Modeling Longitudinal Dynamics of Comorbidities

Basil Maag  
ETH Zurich  
Zurich, Switzerland 
basilmaag@gmail.com

Stefan Feuerriegel  
ETH Zurich  
Zurich, Switzerland 
sfeuerriegel@ethz.ch

Mathias Kraus  
ETH Zurich  
Zurich, Switzerland 
mathiaskraus@ethz.ch

Maytal Saar-Tsechansky  
University of Texas at Austin  
Austin, Texas, USA  
maytal.saar-tsechansky@mccombs.utexas.edu

Thomas Züger  
Inselspital, Bern, University Hospital,  
University of Bern, Bern, Switzerland  
thomas.zueger@insel.ch

ABSTRACT
In medicine, comorbidities refer to the simultaneous occurrence of several diseases affecting patient health [70]. Recent estimates suggest that 50% of hospitalizations in the U.S. involve patients where eleven and more different disease codes are present [71]. Similarly, multiple diseases and, especially, chronic ones simultaneously affect the health of a large proportion of elderly patients [5, 78]. This calls for a better understanding regarding the interaction of diseases, that is, the dynamics behind comorbidities.

Comorbidities are characterized by a concomitant or concurrent evolution, whereby the course of one disease is highly dependent on the course of the other disease [70]. An example are chronic pain and depression. Both often co-occur and their evolution is further highly dependent. Chronic pain often causes a depression and, vice versa, a depression increases typically the perceived level of pain [4, 27]. Another common example of comorbidities are diabetes and chronic liver diseases (e.g., non-alcoholic fatty liver disease) [34, 37, 67]. The evolution of both diseases is known to be subject to profound interactions and metabolic alterations. This is also reflected in a recent proposal to change the nomenclature from non-alcoholic fatty liver disease to metabolic associated fatty liver disease (MAFLD) [21]. To this end, a better understanding of the co-evolution could inform treatment planning and is thus of direct clinical relevance. However, this requires models where the co-evolution is considered, that is, where disease interactions over time are subject to statistical modeling.

Statistical models that are tailored for comorbidities are scarce. One stream of literature has studied which diseases are often found together [32, 40, 73]. However, this question relates to unsupervised learning (e.g., clustering) and does not analyze the longitudinal interactions of the diseases. Another stream of literature is concerned with the impact of comorbidities on patient health [e.g., 11, 27, 51, 59, 66] or readmission risk [71]. However, in these works, comorbidities are merely used as independent variables but not as dependent variables, that is, outcomes of interest. Because of this, such models cannot describe the co-evolution among comorbidities. This motivates our research: how does treating disease A affect the course (disease state) of disease B? To answer this question, statistical inferences are needed that allow for interpretability.

In medicine, the course of diseases is commonly classified by different trajectory phases called “disease states”, whereby diseases undergo different phases of acute and stable states. These disease...
We obtain three main findings. First, modeling disease interactions we specify the Comorbidity-HMM as follows: (1) We account for different states (i.e., acute, stable) in the disease progression by introducing latent states that are of clinical meaning. (2) We model a coupling among the trajectories from comorbidities to capture co-evolution dynamics. (3) We consider between-patient heterogeneity (e.g., risk factors, treatments) in the transition mechanism. We further provide a fully Bayesian estimation framework through Markov chain Monte Carlo (MCMC). Based on our model, we formalize and estimate a spill-over effect which quantifies the indirect effect of treatments through the co-evolution of the diseases. This allows for statistical inferences that ensure interpretability.

Findings: Our Comorbidity-HMM is evaluated based on a large longitudinal dataset. Specifically, we leverage electronic health records from 675 patients over the course of 10 years. The electronic health records document the joint progression of (A) diabetes and (B) chronic liver disease. Patients at risk of developing diabetes mellitus type 2 were chosen as this disease is among the most common diseases worldwide and its presence is often associated with a chronic liver disease as a comorbid condition [34, 37, 67]. We obtain three main findings. First, modeling disease interactions through means coupling results in a superior model fit. Second, an acute state in one of the two diseases is associated with an increased probability of transitioning to an acute state in the other disease. Third, a treatment targeting diabetes has considerable spill-over effects: it is not only associated with a decrease of the risk of an acute disease state in diabetes but also in chronic liver disease. Altogether, the results demonstrate empirically the importance of modeling the co-evolution among comorbidities.

Contributions: Our work advances existing research on modeling longitudinal disease dynamics in the following ways.

1. To the best of our knowledge, our Comorbidity-HMM is the first statistical model that is specifically tailored to capture longitudinal dynamics among comorbidities.

2. For this purpose, we develop a coupled hidden Markov model with a non-homogeneous transition mechanism for longitudinal data and, by deriving the likelihood, provide a Bayesian estimation procedure.

3. We formalize and estimate a spill-over effect that measures the indirect effect of treatments through disease interactions. This allows for direct interpretability as it quantifies empirically how treating disease A affects the course (disease state) of disease B. Thereby, we generate new knowledge for clinical practice (e.g., treatment planning) and research.

2 RELATED WORK

2.1 Comorbidities

In medicine, the term "comorbidity" refers to co-occurring diseases (see [70]). In the literature, comorbidities are sometimes classified into a primary disease and secondary diseases, yet such a classification is not universally applicable [23, 69]. Hence, we later make no such distinction and simply refer to disease A and disease B. Both diseases may simply co-occur by chance, yet they might also be subject to a causal relationship or due to correlated or shared risk factors [69]. This means that interactions among diseases occur. Specifically, by treating one disease, one oftentimes can also expect a better course of co-occurring diseases.

