (test population). Based on this, we collect a set of statistics corresponding to the quantile distances between each of the histories h_n and h'_n, their corresponding lengths, as well as t. These define the kernel policy $\pi_{\texttt{local}}$. We subsequently perform off-policy evaluation using these statistics. For our specific application to HIV therapy selection, we use a therapy history alignment kernel analogous to Bogojeska et al. (2012).

3. We also considered other predictors, such as a parametric POMDP model as in (Parbhoo et al., 2017); see supplement for details.

3.3. When to Transfer: Gating Network between Experts

Our mixture-of-experts uses the properties of a patient's health state to choose whether or not performing a transfer of knowledge is helpful or not to reason about treatment effects. Based on this, it determines when to switch between the kernel and IB-based policies. Specifically, our network uses the following set of input features: reduced covariate representation z, trajectory length, quantile distances between the treatment histories, $CD4^+$ counts and viral load measurements as well as information about the viral strain. These quantities are available in the African cohort.

The mixture-of-experts gating function combines these features x linearly:

$$p_{\texttt{local}} = \text{sigmoid}(w.x) + b \tag{4}$$

$$p_{\texttt{transfer}} = 1 - p_{\texttt{local}}, \tag{5}$$

where $p_{\texttt{local}}$ and $p_{\texttt{transfer}}$ denote the probabilities assigned to using the local cohort (with the kernel-based expert) or attempting transfer (with our novel IB-based framework to account for lack of sequencing data in the African cohort). We optimise the gating network parameters to maximise patient outcomes, with our policy defined as $\pi_e = p_{\texttt{local}}\pi_{\texttt{local}} + p_{\texttt{transfer}}\pi_{\texttt{transfer}}$, and the quality of a policy π_e being determined by off-policy estimators in the next section. The optimisation is performed with gradient descent using the Adam optimiser (Kingma and Ba, 2014).

4. Evaluation Metrics for Measuring Policy Quality

We apply a collection of off-policy evaluation (OPE) estimators in order to evaluate our policies. *Importance Sampling (IS):* The classic *IS estimator* (Kahn and Marshall, 1953; Rubinstein, 1981; Koller and Friedman, 2009) is given as $V^{\pi_e} = \sum_{n=1}^{N} \rho_n Y_n$, where ρ denotes the importance weight computed as $\rho_n = \frac{\pi_e(t_n|h_n)}{\pi_b(t_n|h_n)}$. Histories that are unlikely are thus given a smaller weight when evaluating a policy using IS.

Weighted Importance Sampling (WIS): The standard IS estimator is prone to high variance so a weighted variant exists. Here, the value is given as $V^{\pi_e} = \frac{\sum_{n=1}^{N} \rho_n Y_n}{\sum_{n=1}^{N} \rho_n}$.

Doubly Robust (DR): While the WIS estimate has a lower variance, it is biased. The doubly robust off-policy evaluation scheme (DR) (Jiang and Li, 2015) attempts to trade off bias and variance by coupling the IS weights with a regression estimate $\hat{Q}$ of the function V^{π_e} that is computed on a separate data set. The estimated value of treatment policy π_e is given as, $V^{\pi_e} = V^{\pi_e} + \sum_{n=1}^{N} \rho_n (Y_n - \hat{Q})$. DR works well if either the regression estimate $\hat{Q}$ or IS-weights is fairly accurate.

5. Experiments

Baselines We compare the performance of our proposed approach against several baselines: Two of the baselines did not involve any transfer namely (i) a kernel policy and (ii) an IB policy, both trained and evaluated on only the African cohort. Since (i) and (ii) do not

make use of the transfer mechanism, we refer to these as local experts. We further compare our approach against (iii) a kernel expert from Section 3.2, (iv) an IB transfer expert from Section 3.1 and (v) the mixture-of-experts approach from Parbhoo et al. (2017), all of which are trained on the EuResist cohort and evaluated on the African cohort. We view these as experts that use some notion of knowledge transfer to reason about treatment effects. Unlike the baselines, our proposed approach is a mixture of a local expert and a transfer expert.

