Contrasting Temporal Bayesian Network Models for Analyzing HIV Mutations

Pablo Hernandez-Leal, Lindsey Fiedler-Cameras, Alma Rios-Flores,
Jesús. A González and L. Enrique Sucar
{pablohl,lfiedlerc,jagonzalez,esucar}@inaoep.mx
Instituto Nacional de Astrofísica, Óptica y Electrónica
Coordinación de Ciencias Computacionales
Sta. María Tonantzintla, Puebla, México

Abstract

Evolution is an important aspect of viral diseases such as influenza, hepatitis and the human immunodeficiency virus (HIV). This evolution impacts the development of successful vaccines and antiviral drugs, as mutations increase drug resistance. Although mutations providing drug resistance are mostly known, the dynamics of the occurrence of those mutations remains poorly understood. A common graphical model to handle temporal information are Dynamic Bayesian Networks. However, other options to address this problem exist. This is the case of Temporal Nodes Bayesian Networks. In this paper we used both approaches for modeling the relationships between antiretroviral drugs and HIV mutations, in order to analyze temporal occurrence of specific mutations in HIV that may lead to drug resistance. We compare the strengths and limitations of each of these two temporal approaches for this particular problem and show that the obtained models were able to capture some mutational pathways already known (obtained by clinical experimentation).

1 INTRODUCTION

Viral evolution is an important aspect of the epidemiology of viral diseases such as influenza, hepatitis and human immunodeficiency virus (HIV). HIV is the causal agent for the disease known as Acquired Immunodeficiency Syndrome (AIDS), a condition in which progressive failure of the immune system allows opportunistic life-threatening infections to occur.

This viral evolution impacts the development of successful vaccines and antiviral drugs, as mutations caused by viral evolution increase drug resistance. Although the mutations which result in drug resistance are mostly known, the dynamics of the appearance of those mutations and the time of occurrence remains poorly understood.

Bayesian Networks (BNs) have proven to be successful in various domains, including medicine and bioinformatics. However, classical BNs are not well equipped to deal with temporal information. The common approach to handle temporal information is to construct a Dynamic Bayesian Network (DBN) (Dagum, Galper, and Horvitz, 1992), however other options exist such as Temporal Nodes Bayesian Networks (TNBN) (Arroyo-Figueroa and Sucar, 1999).

In this paper we use both approaches, Dynamic Bayesian Networks and Temporal Bayesian Networks, to model the mutational pathways for four specific HIV antiretrovirals. The objective is to compare the pathways that we obtain with our models against the pathways obtained from experimental testing. In this way, we can see if the models reflect the temporal clinical information reported in many reference sources.

2 BAYESIAN NETWORKS

BNs are directed acyclic graphs used to model conditional dependencies between random variables. The data represented by a BN is typically static, however in many contexts a need arises to model processes whose state variables change throughout the course of time. Dynamic Bayesian Networks evolved to tackle this shortcoming.

2.1 DYNAMIC BAYESIAN NETWORKS

A Dynamic Bayesian Network extends the concept of a Bayesian network to incorporate temporal data. Just as with classic BNs, a static causal model is created to represent a process at a single point in time; multiple copies of this model are then generated for each time

point or *slice* belonging to a temporal range of interest and links between copies are inserted to capture temporal relations.

When modeling dynamic information, DBNs obey the assumption that future states are conditionally independent from past states given the present state (Markov property); additionally they assume that the conditional probabilities which describe the temporal relations between random variables of adjacent time slices do not change (stationary process). By allowing these two basic assumptions, DBNs can offer a more compact model of the dynamic process by defining a 2-time-slice Bayesian network (2-TBN). This 2-TBN can be further unrolled to do inference on the entire temporal range of interest.

The learning of a DBN can be seen as a two stage process (Friedman, Murphy, and Russell, 1998). The first stage refers to the learning of the static model and is done in an identical manner as with classic BNs. The second stage learns the transition network, that is, the temporal relations between random variables of different time slices.

An alternative to DBNs are Temporal Nodes Bayesian Networks (Arroyo-Figueroa and Sucar, 1999) which are another extension of Bayesian Networks.