Comorbidities are widespread in clinical practice. In particular, patients who are hospitalized often suffer from multiple comorbidities (e.g., every second hospitalization lists eleven or more different disease codes [71]). Due to their prevalence, comorbidities are also of direct relevance for clinical decision-making. In particular, the presence of comorbidities might worsen health outcomes and, hence, clinical practice emphasizes that their treatment should be given special consideration [30].

In clinical practice, there are numerous examples of comorbidities. For instance, depression and chronic pain occur often jointly, especially their severity is often dependent [4, 27]. Similarly, up to 40% of cancer patients also suffer from psychological distress such as anxiety or depression [81]. Further, patients with diabetes and chronic liver diseases are also frequently diagnosed with additional comorbidities [41], such as hypertension, cardiovascular, or, as studied in this paper, chronic liver diseases (e.g., non-alcoholic fatty liver disease) [34, 37, 67].

2.2 Statistical Inferences using Comorbidity Data

Despite the importance of comorbidities in medicine, there are only a few studies that have incorporated their structure into the model. We provide a summary of key references below. While they also use data on comorbidities, their objective is different from our work.

One stream in the literature has used statistical models to detect co-occurrence patterns among diseases [e.g., 32, 40, 73]. This allows to make associations of which diseases co-occur and should thus be regarded as comorbidities. However, these studies address the question of which comorbidities appear frequently together but they do not model their course. For this different models can be used. For instance, comorbidities can be represented as networks in which nodes correspond to symptoms and edges to potentially causal relationships [16]. Based on the network, shared or overlapping symptoms of distinct disorders then indicate the onset of

---

1Our code is available via https://github.com/mb2019/Comorbidity-HMM
comorbidities. This hypothesis has also been tested using dynamic structural equation models [9, 31] or deep diffusion processes [56]. This answers the question of which diseases co-occur but not how; that is, without providing insights into longitudinal disease interactions. Similarly, Bayesian networks have been used to investigate comorbidities and their temporal dependencies [22, 46, 47]. However, these works offer only little insights into the actual disease progression such as underlying disease states.

Other works use data on comorbidities as input when making inferences regarding patient outcomes [e.g., 11, 27, 59, 66, 83] including readmission risk [71]. This allows to control for the fact that the co-occurrence of diseases often introduces associations with patient outcomes that are of nonlinear form. For instance, patient survival might be lowered when there are many co-occurring diseases. In clinical practice, this is reflected in different measures such as, e.g., the Charlson comorbidity index [13]. However, the aforementioned works consider comorbidities as independent variables and not as dependent variables.

2.3 Hidden Markov Models

Hidden Markov models (HMMs) are a flexible statistical technique for modeling time series. HMMs model a sequence of latent states that follow a Markov chain, based on which, observations are emitted [29]. HMMs have found widespread application in modeling human behavior. Examples include clickstream data [35], urban activities [80], and hiring decisions [45].

HMMs have previously been applied to data on patient health [e.g., 2, 42, 43, 54, 60, 65]. In these works, HMMs are further adapted in order to reflect patient trajectories through following specification. First, the latent states commonly correspond to different diseases states (i.e., phases) in the trajectory [see 14]. Second, covariates as well as random effects are incorporated in the transition mechanism in order to account for patient heterogeneity and other sources of unobserved heterogeneity [2]. However, the aforementioned HMMs describe the evolution of single-disease outcomes but not the evolution in multi-disease settings, that is, comorbidities.

There exist several variants of HMMs [e.g., 64, pp. 76–77]. For instance, coupled HMMs are an extension of HMMs that could potentially model the evolution of multiple interacting processes. These CHMMs allow the latent states of distinct but possibly correlated processes to be conditional dependent [8]. Here the latent states of different sequences can be linked using a Cartesian product, which gives the basis of a so-called Cartesian product CHMM. In a simulation study, such a Cartesian product CHMM was shown to outperform ordinary and multivariate HMMs when the underlying processes are dependent [55].

CHMMs have been applied in various areas, including speech recognition [53], epileptic seizure detection [15], and sleep staging [58]. Furthermore, they have been used to model the spreading of infectious diseases [19, 68]. In [55], a Cartesian product CHMM is developed for modeling the joint progression of lab measurements in a single-disease setting from intensive care units. In ecology, a CHMM was utilized to investigate the interdependence among voles from different forest regions [61]. More specifically, it was examined if the occurrence of one disease makes the contagion of another one more likely. Further, a multi-chain Markov switching model akin to a CHMM was employed to check if there is volatility spillover between geographical financial markets [26].

Research gap: The importance of comorbidities for clinical practice has been widely recognized, and yet statistical frameworks for modeling their co-evolution over time have received little attention. However, understanding the underlying disease dynamics could be of significant value for treatment planning. To this end, we develop a Comorbidity-HMM that is tailored to capture comorbidity dynamics. Specifically, we model the disease interactions among comorbidities in patients based on longitudinal data.

3 PROPOSED MODEL: Comorbidity-HMM

3.1 Problem Statement

We aim to model the co-evolution of comorbidities. For this, we make use of a longitudinal dataset with health information from patients. We further consider two diseases as given by disease A and disease B. Here we explicitly assume that the evolution of both is characterized by disease interactions. This motivates a better understanding of how the course of both diseases interact. Specifically, we seek to answer the question whether there are potential spill-over effects: how does treating disease A affect the state of disease B? For instance, to what extent does Metformin as diabetes treatment influence the risk of an acute state of a co-occurring liver disease? Formally, we want to measure the influence of one disease on the state of the other. We thus develop a model that ensures interpretability in order to allow for such statistical inferences.

To be of clinical relevance, our model specification accommodates further properties, namely (1) disease states, (2) coupling, and (3) between-patient heterogeneity, as described in the following.