Training Details The mixture-of-experts approach we consider for baseline (v) relies on a POMDP model as one of the experts. Here, we consider a POMDP model of 20 patient states and Gaussian emission probabilities, where the number of states is chosen using the Bayesian Information Criterion and the discount factor is fixed at $\gamma = 0.98$. The IB expert for (iv) is trained using six 5-dimensional Gaussian mixture components where compression parameter $\lambda = 5.8$, based on which we may analyse the cluster compositions. These parameters are selected by examining the mutual information curve of $I(Z;(T,Y))$ against $I(Z;X)$ and selecting the value where the information curve saturates as the compression changes, or where a sufficient covariate representation is available. For the kernel approach we use the history alignment kernel from Bogojeska et al. (2012). The details about this kernel can be found in the supplement.

5.1. Results and Analysis

Locally-trained baselines perform worse than current practice; transfer-based baselines barely reach the level of current practice. Table 6 shows the OPE results of our approach in comparison to the baselines we considered. We include the evaluation over three different OPE methods that have different bias-variance tradeoffs. Additional results based on training and testing the approach on the EuResist cohort and African cohorts individually can be found in the supplement. For comparison purposes, we also include the value of the behaviour policy as the first row of the Table 6. Based on these results, we notice a few things. First, methods that do not use any transfer mechanism tend to perform poorly and have higher variance. This is unsurprising considering the lack of NGS data available for these patients, as well as the size of the African cohort which is significantly smaller than the EuResist data.

Next, each of the methods involving some form of knowledge transfer perform this transfer in different ways: the kernel maps patient histories from the African cohort to histories in the European cohort and uses this as a statistic to reason about treatment effects; in contrast, the mixture-of-experts learns a partition of the space of patients, and conditions on this representation to infer treatment outcomes; the IB performs a sufficient reduction of the covariate in order to reason about treatment effects. As a result, each of these approaches may capture different kinds of information about how to perform the knowledge transfer which ultimately produces the performance differences observed.

Mixing local and transfer-based policies achieves high rewards. The combination of a local kernel expert and a transfer IB expert outperforms all the baselines considered, as it is able to capture both (i) *how to perform a transfer of relevant information* via learning

Type	Method	DR	IS	WIS
Behaviour Policy		5.02 $\pm$ 1.18		
Local	Kernel	3.56 $\pm$ 1.42	1.27 $\pm$ 1.14	1.80 $\pm$ 1.07
	CEIB	3.29 $\pm$ 1.13	3.80 $\pm$ 2.41	3.76 $\pm$ 2.19
Transfer	Kernel	4.17 $\pm$1.4	4.18 $\pm$1.20	4.16 $\pm$1.71
	CEIB	6.29 $\pm$ 0.14	5.17 $\pm$ 0.38	5.27 $\pm$ 0.29
	Mixture-of-Experts	5.28 $\pm$ 0.37	3.42 $\pm$1.39	4.81 $\pm$ 1.25
Local + Transfer	**Ours**	**8.96 $\pm$ 0.39**	**10.64 $\pm$ 1.2**	**10.62 $\pm$1.67**

Table 1: Off-policy evaluation using importance sampling, weighted importance sampling and doubly robust methods for different therapy selection models across the African cohort. Overall, combining a local expert with an expert that performs a transfer of the distribution of relevant covariate information for cases where data are missing, produces the best outcomes.

a sufficiently reduced covariate representation and (ii) *when such a transfer is necessary in the face of few measurements* via the gating network. We discuss both (i) and (ii) in more detail below.