2.2 TEMPORAL NODES BAYESIAN NETWORKS

TNBNs (Arroyo-Figueroa and Sucar, 1999) were proposed to manage uncertainty and temporal reasoning. In a TNBN, each Temporal Node has intervals associated to it. Each node represents an event or a state change of a variable. An arc between two Temporal Nodes corresponds to a causal-temporal relation. One interesting property of this class of models, in contrast to Dynamic Bayesian Networks, is that the temporal intervals can differ in number and size. So, only one (or a few) instance(s) of each variable is required, assuming there is one (or a few) change(s) of a variable state in the temporal range of interest. No copies of the model are needed, thus compacting the representation without losing expressiveness.

A TNBN is composed by a set of TNs connected by arcs. A TN, v_i, is a random variable characterized by a set of states $\mathbf{S}$. Each state is defined by an ordered pair $S = (\lambda, \tau)$, where λ is the particular value taken by v_i during its associated interval $\tau = [a, b]$, corresponding to the time interval in which the state changes, i.e. change in value occurs. In addition, each TN contains an extra default state $s =$ ('no change', $\emptyset$) with no associated interval. Time is discretized in a finite number of intervals, allowing a different number and duration of intervals for each node . Each interval

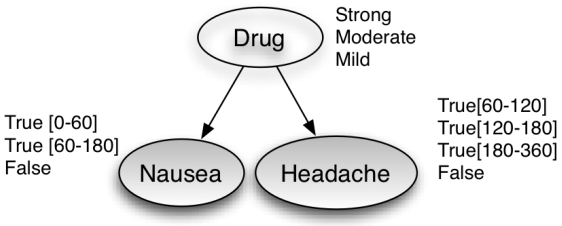


Figure 1: An example of a TNBN. The Drug node is an Instantaneous Node, so it does not have temporal intervals. The Nausea and Headache are temporal nodes with intervals associated to them.

defined for a child node represents the possible delays between the occurrence of one of its parent events and the corresponding child event. If a node lacks defined intervals for all its states then it is referred to as an *instantaneous node*. There is at most one state change for each variable (TN) in the temporal range of interest.

Formally, let $\mathbf{V}$ be a set of temporal and instantaneous nodes and $\mathbf{E}$ a set of arcs between nodes, a TNBN is defined as:

Definition 1. *A TNBN is a pair $B = (G, \Theta)$ where G is a directed acyclic graph, $G = (\mathbf{V}, \mathbf{E})$ and, Θ is a set of parameters quantifying the network. Θ contains the values $\Theta_{v_i} = P(v_i|Pa(v_i))$ for each $v_i \in \mathbf{V}$; where $Pa(v_i)$ represents the set of parents of v_i in G.*

The following is an example of a TNBN of a patient administered with a drug causing two side effects. Its corresponding graphical representation is shown in Figure 1.

Example 1. *Assume that at time $t = 0$, a Drug is administered to a patient. This kind of drug can be classified as* strong, moderate *and* mild. *To simplify the model we will consider only two consequences for the patient, Nausea and Headache. These events are not immediate, we will assume that they depend on the type of drug, therefore, they have temporal intervals associated. For the Nausea node two intervals are defined $[0-60]$, $[60-180]$, for the Headache node three intervals are defined $[60-120]$, $[120-180]$, $[180-360]$. These intervals represent that the state of the node changed during that period of time.*

The learning algorithm for TNBN used in this work has been presented in (Hernandez-Leal, Sucar, and Gonzalez, 2011). We now present a brief description.

1. The algorithm begins by performing an initial discretization of the temporal variables, for example

using an Equal-Width discretization. With this process it obtains an initial approximation of the intervals for all the Temporal Nodes.

2. It then performs a standard BN structural learning using the K2 learning algorithm (Cooper and Herskovits, 1992) to obtain an initial structure. This structure will be used in the third step, the interval learning algorithm.