(1) Disease states. The course of many diseases (especially chronic or otherwise long-lasting diseases) undergoes different phases over time, where the disease is in an “acute” or “stable” phase [14]. In medicine, the phases are typically termed “disease states”. These states themselves cannot be directly identified and, instead, can only be recovered from measurements [14]. To accommodate this, we follow prior literature [e.g., 6, 50, 54, 60, 62, 65] and model disease states through latent states. We later name the latent states “acute” and “stable” to reflect their clinical meaning. Consistent with prior literature [e.g., 6, 50, 54, 60, 62, 65], we thus build upon the HMM-based framework [57] for modeling the course of diseases A and B. Intuitively, we represent both through two separate HMMs.

(2) Coupling. We expect disease interactions among comorbidities and, therefore, explicitly assume the states of disease A and B to be dependent. As an example, when a disease A transitions from a stable to an acute state, it should also increase the probability of disease B to move from a stable to an acute state. For this reason, we later introduce a coupling between the HMMs from disease A and B.

(3) Between-patient heterogeneity. Disease dynamics are known to vary extensively across patients [e.g., 52, 75]. In medicine, this is described via risk factors (e.g., age, gender) and the presence of treatments [e.g., 36]. Therefore, our model accommodates between-patient heterogeneity as follows. On the one hand, we incorporate patient-level covariates that denote risk factors and treatments.
3.2 Model Specification

Our Comorbidity-HMM consists of the following five components: (i) latent states representing the disease states; (ii) observations representing the measurements or symptoms; (iii) an emission component linking states and observations; (iv) a transition mechanism; and (v) a coupling. Based on this model, we develop a model named Comorbidity-HMM. Our Comorbidity-HMM accommodates the above properties (1)–(3). Based on the Comorbidity-HMM, we define a spill-over effect which estimates the influence of a treatment from disease A on the state from disease B. Here we follow clinical guidelines that base treatment planning primarily on disease states instead of measurements or symptoms [14, 48, 63]. For that reason, we define the spill-over effect with regard to disease states as opposed to measurements or symptoms. This allows us to make statistical inferences of how a treatment targeting disease A render the risk of disease B to become acute vs. stable.

We now define a global latent state based on the Cartesian product of the two separate latent states. Formally, we define $S^{(G)} = (S^{(A)}_t, S^{(B)}_t)$. This is later necessary when we detail the coupling as part of component (v). The underlying mapping is listed in Tbl. 1. For instance, a global state $S^{(G)}_t = (2, 2)$ means that both diseases are in an acute state. As a short form, we use the notation $S^{(G)}_t \in S^{(G)} = \{1, \ldots, 4\}$.

(iii) Emission component: The emission component defines the probability of an observation conditional on the latent state. In our case, it defines the probability of observing a measurement or symptom conditional on disease state. For this, both observations $Y^{(A)}_t$ and $Y^{(B)}_t$ of the two diseases are assumed to depend on the latent states $S^{(A)}_t$ and $S^{(B)}_t$, respectively. Given a latent state $s$, we model them to be conditionally normally distributed via

$$ Y^{(A)}_t \mid S^{(A)}_t = s \sim N(\mu^{(A)}_s, \sigma^{2(A)}_s) $$

$$ Y^{(B)}_t \mid S^{(B)}_t = s \sim N(\mu^{(B)}_s, \sigma^{2(B)}_s) $$

with mean $\mu^{(A)}_s$, $\mu^{(B)}_s$ and variance $\sigma^{2(A)}_s$, $\sigma^{2(B)}_s$. Here the observations are modeled with an individual mean per state $s$ and disease. Hence, an acute state from disease A might have, on average, larger values as observed measurements or symptoms. The variance is disease-specific but independent of the state. In the above specification, we used a normal distribution to model the observations as, in our experiments, the measurements (e.g., blood glucose levels) follow a normal distribution.

(iv) Transition mechanism: The transition mechanism specifies the probability of moving from one latent disease state to another latent disease state. As typical for HMMs, we assume the latent states belong to disease A and B, respectively. As such, each disease has a separate latent state and can thus attain a different state. For each disease, we assume two different states “1” and “2”, i.e., $S^{(A)} = \{ 1, 2 \}$. For ease of notation, we refer to state 1 as “stable” and state 2 as “acute”. In total, this allows for $|S^{(A)}| \times |S^{(B)}|$ different combinations.

Table 1: Mapping of latent states.

| Disease states | Latent states | Global state |
|----------------|---------------|--------------|
| Disease A      | Disease B     | $S^{(A)}_t$  |
| stable         | stable        | (1,1)        |
| stable         | acute         | (1,2)        |
| acute          | stable        | (2,1)        |
| acute          | acute         | (2,2)        |

The choice that risk factors (and treatments) should enter not the emission but the transition mechanism is due to clinical research. The reasoning is the following: variables in the emission only affect observed measurements (e.g., they affect pain resistance) as they moderate how acute vs. stable disease states link to measurements; yet they cannot affect the course of the disease progression. This choice is later validated in our numerical experiments.

The choice that risk factors (as well as treatments) affect the disease dynamics: this allows risk factors (as well as treatments) to increase the propensity of acute vs. stable disease states. On the other hand, we account for unobserved heterogeneity. Here we follow prior literature [7, 17, 20, 33, 38, 76] and control for within effects in the treatment variable (see Sec. 4.2 for details).