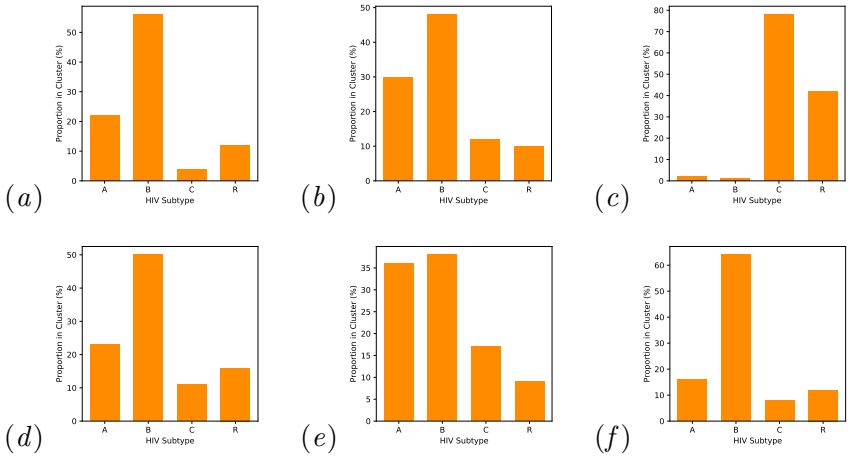

Figure 4: Illustration of the proportion of HIV patients in each cluster according to their subtype. Only the most frequently occurring subtypes in each cluster are shown. A, B, C denote the respective subtype, while R denotes a recombinant subtype. The clusters differ in their composition by subtype. In particular, cluster (c) shows an enrichment of information in the Subtype C group as well as those patients with recombinant forms from the African cohort. These patients are likely to vary from those patients in the EuResist cohort and form their own cluster.

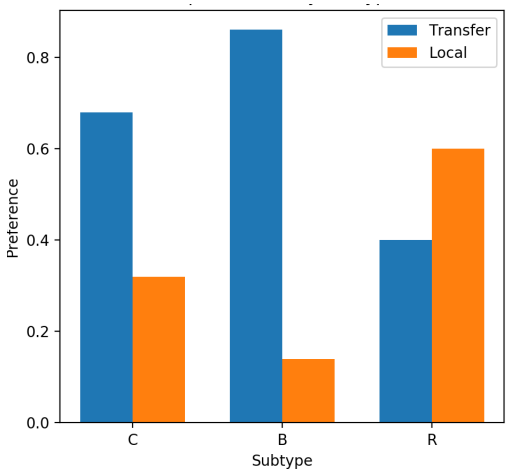

Figure 5: Expert chosen for recombinant patients and those with Subtypes B and C in African cohort. Information transfer is possible for 68% of patients with subtype C and 40% of recombinant subtypes. Information transfer is also possible for a large proportion of subtype B patients that have significant overlap with patients in the EuResist cohort. For the remaining patients, the local expert is preferred.

Interpreting the Bottleneck: IB determines that HIV subtype is an important bottleneck parameter for transfer. Thus far, we showed that combining local and transfer experts using a mixture-of-experts outperforms existing approaches for HIV therapy selection. Here, we validate when the gating network chooses to switch between each expert, as well as what information can be transferred from European populations when reasoning about treatment effects for patients in Africa. We begin by performing an analysis of the compressed representation learnt by the transfer IB expert. We examine the predominant features of the clusters obtained. These results are shown in Figure 4 according to the proportion of subtypes within in each cluster. Here, cluster (c) shows an enrichment of patients with Subtype C in comparison to other subtypes. This subtype is more prevalent in African population, while Subtypes B and A dominate the other populations. The IB approach identifies this as relevant information as a basis to group patients and reason about treatment effects. We also observe a fairly high number of recombinant strains in cluster (c). We attribute this to the fact that many of the patients in the African cohort experience therapy failures from drug resistance and have to undergo several therapy switches to manage viral mutation. Clusters (a) and (f) are very similar in terms of their distribution according to their subtype, however tend to differ along several other dimensions such as age distribution and risk group.