3. The interval learning algorithm refines the intervals for each TN by means of clustering. For this, it uses the information of the configurations of the parent nodes. To obtain the initial set of intervals a Gaussian mixture model is used as a clustering algorithm for the temporal data. Each cluster corresponds, in principle, to a temporal interval. The intervals are defined in terms of the mean and the standard deviation of the clusters. The algorithm obtains different sets of intervals that are merged and combined, this process generates different interval sets that will be evaluated in terms of the predictive accuracy (Relative Brier Score). The algorithm applies two pruning techniques in order to remove some sets of intervals that may not be useful and also to keep a low complexity of the TNBN. The best set of intervals (that may not be those obtained in the first step) for each TN is selected based on predictive accuracy. When a TN has as parents other Temporal Nodes, the configurations of the parent nodes are not initially known. In order to solve this problem, the intervals are sequentially selected in a top-down fashion according to the TNBN structure.

The algorithm then iterates between the structure learning and the interval learning. However, for the experiments presented in this work, we show the results of only one iteration.

3 HIV AND ANTIRETROVIRAL THERAPY

Viral evolution impacts the development of successful vaccines and antiviral drugs, as mutations (caused by viral evolution) increase drug resistance. In HIV, this is particularly relevant as the virus ranks among the fastest evolving organisms (Freeman, Herron, and Payton, 1998). In viral diseases, such as HIV, it would be important to develop proactive therapies that predict the advent of mutations, thus reducing the possibility of drug resistance, which will then help to predict the duration of a new treatment. Viral therapy failure in patients treated for the HIV-1 infection is commonly associated with the emergence of mutations which are resistant to specific drugs. In addition, a troublesome cross-resistance within the same class of medications complicates the therapeutic options for patients who have treatment regimen failure. Cross-resistance is particularly common among the protease inhibitors (PIs), making the sequential use of these agents frequently problematic. Although the mutations which result in drug resistance are mostly known, the dynamics of the appearance of those mutations on the time of occurrence remains poorly understood.

The relationship between phenotypic susceptibility to some inhibitors and the genotypic pattern was investigated in the same inhibitors. From these studies we now know the resistant patterns associated with the inhibitors most frequently used. This information has led to the ability to select a new salvage therapy. In addition, if we know the pathway and time of occurrence of resistant mutations of common and well known therapies, then this might lead to predicting the duration of new therapies, that use inhibitors that have similar structures or belong to the same class. It is also interesting to compare two or more mutational patterns to see if they share the same mutational pathways, which at the end will help to reduce the possibility of drug resistance.

To combat HIV infection several antiretroviral (ARV) drugs belonging to different drug classes that affect specific steps in the viral replication cycle have been developed. Antiretroviral therapy (ART) generally consists of well-defined combinations of three or four ARV drugs. Due to its remarkable variation capabilities, HIV can rapidly adapt to the selective pressure imposed by ART through the development of drug resistant mutations, that are fixed in the viral population within the host in known mutational pathways. The development of drug resistant viruses compromises HIV control, with the consequence of a further deterioration of the patient's immune system. Many of these ARV drug resistant mutations reduce HIV susceptibility to ARV drugs by themselves, while others need to accumulate in order to cause resistance.

3.1 RELATED WORK

There are several works describing computational models aimed to better understand HIV evolution and immunopathogenesis. A portion of these models is devoted to predict phenotypic HIV resistance to antiretroviral drugs using different approaches such as decision trees (Beerenwinkel et al., 2002) or neural networks (Draghici and Potter, 2003). Other works try to identify relevant associations between clinical variables and HIV disease (Ramirez et al., 2000). In (Chausa et al., 2009), association rules between clinical variables and the failure of the treatment are extracted. The results obtained are temporal rules that have as

Table 1: An example of the HIV patient data. It presents two patients P_1 with 3 temporal studies, and P_2 with two temporal studies.

Patient	Treatment	Mutations	Weeks
P_1	IDV, RTV	L90M, V82A	15
		I54V	45
		M46I	55
P_2	LPV,IDV, RTV	V82A	25
		I54V	45

antecedent the increasing of a subset of clinical variables and as consequent the failure of the treatment, given by side effects of the drugs or by the elevated viral count (unsuccessful therapy). None of the clinical variables considered are HIV mutations. Finally, in (Hernandez-Leal et al., 2011) TNBNs are used to analyze the temporal relationships between all the protease inhibitors and some high frequency mutations. In contrast, the present work uses two different temporal models: DBN and TNBN. Moreover, the experiments presented here are aimed to analyze specific and highly used antiretrovirals (IDV, APV, LPV, RTV), and its corresponding known resistance mutations in order to analyze the mutational pathways.