In the following, we develop a model named Comorbidity-HMM. Our Comorbidity-HMM accommodates the above properties (1)–(3). Based on the Comorbidity-HMM, we define a spill-over effect which estimates the influence of a treatment from disease A on the state from disease B. Here we follow clinical guidelines that base treatment planning primarily on disease states instead of measurements or symptoms [14, 48, 63]. For that reason, we define the spill-over effect with regard to disease states as opposed to measurements or symptoms. This allows us to make statistical inferences of how a treatment targeting disease A render the risk of disease B to become acute vs. stable.

These are entered in the transition mechanism of the HMMs, so that risk factors (as well as treatments) affect the disease dynamics: this allows risk factors (as well as treatments) to increase the propensity of acute vs. stable disease states. On the other hand, we account for unobserved heterogeneity. Here we follow prior literature [7, 17, 20, 33, 38, 76] and control for within effects in the treatment variable (see Sec. 4.2 for details).

By the choice that risk factors (and treatments) should enter not the emission but the transition mechanism is due to clinical research. The reasoning is the following: variables in the emission only affect observed measurements (e.g., they affect pain resistance) as they moderate how acute vs. stable disease states link to measurements; yet they cannot affect the course of the disease progression. This choice is later validated in our numerical experiments.
where

$$
\eta_{jk} = \begin{cases} 
\alpha_{jk} + \beta_{jk} x_{it}^j, & \text{if } j \neq k, \\
0, & \text{if } j = k.
\end{cases}
$$

(3)

The above logit makes use of further variables: The variable $\alpha_{jk}$ in Eq. 3 corresponds to the intercept of the transition from global state $j$ to $k$. All else equal, it captures how likely a certain disease state transition is. The covariates $x_{it}$ describe the between-patient heterogeneity (e.g., risk factors, treatments). The coefficients of the covariates $x_{it}$ are set to zero in order to ensure identifiability. Hence, transitions where the same latent disease state is maintained (i.e., recurrent transitions) serve as the reference category; see Tbl. 6.

The initial probability of latent state $s$ is given by $\pi_s$ (i.e., the probability of latent states in time step $t = 1$). The initial probabilities were estimated independently of each patient and any risk factors. This approach follows previous works [77].

**Coupling:** We model the interaction of the two diseases by introducing a coupling among their latent states. Specifically, we build upon the concept of a Cartesian product coupled HMM [55]. For this, we draw upon the Cartesian product of the underlying latent states from above. Recall that $S_{it}^{(G)} = (S_{it}^{(A)}, S_{it}^{(B)})$, while $S_{it}^{(G)}$ is referred to as global state. Then we assume that the global states follow a Markov chain (Fig. 1).

The coupling among the latent states is best seen in an example. Recall that the transition mechanism depends on an intercept but it is also influenced by further covariates (i.e., risk factors). We now discuss how the transitions of one disease differ depending on the other diseases. For instance, it might be the case that the probability of a transition from a stable to an acute state for disease B is more likely given that disease A is in an acute state. Such a case could, for instance, arise if the corresponding intercept $\alpha_{34}$ [for transition $(2_{=afe,1_{=stable}}) \rightarrow (2_{=acute,2_{=acute}})$] is substantially larger than the intercept $\alpha_{12}$ [for transition $(1_{=stable,1_{=stable}}) \rightarrow (1_{=stable,2_{=acute}})$]. This means that an acute state for disease A is associated with an increased propensity of disease B also moving to an acute state (rather than a stable state). Thus, comparing the estimated intercepts of such transitions can offer insights about the underlying disease interactions. Hence, we later report our estimations results for $\alpha_{jk}$ as this indicates the strength of the underlying coupling, i.e., the underlying disease interaction.

### 3.3 Spill-Over Effect

Treatments such as medications can also be included in the covariates. If such a treatment is specifically targeted at one comorbidity, it will have a direct effect on the state transitions of that disease. For instance, a drug designed for disease A may prevent the transition from state $(1, 1)$ to $(2, 1)$ (i.e., prevent A from becoming acute). However, the same drug might also implicitly effect the state transitions of disease B through their coupling. We will therefore refer to the former as direct treatment effect and to the latter as spill-over effects.

An example of the a spill-over effect is the following. Let us consider the state transitions $(2, 2) \rightarrow (1, 2) \rightarrow (1, 1)$. It represents the scenario in which a treatment makes disease A transition from an acute state to a stable one (i.e., direct effect). Subsequently, disease B also transitions to a stable state in the following time period (i.e., as an additional indirect effect).

We formalize the above spill-over effect as follows. Let $z$ denote a treatment which is targeted at disease A and affects the transition probabilities via the covariates $x_{it}$ as denoted in Equation 3. Then, the probability of the aforementioned state transitions of patient $i$ under treatment $z$ is given by

$$
\xi_{it}(z) = y_{it}^{(4_{=2})'} \times y_{it+1}^{(2_{=1})'}
$$

$$= P(S_{it+1}^{(G)} = 2 \mid S_{it}^{(G)} = 4, Z = z)
$$

$$\times P(S_{it+2}^{(G)} = 1 \mid S_{it+1}^{(G)} = 2, Z = z).
$$

(4)

Analogously, the probability under no treatment denoted with $z'$ is given by

$$
\xi_{it}(z') = y_{it}^{(4_{=2})'} \times y_{it+1}^{(2_{=1})'},
$$

$$= P(S_{it+1}^{(G)} = 2 \mid S_{it}^{(G)} = 4, Z = z')
$$

$$\times P(S_{it+2}^{(G)} = 1 \mid S_{it+1}^{(G)} = 2, Z = z').
$$

(5)

Thus, if the treatment $z$ increases the probability $\xi_{it}(z)$ but is specifically designed for disease A, it also affects disease B positively through their coupling. Hence, we refer to

$$\text{SpillOver}_{(2, 2)\rightarrow(1, 2)\rightarrow(1, 1)} := \xi_{it}(z) - \xi_{it}(z')
$$

as spill-over effect for $(2, 2) \rightarrow (1, 2) \rightarrow (1, 1)$. If $\xi_{it}(z) - \xi_{it}(z') > 0$, there exists a positive spill-over effect.
Without loss of generality, potential spill-over effects could also arise in different state transitions such as \((1, 1) \rightarrow (2, 1) \rightarrow (2, 2)\) and, hence, the above formulation can be calculated for them analogously.