Interpreting the Gating: Network learns to transfer when there exist similar patients in European cohort. Here, we validate when knowledge transfer using CEIB is necessary, by examining the choices made by our proposed mixture-of-experts model and the features that contribute to the decisions made by the gating network. In particular,

we find the mixture-of-experts has a preference for the kernel policy 37.2% of the time, while choosing the IB expert 62.8% of the time. These proportions seem consistent with the relative proportions of subtypes in the South African cohort. Specifically, the South African cohort contains approximately 36% patients with Subtype C, 31% with Subtype B and a remaining 23% of patients with recombinant forms. Table 2 shows a ranked list of the most important features identified by our mixture-of-experts for predicting patient outcomes. Overall, the model relies heavily on the subtype information, a few resistance mutations and the viral load to reason about treatment effects.

Next, we analyse when the transfer expert is used for different subtypes. These results are shown in Figure 5. We observe that the model makes use of the transfer expert for 68% of patients with Subtype C, 86% of patients with Subtype B, and 40% of patients with recombinant forms. Of the patients with recombinant forms, the vast majority (72%) of these patients have a recombinant strain involving Subtype C. Evidently for patients of Subtype B within the African cohort, it is relatively straightforward to map to similar patients from the European cohort and transfer the relevant information via the information bottleneck. Interestingly however, it is also possible to perform such an information transfer for a significant proportion of patients with Subtype C and recombinants involving Subtype C. For the remaining patients, the kernel expert is the preferred approach. These are patients with complex recombinants of multiple strains for whom distribution transfer from the European cohort does not make sense. These patients also tend to experience significantly more therapy failures and have higher viral loads. For these individuals, the only alternative is to map to other patients within the African population that exhibit similar response patterns.

Feature	W_k
Quantile distance	**0.4720**
$CD4^+$ count	0.1612
Viral load	**-0.4579**
Viral strain	**2.6179**
History length	0.1724
RT67N	-0.1161
RT 215YF	-0.128

Table 2: Feature weights for gating function. RT67N and RT215YF are HIV mutations. The patient subtype, quantile distance to similar patients, and viral load have the largest weights.

Interpreting the Policy: Network learns policies that are consistent with current therapy guidelines We compared the treatment policy produced by our approach to existing WHO and IAS-USA clinical guidelines for treating HIV (Saag et al., 2018) to assess how consistent these were. For the purposes of this analysis, we restricted our focus to the adult patients (aged over 18) in the African cohort. We further classify whether the learnt treatment policy follows a) the recommended regimen, b) an alternative regimen that is not recommended but not in violation of the guideline or c) a regimen in violation

of the guideline. Overall, we find that our policy observes the guidelines 73% of the time, while suggesting an alternative treatment 15% the time. For the cases where an alternative treatment is learnt, these treatments tend to comply with previous guidelines (Günthard et al., 2016; Günthard et al., 2014), which have since been modified as a result of reformulating drug compounds into simpler regimens for improved adherence. The remaining 12% were patients with more complex recombinants, where our learnt treatments contained more drug compounds than outlined by the guidelines to combat resistance. For these cases in the African cohort, genetic sequencing would be crucial to reason about treatment effects.

6. Discussion

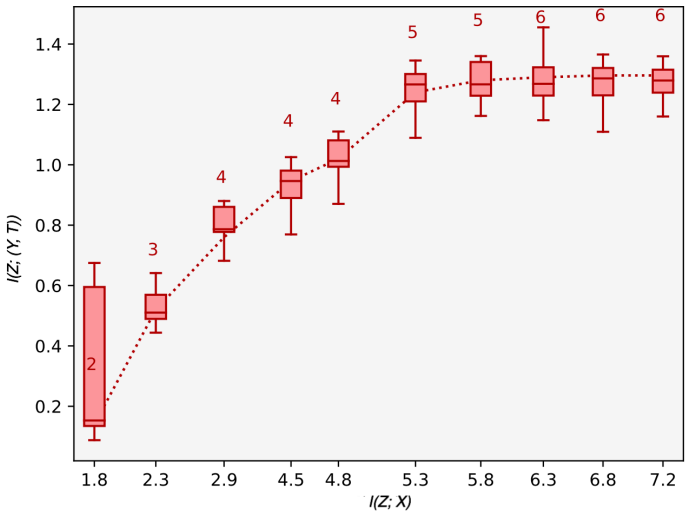

Figure 6: Information curve of $I(Z;(T,Y))$ against $I(Z;X)$. The numbers on top of the box plots indicate the dimensions of the compressed representation. We perform a sufficient reduction of the covariate where the information curve stabilises. Based on the reduced covariate, we can predict outcomes Y and perform OPE on the learnt policy.