4 EXPERIMENTS

In this section we present the data used in the experiments, along with the selection method for the drugs and mutations. The first experiment presents the results using a DBN, while the second experiment uses a TNBN. Finally, we contrast the results and models obtained.

4.1 DATA AND PREPROCESSING

Data was gathered from the Stanford HIV Database (HIVDB) (Shafer, 2006) obtained from longitudinal treatment profiles reporting the evolution of mutations in individual sequences.

We retrieved data from patients with HIV subtype B. We choose to work with this subtype because it is the most common in America (Hemelaara et al., 2006), our geographical region of interest. For each patient data retrieved contains a history consisting of a variable number of studies. Information regarding each study consists of an initial treatment (cocktail of drugs) administered to the patient and the list of the most frequent mutations in the viral population within the host at different times (in weeks) after the initial treatment. An example of the data is presented in Table 1.

The number of studies available varies from 1 to 10 studies per patient history. Since we are interested in temporal evolution of the mutational networks, we filtered those patients having less than 2 studies. The filtered dataset consisted of approximately 300 patients.

Antiretrovirals are usually classified according to the enzyme that they target. We focus on protease as this is the smallest of the major enzymes in terms of number of aminoacids. For the experiments we used: Atazanavir (ATV), Lopinavir (LPV), Indinavir (IDV), Ritonavir (RTV). According to the expert's opinion ATV and LPV are the most commonly used antiretrovirals nowadays. IDV was selected because it shares mutational pathways with LPV, and RTV was selected because its frequently used in combination with the other three.

To define the target set of mutations of interest, we used the Major HIV Drug Resistance Mutations according to (Stanford University, 2012). The mutations selected for both experiments are: L90M, V82A, I54V, I84V, V32I, M46I, M46L, I47V, G48V.

Ir order to evaluate the models and to measure the statistical significance of edge strengths we used nonparametric bootstrapping. For this we obtained a reshuffled (re-sampling with replacement) dataset generated from the original dataset and learned the models from this new dataset; this procedure was repeated 10 times. Confidence in a particular directed edge is measured as a percentage of the number of times that edge actually appears in the set of reconstructed graphs. We used two thresholds for considering a relation as important. The first one is a strong relation that appears at least in 90% of the graphs, and the other is a suggestive relation, this occurs with values between 70 and 90%. In Figures 2(a)-2(b) a suggestive relation is shown as an arrow labeled with *, and a strong relationship is presented as an arrow label with **.

4.2 DYNAMIC BAYESIAN NETWORK

In order to obtain the corresponding DBN from the data we began by learning the structure of the static network. For this stage each variable in a patient record was seen to have a binary value, where this value was equal to 1 if that variable was present and 0 if not present. While there are many approaches for learning DBNs such as (Friedman, Murphy, and Russell, 1998; Wang, Yu, and Yao, 2006; Gao et al., 2007) we decided to use a simple approach, therefore the structure of the static network was learned by applying (Chow and Liu, 1968) from which a fully connected tree is obtained. Since the Chow-Liu algorithm does not provide the direction of the arcs, we subsequently applied (Rebane and Pearl, 1987) in conjunction with expert knowledge in order to obtain the final directed acyclic graph. We mention that Rebane and Pearl's

algorithm found the antiretroviral RTV and the mutation L90M to be parent nodes of the antiretroviral IDV; however this relation was found to be invalid and was not established since we know that a mutation cannot be a cause of a medication (expert knowledge). By establishing mutations as effects of medications the directions of the remaining arcs are easily determined.

To obtain the structure of the transition network each record was discretized into equal length time intervals, where the value of a variable was set to 1 if it was present during that time interval and 0 otherwise. Once a variable is observed it remained set to 1 for all subsequent time intervals. If a variable was observed at a time point between state changes, its value was set to 1 for all time intervals greater than the observed time. For our experiments we used different granularities: 5, 8,10 and 20, that corresponds to different number of weeks. The obtained structures were the same except for one new arc in one experiment. When learning the transition network, a node in a time slice can only choose its parents from the previous time slice. In order to choose the best set of parents we applied the Bayesian Information Criterion (BIC) scoring metric to evaluate each selection. This metric returns the probability of the data given the model penalizing the complexity of the model, in other words it favors simpler models. The learning of the transition network was done by using Kevin Murphy's Bayesian Network Toolbox for Matlab (Murphy and others, 2001). Figure 2(a) presents the resulting DBN.