### 3.4 Model Performance

**Model fit:** We evaluate the fit of our Comorbidity-HMM using standard performance metrics for Bayesian modeling [28], that is, the expected log pointwise predictive density (elpd). The elpd can be efficiently computed using the help of Bayesian leave-one-out cross-validation and Pareto smoothed importance sampling [72]. In our case, we measure model performance for the hold-out sample at patient level (rather than at observation level). For this, we used the loo package as part the statistical software R. Furthermore, we report the widely-applicable information criterion (WAIC).

**Model variants:** We compare our proposed Comorbidity-HMM against alternative model variants. The model variants: (1) A simplified variant where disease \(A\) has only a single latent state, i.e., \(|S(A)| = 1\). (2) A simplified variant where disease \(B\) has only a single latent state, i.e., \(|S(B)| = 1\). These represent simplified disease dynamics with only reduced set of states. Note that both model variants have access to identical data as our Comorbidity-HMM. As we shall see later, the proposed Comorbidity-HMM with \(|S(A)| = |S(B)| = 2\) is superior, thus implying the necessity of accommodating different disease states.

### 3.5 Estimation Procedure

**Sampling:** In the Comorbidity-HMM, all model parameters are obtained via a fully Bayesian estimation [28], specifically through the use of Markov chain Monte Carlo (MCMC) sampling. For this, we used the probabilistic language Stan [12]. Based on it, we obtained posterior estimates from the Hamiltonian Markov chain algorithm together with the No-U-Turn sampler [39]. We ran four chains with 1500 iterations each (1500 additional iterations were discarded as typical hardware amounted to around 12 hours. Switching and to promote identifiability [see 44]. The runtime on typical hardware amounted to around 12 hours.

**Priors:** We used weakly informative priors for the elpd component and the initial state probabilities:

\[
\mu_1(A), \mu_1(B) \sim N(0, 10^2), \quad \sigma_1(A), \sigma_1(B) \sim N(0, 1), \quad \pi_s \sim \text{Dirichlet}(1, 1, 1, 1).
\]

Slightly narrow priors were chosen for the intercepts \(\sigma_{jk}\) with the aim to stabilize the model fit. Instead of specifying priors for the coefficients \(\beta_{jk}\), they were placed on the transformed parameters \(\hat{\beta}_{jk}\), which resulted from a thinned QR decomposition of the centered covariates \(x_{ijt}\). This corresponds to

\[
\sigma_{jk} \sim N(0, 2.5^2), \quad \hat{\beta}_{jk} \sim N(0, 1).
\]

**Diagnostics:** We follow common guidelines in Bayesian modeling [28] and inspect the convergence of the Markov chains. We further conduct posterior predictive checks of whether we can simulate reasonable new observations based on the estimated model parameters [25]. Both are reported as part of the model diagnostics in Sec. B. In addition to that, we checked if our Comorbidity-HMM was able to recover the specified parameters from entirely simulated data. If not stated otherwise, we later report the posterior mean, as well as the 50 % and 90 % credible intervals (CrI).

### 4 EMPIRICAL SETUP

#### 4.1 Data

We evaluate our Comorbidity-HMM based on two co-occurring diseases: (A) diabetes mellitus type 2 and (B) chronic liver disease. Both often co-occur as comorbidities [34, 37, 67]. Diabetes is one of the most common chronic disease, impairing the lives of millions of people worldwide. Diabetes further causes substantial cost during the provision of care [79]. On the other hand, chronic liver disease is a medical condition often co-occurring with diabetes [34, 37, 67]. In our analysis, we use blood glucose level as lab measurements for diabetes and the level of alanine aminotransferase, an enzyme predominantly found in the liver [24, 74], as a lab measurement for chronic liver disease.

We later compute spill-over effects following the treatment of increased blood glucose level with Metformin. Metformin is the first line treatment for diabetes type 2 and lowers overall blood glucose levels. In the literature, Metformin is also considered to have a beneficial impact for patients with liver disease [3, 10, 34]. Hence, this raises the question to what extent spill-over effect influence the disease state for chronic liver disease.

In our analysis, we utilize a longitudinal dataset of patients with prediabetes. The dataset consists of annual electronic health records [1, 82] that also include recordings on lab measurements, treatments, and risk factors (see next section). Overall, the dataset includes 675 patients. Each patient has a time series with recordings between 4 and risk factors (see next section). Overall, the dataset includes 675 patients. Each patient has a time series with recordings between 4 and 10 consecutive years. This amounts to 3253 observations.

#### 4.2 Model Variables

Our Comorbidity-HMM is estimated using the following variables (Tbl. 2).
The observations $y_{it}^{(A)}$ and $y_{it}^{(B)}$ correspond to the glucose level and the alanine aminotransferase level as measurements of diabetes and chronic liver disease, respectively. We log-transform both variables due to their right skewed distributions.