Tuning λ helps us understand whether information transfer is meaningful. The selection of compression parameter λ is key to the performance of the IB approach. Adjusting λ allows us to explore a range of covariate representations by examining the mutual information curves of $I(Z;(T,Y))$ against $I(Z;X)$ such as in Figure 6. This is key to the model's interpretability as it enables us to examine a range of solutions from having a completely insufficient representation at the start of the curve, to a completely sufficient representation where the curve stabilises. Adjusting λ can also give us insights about when information transfer is *not* helpful. Such a case would arise if our dataset contained co-

variate information that could not be further compressed when reasoning about treatment outcomes. In this instance, the information curve would appear relatively constant, while the corresponding distributions of clusters would be fairly similar for varying values of λ. In such a situation, no knowledge transfer should be necessary, and one should rather apply a local approach to reason about outcomes.

Our approach may offer broader insights about when medical guidelines are helpful. While our approach was used for treating patients with HIV, we believe it could offer insights about applying existing therapy guidelines for various diseases to different populations. Specifically, by fixing treatments according to various medical guidelines, one could assess how outcomes change across different populations as well as identify the information relevant to predicting these outcomes. Such analyses could reveal nuances about when guidelines should be followed, which guidelines work for different types of patients, as well as when they may be unhelpful. Together, these insights could lead to developing new guidelines and better practices for treating populations with less resources.

7. Conclusion

While sequencing technologies have helped to advance HIV therapy in parts of Europe and North America, these therapies have been less successful for managing HIV in Africa as a result of the high viral heterogeneity and the costs associated with genetic sequencing. In this paper, we introduced a novel mixture-of-experts approach to tackle this challenge; first, we trained a transfer expert for identifying information that may be relevant to transfer across different populations when reasoning about treatment effects; for cases where this transfer is not possible, we trained a local expert for identifying within-population effects that are useful to infer patient outcomes. In general, our approach is applicable to various domains that suffer from similar problems with missing measurements, as well as scenarios where we wish to evaluate existing treatment guidelines across different populations.

Limitations and Future Work An important assumption we made across all our models was the fact that confounding factors may be measured. In reality, there is no way to check the validity of this assumption as there may be many factors that directly or indirectly influence a patient's outcomes. Hence future work could explore ways to relax this assumption through explicit modelling of latent confounders. We also performed various types of off-policy evaluation to evaluate the performance of the policies we learn. These methods have different trade-offs and can have high-variance when there is little overlap between the behaviour policy in the data and the evaluation policy. To address this issue, future work could consider different strategies for evaluation that do not make use of importance weighting. In our work, we combined transfer and local experts here using a simple linear combination gating network. Future work could use more sophisticated strategies for combining these as well as how to incorporate some notion of uncertainty between patient groups when performing information transfer.

References

Los Alamos. Los alamos HIV Sequence Database: HIV Circulating Recombinant Forms, March 2020. URL https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html.

Steffen Bickel, Jasmina Bogojeska, Thomas Lengauer, and Tobias Scheffer. Multi-task learning for hiv therapy screening. In *Proceedings of the 25th international conference on Machine learning*, pages 56–63. ACM, 2008.

Jasmina Bogojeska. History distribution matching method for predicting effectiveness of hiv combination therapies. In *Advances in neural information processing systems*, pages 424–432, 2011.

Jasmina Bogojeska, Steffen Bickel, André Altmann, and Thomas Lengauer. Dealing with sparse data in predicting outcomes of hiv combination therapies. *Bioinformatics*, 26(17): 2085–2092, 2010.