From the model in Figure 2(a) we can observe, that all the nodes, except ATV, have persistent arcs between time slices. In the transition network, arcs from mutations in time slice t appear to be catalysts to the mutations they point to in time slice $t+1$. The static network provides more information on which antiretrovirals are the causing agents of certain mutations. For example, we can observe that the mutation L90M is a reaction to the IDV drug. By following the arcs in the static network and moving through the transition network we can begin to detect mutational pathways.

4.3 TEMPORAL NODES BAYESIAN NETWORKS

To apply the learning algorithm for the TNBN the data is arranged as a table where each column represents a drug or mutation and each row represents a patient case. For the drugs the values are USED or NOT USED, and for the mutations the values are: APPEAR, with the number of weeks in which the mutation appeared for the first time, or NOT, if the mutation did not appear in that case. Thus, the drugs are instantaneous nodes, and the mutations are temporal nodes of the TNBN.

We evaluated different orderings for the K2 algorithm. Specifically, we evaluated all the different combinations for the first two mutations and the order of the rest was chosen randomly. In Figure 2(b) the model with highest predictive accuracy is presented.

The model shows a relation between IDV and RTV, this may suggest that they are mainly used together. ATV is shown isolated from the rest. A reason for this may be that the number of cases that used this drug was low and the algorithm could not find any relations with other drugs or mutations.

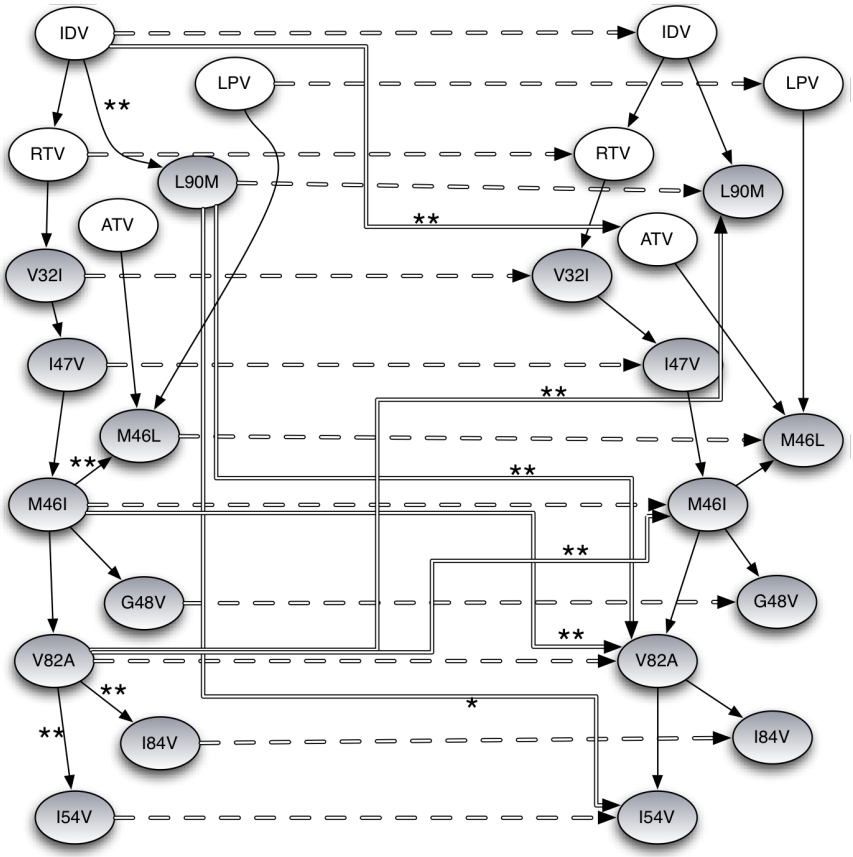
The mutations L90M, I54V and I84V appear to be the first mutations caused by the effects of the drugs. Mutation V82A appears to be important since it has three arcs directed to other mutations. In this model, the mutations M46L, I47V, V32I, V82A and M46I had only a causing mutation as parent. Finally, the mutation G48V appears isolated; this may happen due to the fact that this mutation was infrequent in the data.