The covariates $x_{it}$ characterize between-patient heterogeneity. These represent risk factors and data on treatments. Here we use the body mass index (BMI), sex, age (at time step $t = 1$), and time-since-prediabetes. We explicitly decompose temporal information into both age and time-since-prediabetes as this allows us to isolate age as a risk factor and potential trends. Data on treatments is given by medication with Metformin. This is a binary variable of whether a Metformin treatment (i.e., Glucomin or Glucophage) was prescribed.

| Table 2: Overview of model variables. |
|-----------------|-----------------|
| Model component | Variables       |
| Observations    |                 |
| Diabetes $y_{it}^{(A)}$ | Blood glucose level (log) |
| Chronic liver disease $y_{it}^{(B)}$ | Alanine aminotransferase level (log) |
| Transition mechanism |                 |
| Covariates $x_{it}$ | BMI, sex, age, time-since-prediabetes, Metformin treatment (centered), Metformin treatment (log=1; centered) |

The following procedure was used to infer the so-called within effect of Metformin. Metformin is generally expected to reduce or stabilize the glucose levels within patients [3]. However, patients who take Metformin might have higher glucose levels compared to other ones. Thus, the so-called within and between effects might be conflicting. For that reason, the Metformin variable is centered with the corresponding patient specific mean in order to infer about the within effect. This approach is consistent with previous literature [7, 17, 20, 33, 38, 76]. Besides the centered Metformin variable, a lagged version of it was also included. These two Metformin variables will be used to estimate potential treatment and spill-over effects over time.

4.3 Summary Statistics

Summary statistics of the model variables are given in Tbl. 7. Blood glucose levels tend to be elevated as compared to recommendations from the World Heath Organization as expected in this population of individuals with (pre-)diabetes. Similarly, alanine aminotransferase values up to 143.40 u/l, suggesting the presence of liver disease in some of the patients [e.g., 34]. Overall, patients are characterized by rather high BMI implying a frequent presence of obesity. This is expected as a high BMI is a known risk factor for diabetes. Only a subset of patients received a Metformin treatment.

We observe a significant correlation between both measurements, namely blood glucose and alanine aminotransferase levels. The Pearson’s correlation coefficient amounts to 0.11 ($p < 0.001$). Hence, a higher blood glucose level coincides with a higher alanine aminotransferase level. This points towards potential interactions among both diseases, yet it does not allow for statistical inferences regarding the underlying longitudinal dynamics.

5 NUMERICAL RESULTS

5.1 Model Performance

The results of the performance comparison are reported in Tbl. 3. The models are evaluated based on the expected pointwise predictive density (elpd). The elpd assesses predictive accuracy in a Bayesian framework and, therefore, the overall fit of a model [28, 72]. A higher elpd indicates a better fit. Among the models which converged, the highest elpd (2,148.78) is registered for the proposed Comorbidity-HMM. It thus achieves the best overall model fit.

We use baseline models that are suitable for jointly modeling multiple time series. Specifically, we draw upon the baselines from [55]. Alternative models are typically designed for single time series or lack a mechanism for disease interactions (see Sec. 2). Hence, their choice for the purpose of our paper is precluded.

Which number of latent states should be selected? We compare our Comorbidity-HMM against baselines with a simplified latent state space, that is, a Comorbidity-HMM where the latent states for either disease $A$ or $B$ are removed. Here simplified $A$ refers to variant with a single latent state for diabetes and, analogously, simplified $B$ to a variant with a single latent state for chronic liver disease. As a result, the diseases are assumed to have only a single state (yet which would oppose the Corbin-Straus trajectory framework [14, 63]). Both the simplified $A$ (elpd: 1,620.96) and a simplified $B$ (elpd: 1,450.83) are outperformed by the proposed Comorbidity-HMM by a considerable margin. Taken together, this confirms the importance of including latent states and thus accounting for different trajectory phases (i.e., disease states). In subsequent sections, we thus report estimation results from the proposed Comorbidity-HMM with $2 \times 2$ states as this model is favored in the model comparison.

How should between-patient heterogeneity be modeled? To answer this, we run additional comparisons. First, we draw upon a naive coupled HMM [8]. This model does not use any covariates for modeling between-patient heterogeneity (i.e., neither in the emission component nor in the transition mechanism). Evidently, the model is inferior. Second, we use the emission coupled HMM from Pohle et al. [55]. This model includes covariates in the emission (whereas our model includes covariates in the transition). However, the chains during MCMC sampling did not converge and, because of this, interpretability for model coefficients is precluded. Moreover, the estimate of the elpd may be unreliable. Third, as a remedy, we modified the emission coupled HMM from Pohle et al. [55]. Here we additionally introduce an ordering of the estimated means besides ordered intercepts in order to prevent label-switching across latent states, set initial values for intercepts based on a k-means procedure [18], and follow common practice in Bayesian modeling whereby the covariates are subject to a QR decomposition. The latter should remove posterior correlations among covariates and thus facilitate convergence of the MCMC chains [28]. To confirm the accuracy of our implementation, we tested the modified emission coupled HMM based on simulated data, where MCMC sampling converged and where the model coefficients were recovered successfully. The model converged successfully for $1 \times 2$ states (but not for the other state combinations). Overall, the model has a poorer fit compared to our proposed model. Altogether, this implies that that covariates
should be included in the transition mechanism (and not in the emission component).

What is the utility of modeling disease interactions? By definition, coupled HMMs with only a single latent state cannot account for disease interactions. Hence, we now assess the corresponding benefit. For the naive coupled HMM, we find that model with $2 \times 2$ states performs best, implying that the capturing disease interactions is beneficial. (We cannot comment on the emission coupled HMM due to divergence in the chains.) For the proposed Comorbidity-HMM, we also find that the model with $2 \times 2$ states is best. Hence, it must be assumed that both diabetes and chronic liver disease are subject to longitudinal disease interactions. This is later also tested statistically by interpreting the transition probability.