Jasmina Bogojeska, Daniel Stöckel, Maurizio Zazzi, Rolf Kaiser, Francesca Incardona, Michal Rosen-Zvi, and Thomas Lengauer. History-alignment models for bias-aware prediction of virological response to hiv combination therapy. In *AISTATS*, pages 118–126, 2012.

A Philip Dawid. Fundamentals of statistical causality. Technical report, Department of Statistical Science, University College London, 2007.

Steven G Deeks, Sharon R Lewin, and Diane V Havlir. The end of aids: Hiv infection as a chronic disease. *The Lancet*, 382(9903):1525–1533, 2013.

Finale Doshi-Velez and George Konidaris. Hidden parameter markov decision processes: A semiparametric regression approach for discovering latent task parametrizations. In *IJCAI: proceedings of the conference*, volume 2016, page 1432. NIH Public Access, 2016.

Damien Ernst, Guy-Bart Stan, Jorge Goncalves, and Louis Wehenkel. Clinical data based optimal sti strategies for hiv: a reinforcement learning approach. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 667–672. IEEE, 2006.

Huldrych F. Günthard, Judith A. Aberg, Joseph J. Eron, et al. Antiretroviral treatment of adult hiv infection: 2014 recommendations of the international antiviral society usa-panel. *JAMA*, 312(4):410–425, 2014. doi: 10.1001/jama.2014.8722. URL +http://dx.doi.org/10.1001/jama.2014.8722.

Huldrych F Günthard, Michael S Saag, Benson Constance A, et al. Antiretroviral drugs for treatment and prevention of hiv infection in adults: 2016 recommendations of the international antiviral societyâ“usa panel. *JAMA*, 316(2):191–210, 2016. doi: 10.1001/jama.2016.8900. URL +http://dx.doi.org/10.1001/jama.2016.8900.

Hui Guo and A Philip Dawid. Sufficient covariates and linear propensity analysis. In Yee Whye Teh and Mike Titterington, editors, *Proceedings of the Thirteenth International*

Conference on Artificial Intelligence and Statistics, volume 9 of *Proceedings of Machine Learning Research*, pages 281–288, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*, 2016.

Nan Jiang and Lihong Li. Doubly robust off-policy evaluation for reinforcement learning. *arXiv preprint arXiv:1511.03722*, 2015.

Fredrik D. Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, ICML'16, pages 3020–3029. JMLR.org, 2016.

Herman Kahn and Andy W Marshall. Methods of reducing sample size in monte carlo computations. *Journal of the Operations Research Society of America*, 1(5):263–278, 1953.

Taylor W Killian, Samuel Daulton, George Konidaris, and Finale Doshi-Velez. Robust and efficient transfer learning with hidden parameter markov decision processes. In *Advances in neural information processing systems*, pages 6250–6261, 2017.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. abs/1412.6980, 2014.

Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences*, 116(10):4156–4165, 2019.

Sonali Parbhoo, Jasmina Bogojeska, Maurizio Zazzi, Volker Roth, and Finale Doshi-Velez. Combining kernel and model-based learning for HIV therapy selection. *In Proceedings of the AMIA Summit on Clinical Research Informatics (CRI)*, 2017.

Sonali Parbhoo, Omer Gottesman, Andrew Slavin Ross, Matthieu Komorowski, Aldo Faisal, Isabella Bon, Volker Roth, and Finale Doshi-Velez. Improving counterfactual reasoning with kernelised dynamic mixing models. *PloS one*, 13(11):e0205839, 2018.

Sonali Parbhoo, Mario Wieser, Aleksander Wieczorek, and Volker Roth. Information bottleneck for estimating treatment effects with systematically missing covariates. *Entropy*, 22(4), 2020. ISSN 1099-4300. URL https://www.mdpi.com/1099-4300/22/4/389.

Reuven Y Rubinstein. Simulation and the monte carlo method. Technical report, 1981.