4.4 CONTRASTING THE MODELS

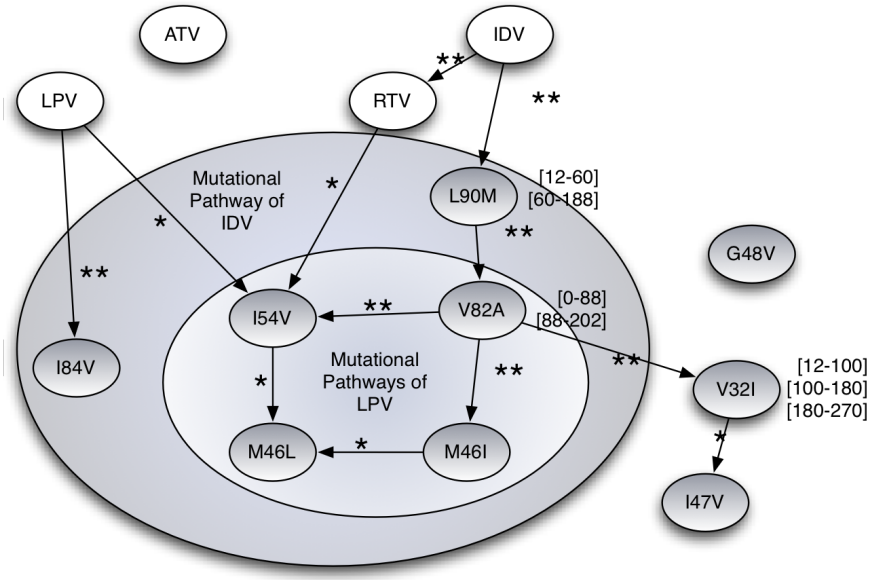
Structure

In order to compare the two models we begin by contrasting the structure displayed by each one. A DBN is typically represented as a 2-TBN in order to give a smaller representation and therefore inference for future times requires the network to be unrolled. Unlike the DBN, a TNBN only has one base network, which can be interpreted as the causal temporal relationships existing between random variables. In a TNBN there is no need to repeat the structure. Therefore, TNBNs offer a more compact representation than DBNs, as DBNs can grow to become increasingly complex, as they are further unrolled to include greater time intervals. Unrolling a DBN can also result in the repetition of nodes whose state has not changed, thus generating unnecessary replications that clutter the model.

The way in which a TNBN is constructed also provides us with different visual information. Given that a TNBN is learned using the K2, the nodes of the resulting model have a temporal ordering, and because the TNBN only consists of one base structure, order of occurrence between different variables is easily visualized. For example, in the context of HIV, pathways formed between mutations can be determined by following the directions of the arcs. In Figure 2(b) we can detect the pathway L90M→V82A→M46I. In contrast, in a DBN temporal orderings are more difficult to visualize solely from the model. However, DBNs offer their own distinct interpretation of the process being modeled. In DBNs, variables that are strongly related can be visualized from the arcs in both the static and



(a) A learned DBN model. Discontinuous lines represent persistent arcs.



(b) A learned TNBN model. Some intervals associated with their respective temporal nodes are shown.

Figure 2: The two temporal models: DBN and TNBN. White nodes represent drugs and grey nodes represent mutations of protease. An arc labeled with a * represents a suggestive relation. An arc with ** represents a strong relation.

transition networks. For example, in Figure 2(a), in time slice t arcs go from mutations L90M and M46I to V82A in time slice $t+1$, indicating a correlation among V82A and both of these other two mutations.

Bootstrapping results

The relations found after performing bootstrapping in the models can be seen in Figures 2(a)-2(b). Both models successfully detected several well known relations among mutations, and while they both coincide in many of these, each one also displays a set of unique relations found. For example, the TNBN detected V82A→V32I as a strong relation, this is consistent with the literature, but was not found by the bootstrapping preformed for the DBN. On the other hand, the DBN was able to detect the relation L90M→I54V; this relation is not present in the TNBN but exists in the literature.