5.2 Estimation Results

Emission component: The estimated parameters from the emission component are reported in Tbl. 4. For diabetes, the stable state is associated with an average blood glucose level of only $\exp(4.55) = 94.63$ mg/dl. As expected, the acute state has a higher value, with a mean of $\exp(4.70) = 109.95$ mg/dl. For both, the credible intervals do not overlap and are thus different at a statistically significant level.

For chronic liver disease, the mean of the stable state amounts to $\exp(2.86) = 17.46$ u/l. This is lower than the mean of the acute state (30.88 u/l). The credible intervals of do not overlap and, hence, the posterior means of the stable and acute states are different at the statistically significant level.

Transition mechanism: Estimates of the parameters regarding the state transition probabilities as follows. Due to high number of coefficients, we refrain from presenting them all and rather focus on how Metformin treatment influences the transition probability. This thus quantifies the direct treatment effect of Metformin on the disease state. Informed by clinical research, we expect the following. (1) Recall that Metformin is a diabetes treatment. Hence, it should only have an effect on diabetes and not on chronic liver disease. This is confirmed (all corresponding CrI do not include zero). (2) Metformin has a short-term effectiveness and, hence, we expect to see treatment effects in the non-lagged variable (rather than in the lagged). Again, we yield positive evidence. (3) We expect a positive treatment effect which would be seen in the transition (acute,acute) → (stable,acute). Here a large portion of the probability mass is above zero, indicating a positive treatment effect of Metformin. Here the posterior mean corresponds to an increase of the log-odds of 3.29, resulting in a $\exp(3.29) = 26.84$ odds ratio for a one unit increase in the centered Metformin variable and, hence, making the aforementioned transition more likely than staying in the state (acute,acute).

Coupling: We now investigate the coupling among the two comorbidities. For that, we report the baseline probability in the transition mechanism (i.e., the intercept $\alpha_{ij}^z$). It quantifies the transition conditional on the latent states of both disease A and B. To estimate this, all covariates were set to respective sample mean. This corresponds to $x_{it} = \bar{x}_{it}$ in Eq. 3.

Fig. 2 shows the coupling for diabetes. The plot gives the probability of disease A (diabetes) to transition from a stable to an acute latent state conditional on the latent state of both diseases. If both disease are independent, then the credible intervals for the conditional transition probability should include zero. Larger conditional transition probability values indicate a stronger coupling. From Fig. 2, we find that an acute state for chronic liver disease is associated with an increased probability of moving from stable to acute in diabetes; see (1, 2) → (2, 2). Put simply, if chronic liver disease is acute, then it is likely that diabetes also transitions from stable to acute. Overall, we find strong evidence that diabetes is coupled with chronic liver disease.

The interpretation holds analogously when exchanging A and B. Here similar effects can also be detected in Fig. 3. Put simply, if diabetes is already acute, then it is somewhat likely that chronic liver disease also becomes acute. However, the effect is less pronounced and the 50% and 90% credible interval overlap to a great extent. Nevertheless, these findings suggest that the two comorbidities indeed interact with each other.

Note: Here shown are the transitions probabilities that diabetes moves from a stable to an acute latent state, i.e., $(1, X) \rightarrow (2, X)$. Hence, this compares the transition probability when exiting a stable chronic liver disease (top) and an acute chronic liver disease (bottom). Estimations show the posterior means (thick line), as well as the 50% (shaded area) credible interval. All covariates are set to the respective sample mean during estimation.

Figure 2: Estimation results for coupling conditional on diabetes.

5.3 Estimated Spill-Over Effect

Lastly, spill-over effects are presented. Specifically, we assess the effect of Metformin on liver disease through the co-evolution of the two comorbidities. For this, we consider the state transitions (2, 2) → (1, 2) → (1, 1). It reflects the transition from acute to stable in diabetes followed by the transition from acute to stable in liver disease. In particular, we report the following: (1) Metformin has been prescribed, i.e., $\xi_{it}(z)$. Here the centered variable is set to 0.5, whereby the remaining covariates are also assumed to be equal to the sample mean. (2) No Metformin has been prescribed, i.e., $\xi_{it}(z')$. This corresponds to a centered Metformin variable equal to zero and all remaining covariates equal to the respective sample mean. (3) The spill-over effect, which amounts to the absolute difference between both. (4) A quotient. It denotes how many times more likely a spill-effect is when comparing Metformin vs. no Metformin.
The corresponding quantiles of the transition probabilities of these two scenarios as well as their difference are reported in Tbl. 5.

Overall, we find the following. If no Metformin is prescribed, there is hardly any change in the state for liver disease through the underlying coupling. Here all estimated quantiles are close to zero. If Metformin is prescribed, we observe a non-zero probability that the state from chronic liver disease also improves. By comparing the different quantiles, we observe that the distribution for \( \xi_{1t}(z) \), i.e., Metformin prescribed, has a very long tail. For many patients on the left-tail, the spill-over effect is small or absent (e.g., due to non-adherence or simply because the treatment is not effective). Based on our results, it cannot be ruled out that, for the 5% quantile, the Metformin treatment has also negative implications for a co-occurring liver disease.

However, for the patients on the right-tail, there is a large estimated spill-over effect. For instance, for the 50% quantile, prescribing Metformin (as opposed to non-prescribing) makes a improvement in the liver state 2.898 times more likely. In the 75% quantile, the relative quotient increases even to 13.695. As shown here, our model allows to quantify the spill-over effect of diabetes on liver disease due to the use of Metformin. In sum, we find evidence that Metformin stabilizes not only diabetes but, through the coupling, also a co-occurring chronic liver disease.