Michael S Saag, Constance A Benson, Rajesh T Gandhi, Jennifer F Hoy, Raphael J Landovitz, Michael J Mugavero, Paul E Sax, Davey M Smith, Melanie A Thompson, Susan P Buchbinder, et al. Antiretroviral drugs for treatment and prevention of hiv infection in

adults: 2018 recommendations of the international antiviral society–usa panel. *Jama*, 320(4):379–396, 2018.

Hiroto Saigo, Andre Altmann, Jasmina Bogojeska, Fabian Müller, Sebastian Nowozin, and Thomas Lengauer. Learning from past treatments and their outcome improves prediction of in vivo response to anti-hiv therapy. *Statistical applications in genetics and molecular biology*, 10(1), 2011.

Peter Schulam and Suchi Saria. Reliable decision support using counterfactual models. In *Advances in Neural Information Processing Systems*, pages 1697–1708, 2017.

Naftali Tishby and Noga Zaslavsky. Deep learning and the information bottleneck principle. *CoRR*, abs/1503.02406, 2015.

Naftali Tishby, Fernando C Pereira, and William Bialek. The information bottleneck method. *arXiv preprint physics/0004057*, 2000.

WHO World Health Organisation. The global hiv aids epidemic. 2018. URL https://www.who.int/gho/hiv/.

Summary statistics of patients from EuResist and African cohorts

Characteristics	**EuResist** N (%)	**Africa** N (%)
Sex		
Male	19743 (59.9)	1917(63.9)
Female	13216 (40.1)	1083(36.1)
Age at start		
$18-30$	757(2.3)	1245(41.5)
$31-40$	3659(11.1)	993(33.1)
$41-50$	8504(25.8)	378(12.6)
>50	20040(60.8)	384(12.8)
Ethnicity		
White	6724(20.4)	465(15.5)
Black	13975(42.4)	2 355(78.5)
Asian	890(2.7)	36(1.2)
Hispanic	2011(6.1)	30(1.0)
Other	9360(28.4)	114(3.8)
Risk		
MSM	7613(23.1)	765(25.5)
Heterosexual	7185(21.8)	1449(48.3)
IDU	8801(26.7)	336(11.2)
Other	9361(28.4)	450(15.0)
Baseline CD4$^+$ count		
<200	18985(57.6)	1926(64.2)
$200-349$	8701(26.4)	393(13.1)
>350	5274(16.0)	681(22.7)
Has AIDS at start	8965(27.2)	1947(64.9)
Baseline Log viral RNA		
≤ 4	791(2.4)	462(15.4)
$4-5$	13250(52.8)	366(12.2)
>5	18919(57.4)	2172(72.4)
HIV Subtype		
B	15853(48.1)	1080(30.6)
C	4977(15.1)	1092(36.4)
Recombinants	7185(21.8)	690(23.0)

Table 3: Summary statistics of patients undergoing ART between 1983 and 2016 in the EuResist and African cohorts respectively.

The History Alignment Model

The history alignment model (Bogojeska et al., 2012) first constructs a *resistance mutations kernel* to quantify the pairwise similarities between different therapy combinations. Formally, the kernel may be defined as follows. Let G denote the set of different drug groups, and and u_{ag} and $u_{a'g}$ be the sets of resistance-relevant mutations for the drugs occurring in drug group $g \in G$ of the therapies a and a', respectively. The pairwise similarity between the drug-g mutations of the drug combinations a and a' is then calculated using the Jaccard index:

$$sim_g(a, a') = \frac{|u_{ag} \bigcap u_{a'g}|}{|u_{ag} \bigcup u_{a'g}|}, \tag{6}$$

where $|\cdot|$ denotes set cardinality. We then derive the similarity $k_m(a, a')$ between the therapies a and a' by averaging the similarities of their corresponding drug groups:

$$k_m(a, a') = \sum_{g \in G} \frac{sim_g(a, a')}{|G|}. \tag{7}$$

The resistance mutations kernel is subsequently used together with the Needleman Wunsch algorithm to deduce a kernel over patient histories. This kernel can subsequently be used to perform non-parametric policy learning.