Overall the TNBN was more successful at detecting mutational reactions to specific antiretrovirals. The mutational effects of the medications used are well documented and the TNBN reflects this knowledge. For example, RTV→I54V was found as a suggestive relation; however it is known that when RTV is taken as a booster in combination with IDV, I54V is a common mutational reaction. We note that the TNBN also displays the relation between IDV and RTV.

We also mention that not all relations found by the models have been previously reported. V82A→I84V (found in the DBN) and M46I→M46L (found in both) are as far as the authors know unreported. Further research is needed to determine the correctness of these relations.

Clinical relevance

Both learned models were capable of obtaining known mutational pathways. For example, it is known that the LPV drug causes the pathways:(i) M46I/L, I54V/T/A/S and V82T/F/S (Kempf et al., 2001) , and (ii) V32I, I47V/A, I50V, I54L/M and L76V (Nijhuis et al., 2007; Parkin, Chappey, and Petropoulos, 2003). For IDV the main mutations are V82A/T/F/S/M, M46I/L, I54V/T/A, I84V and L90M (Bélec et al., 2000; Descamps et al., 2005). Moreover M46I/L, I54V/T/A/S and V82T/F/S are reported as major mutations both to IDV and LPV and we can see that in Figures 2(a) and 2(b) the models were able to discover these shared pathways. These results suggest that the models could be applied to new ARVs with structural similarities to determine the duration of the treatment.

5 CONCLUSIONS

Mutational pathways provide important information for decision making in multidrug therapy. In our research, we used HIV data from several patients in order to analyze the temporal occurrence of mutations and create such mutational pathways. We present a comparison of the DBNs and TNBNs models created with this data. Even when Dynamic Bayesian Networks have become a standard for time series modeling, TNBNs offer different advantages. We show why they should be considered as an option when facing problems with dynamic information. Both models were able to capture pathways previously discovered by clinical experiments. These results suggest that temporal BNs are models that can have a significant impact in the battle against the HIV disease. For example, we could use these models to predict mutational pathways and how long new antiretrovirals can be used in specific cases. These models would also help physicians to follow up on patients that are undergoing a therapy that shares similar chemical properties with another treatment whose mutational pathways are already known. As future research, it would be interesting to compare two different cocktail treatments along with the temporal occurrence of drug resistant mutations, in order to predict the most effective treatment. We believe this could aid the experts in the selection of the best treatment for the patient.

Acknowledgements

We would like to thank Dr. Santiago Ávila-Rios and Dr. Gustavo Reyes-Terán from CIENI-INER for their valuable comments and suggestions.

References

Arroyo-Figueroa, G., and Sucar, L. E. 1999. A temporal Bayesian network for diagnosis and prediction. In *Proceedings of the 15th UAI Conference*, 13–22.

Beerenwinkel, N.; Schmidt, B.; Walter, H.; Kaiser, R.; Lengauer, T.; Hoffmann, D.; Korn, K.; and Selbig, J. 2002. Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype. *Proceedings of the National Academy of Sciences of the United States of America* 99(12):8271–8276.

Bélec, L.; Piketty, C.; Si-Mohamed, A.; Goujon, C.; Hallouin, M.; Cotigny, S.; Weiss, L.; and Kazatchkine, M. 2000. High Levels of Drug-Resistant Human Immunodeficiency Virus Variants in Patients Exhibiting Increasing CD4+ T Cell Counts Despite Virologic Failure of Protease Inhibitor Containing Antiretroviral Combi-

nation Therapy. *Journal of Infectious Diseases* 181(5):1808–1812.

Chausa, P.; Cáceres, C.; Sacchi, L.; León, A.; García, F.; Bellazzi, R.; and Gómez, E. 2009. Temporal Data Mining of HIV Registries: Results from a 25 Years Follow-Up. *Artificial Intelligence in Medicine* 56–60.

Chow, C., and Liu, C. 1968. Approximating discrete probability distributions with dependence trees. *Information Theory, IEEE Transactions on* 14(3):462–467.

Cooper, G., and Herskovits, E. 1992. A Bayesian method for the induction of probabilistic networks from data. *Machine learning* 9(4):309–347.

Dagum, P.; Galper, A.; and Horvitz, E. 1992. Dynamic network models for forecasting. In *Proceedings of the 8th Workshop UAI*, 41–48.

Descamps, D.; Joly, V.; Flandre, P.; Peytavin, G.; Meiffrédy, V.; Delarue, S.; Lastère, S.; Aboulker, J.; Yeni, P.; and Brun-Vézinet, F. 2005. Genotypic resistance analyses in nucleoside-pretreated patients failing an indinavir containing regimen: results from a randomized comparative trial. *Journal of clinical virology* 33(2):99–103.

Draghici, S., and Potter, R. B. 2003. Predicting HIV drug resistance with neural networks. *Bioinformatics* 19(1):98–107.

Freeman, S.; Herron, J.; and Payton, M. 1998. *Evolutionary analysis*. Prentice Hall Upper Saddle River, NJ.

Friedman, N.; Murphy, K.; and Russell, S. 1998. Learning the structure of dynamic probabilistic networks. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 139–147. Morgan Kaufmann Publishers Inc.

Gao, S.; Xiao, Q.; Pan, Q.; and Li, Q. 2007. Learning dynamic bayesian networks structure based on bayesian optimization algorithm. *Advances in Neural Networks–ISNN 2007* 424–431.

Hemelaara, J.; Gouws, E.; Ghys, P. D.; and Osmanov, S. 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 20:W13–W23.

Hernandez-Leal, P.; Rios-Flores, A.; Ávila-Rios, S.; Reyes-Terán, G.; González, J.; Orihuela-Espina, F.; Morales, E.; and Sucar, L. 2011. Unveiling HIV mutational networks associated to pharmacological selective pressure: a temporal Bayesian approach. *Probabilistic Problem Solving in BioMedicine* 41.

Hernandez-Leal, P.; Sucar, L. E.; and Gonzalez, J. A. 2011. Learning temporal nodes Bayesian networks. In *The 24th Florida Artificial Intelligence Research Society Conference (FLAIRS-24)*.

Kempf, D.; Isaacson, J.; King, M.; Brun, S.; Xu, Y.; Real, K.; Bernstein, B.; Japour, A.; Sun, E.; and Rode, R. 2001. Identification of genotypic changes in human immunodeficiency virus protease that correlate with reduced susceptibility to the protease inhibitor lopinavir among viral isolates from protease inhibitor-experienced patients. *Journal of Virology* 75(16):7462–7469.

Murphy, K., et al. 2001. The bayes net toolbox for matlab. *Computing science and statistics* 33(2):1024–1034.

Nijhuis, M.; Wensing, A.; Bierman, W.; De Jong, D.; van Rooyen, W.; Kagan, R.; et al. 2007. A novel genetic pathway involving l76v and m46i leading to lopinavir/r resistance. *HIVDRW2007* 12:140.

Parkin, N.; Chappey, C.; and Petropoulos, C. 2003. Improving lopinavir genotype algorithm through phenotype correlations: novel mutation patterns and amprenavir cross-resistance. *Aids* 17(7):955.

Ramirez, J.; Cook, D.; Peterson, L.; and Peterson, D. 2000. Temporal pattern discovery in course-of-disease data. *Engineering in Medicine and Biology Magazine, IEEE* 19(4):63–71.

Rebane, G., and Pearl, J. 1987. The recovery of causal poly-trees from statistical data. In *Third Conference on Uncertainty in Artificial Intelligence (Vol. 87, pp. 222-228)*., volume 86, 222–228.

Shafer, R. 2006. Rationale and uses of a public hiv drug-resistance database. *Journal of Infectious Diseases* 194(Supplement 1):S51–S58.

Stanford University. 2012. Major HIV Drug Resistance Mutations. `http://hivdb.stanford.edu/pages/download/resistanceMutationshandout.pdf`.

Wang, H.; Yu, K.; and Yao, H. 2006. Learning dynamic bayesian networks using evolutionary mcmc. In *Computational Intelligence and Security, 2006 International Conference on*, volume 1, 45–50. IEEE.