Additional Results and Experiments

Performance of training a POMDP (additional baseline): In addition to comparing our proposed approach against using a kernel trained on the South African cohort and an IB trained only on the South African cohort, we also tried to train a POMDP model solely on the South African cohort. Note that it is not possible to train the POMDP on the EuResist cohort and transfer this policy to the case where data are missing at test time, since we would need these dimensions to estimate and update our belief state. As a result, the performance of the POMDP is far worse since there is limited data to learn from and training the model becomes difficult. We used a similar POMDP to the one of the experts in the original mixture-of-experts model of (Parbhoo et al., 2017) with 20 states and Gaussian emissions. Unfortunately, however since it is difficult to construct a POMDP using a limited source of data, the POMDP does not perform well against the other models.

Performance of methods when training and testing on EuResist: We can perform a similar experiment with training and testing only on the European cohort. All the baselines perform considerably better given the fact that the data set is well sized and genetic information is available for all patients.

Performance of methods when training and testing on African cohort: Repeating the same experiments but while trying to train and test on African cohorts is significantly worse for all baselines as the data set is limited in size. Specifically, all the methods that involve training neural networks perform very poorly here as a result. Here a non-parametric kernel estimate works best.

Type	Method	DR	IS	WIS
Behaviour Policy	5.02 ± 1.18			
Local	Kernel	3.56 ± 1.42	1.27 ± 1.14	1.80 ± 1.07
	CEIB	3.29 ± 1.13	3.80 ± 2.41	3.76 ± 2.19
	POMDP (SA)	-1.47 ± 2.38	-1.72 ± 2.71	-1.14 ± 2.08
Transfer	Kernel	4.17 ±1.4	4.18 ±1.20	4.16 ±1.71
	CEIB	6.29 ± 0.14	5.17 ± 0.38	5.27 ± 0.29
	Mixture-of-Experts	5.28 ± 0.37	3.42 ±1.39	4.81 ± 1.25
Local + Transfer	**Ours**	**8.96 ± 0.39**	**10.64 ± 1.2**	**10.62 ±1.67**

Table 4: Off-policy evaluation using importance sampling, weighted importance sampling and doubly robust methods for different therapy selection models across the African cohort. Overall, combining a local expert with an expert that performs a transfer of the distribution of relevant covariate information for cases where data are missing, produces the best outcomes.

Type	Method	DR	IS	WIS
Behaviour Policy	5.6 ± 1.18			
Local	Kernel	9.47 ± 1.5	5.61 ± 1.41	6.7 ± 1.36
	CEIB	**10.24 ± 1.18**	**11.61 ± 1.41**	**12.76 ± 1.51**
	POMDP	6.07 ± 2.46	4.81 ± 2.75	6.92 ± 1.78
	Mixture-of-Experts	**11.28 ± 0.21**	**12.42 ±1.19**	**11.83 ± 1.45**
	Ours	**12.73 ± 0.39**	**10.79± 1.2**	**11.71 ±1.38**

Table 5: Off-policy evaluation using importance sampling, weighted importance sampling and doubly robust methods for different therapy selection models across the Eu-Resist cohort. Both the CEIB method and the combination of CEIB and Kernel (Ours) produce the best outcome.

Type	Method	DR	IS	WIS
Behaviour Policy	5.02 ± 1.18			
Local	**Kernel**	**5.56 ± 1.02**	**4.7 ± 0.28**	**4.23 ± 0.15**
	CEIB	1.29 ± 1.13	-3.80 ± 2.41	-3.76 ± 2.19
	POMDP	-1.47 ± 2.38	-1.72 ± 2.71	-1.14 ± 2.08
	Mixture-of-Experts	-2.17 ± 3.86	-1.05 ±1.03	0.17 ±0.5
	Ours	3.96 ± 1.39	3.64 ± 1.66	3.82 ± 2.7

Table 6: Off-policy evaluation using importance sampling, weighted importance sampling and doubly robust methods for different therapy selection models across the African cohort. The kernel policy outperforms each of the other baselines here. All the approaches do worse in comparison to the behaviour policy.