### 6 DISCUSSION

**Summary of findings**:
(1) Our results find that our Comorbidity-HMM outperforms alternative model specifications. In particular, it is superior over variants where no disease interaction is modeled. Here we obtain substantial improvements in the model performance as measured by predictive accuracy for Bayesian modeling (i.e., the expected pointwise predictive density, elpd). (2) We provide empirical evidence that both diabetes and chronic liver disease are dependent. This is revealed by the fact that transitions of one disease depend on the state of the other disease at a statistically significant level. Thereby, we confirm the importance of coupling

---

**Table 3: Model performance comparison.**

| Disease interactions | #Latent states | elpd | SE(elpd) | WAIC |
|-----------------------|---------------|------|----------|------|
| Naive coupled HMM [8]  | ✗             | (1 × 2) | 1,622.54 | 150.1 | -3,245.17 |
| Naive coupled HMM [8]  | ✗             | (2 × 2) | 1,430.12 | 140.83 | -2,860.39 |
| Naive coupled HMM [8]  | ✓             | (2 × 2) | 1,333.57 | 148.64 | -2,467.34 |
| Emission coupled HMM (unmodified) [55] | ✗ | (1 × 2) | (1775.55) | (152.19) | (-3491.69) |<sup>†</sup> |
| Emission coupled HMM (unmodified) [55] | ✗ | (2 × 2) | (1462.99) | (148.61) | (-2483.17) |<sup>†</sup> |
| Emission coupled HMM (unmodified) [55] | ✓ | (2 × 2) | (2280.02) | (153.16) | (-4509.16) |<sup>†</sup> |
| Emission coupled HMM (modified) [55] | ✓ | (1 × 2) | 1,932.96 | 148.19 | -3,864.83 |
| Emission coupled HMM (modified) [55] | ✗ | (2 × 2) | (1383.15) | (142.6) | (-2528.34) |<sup>†</sup> |
| Emission coupled HMM (modified) [55] | ✓ | (2 × 2) | (1424.63) | (140.14) | (-2563.72) |<sup>†</sup> |
| Comorbidity-HMM (simplified A) | ✗ | (1 × 2) | 1,620.96 | 149.61 | -3,241.99 |
| Comorbidity-HMM (simplified B) | ✓ | (2 × 1) | 1,450.83 | 139.70 | -2,902.97 |
| Comorbidity-HMM (simplified B) | ✗ | (2 × 2) | 2,148.78 | 147.29 | -4,298.73 |

<sup>†</sup> MCMC sampling did not converge
elpd: expected log pointwise predictive density (higher values are better); SE: standard error
WAIC: widely-applicable information criterion (lower values are better)

---

**Table 4: Posterior estimates of emission component.**

| Latent state Mean | Credible interval |
|------------------|------------------|
|                  | 5%               | 95%               |
| **Diabetes** (blood glucose level; in log) |
| \( \mu_{1(A)} \) stable | 4.55 | 4.54 | 4.55 |
| \( \mu_{2(A)} \) acute | 4.70 | 4.69 | 4.70 |
| \( \sigma_{(A)} \) --- | 0.09 | 0.09 | 0.09 |
| **Chronic liver disease** (alanine aminotransferase level; in log) |
| \( \mu_{1(B)} \) stable | 2.86 | 2.85 | 2.88 |
| \( \mu_{2(B)} \) acute | 3.43 | 3.41 | 3.45 |
| \( \sigma_{(B)} \) --- | 0.30 | 0.29 | 0.31 |

Note: Here shown are the transitions probabilities that chronic liver disease moves from a stable to an acute latent state, i.e., (G1, 1) → (G1, 2). Hence, this compares the transition probability when exiting a stable diabetes (top) and an acute diabetes (bottom). Estimations show the posterior means (thick line), as well as the 50% (shaded area) credible interval. All covariates are set to the respective sample mean during estimation.

**Figure 3: Estimation results for coupling conditional on chronic liver disease.**
in our model formation. (3) Our model returns empirical evidence characterizing the spill-over effect of a diabetes treatment on the state from chronic liver disease.

Prior literature suggested mixed findings on the role of Metformin for patients with chronic liver disease [3, 10, 34], especially as a direct effect would not necessarily be supported by the molecular mechanism of action. Here we contribute empirical evidence whereby Metformin has a direct effect on the state for diabetes, while we find no support for a direct effect on the state for chronic liver disease. Instead, we find an indirect effect through the coupling.

**Interpretability:** Our model is based on a parsimonious specification in order to warrant for interpretability. Here we build upon a central concept from medicine: diseases form a so-called trajectory which undergoes different phases with acute and stable disease states [14, 63]. These disease states are directly relevant for clinical decision-making as different treatment plans should be applied depending on whether a patient is in an acute vs. a stable state. Specifically it suggests that clinical decision-making should tailor care to the underlying disease state (and not measurements or symptoms as these are only noisy observations of the state) [48]. Owing to this, prior research has modeled health trajectories through the use of latent states {e.g., 6, 50, 54, 60, 62, 65}. In line with this, we also introduce latent disease states and use them for coupling the states of different diseases, and, on top of that, we define our spill-over effect with regard to latent states (as opposed to measurements or symptoms).

**Limitations:** As with other research in medicine, ours is not free of limitations and we thus make suggestions for future research. First, our data comprises a unique, large-scale collection of electronic health records from a Western health provider. Future research could replicate our approach with other cohorts. Second, we used an extensive set of risk factors informed by prior literature. Nevertheless, other variables might also represent risk factors and, hence, future research could test how the model performance changes by including them. Third, our analysis is based on specific settings, namely the interaction among diabetes and chronic liver disease under Metformin treatment. Here future research could validate our model based on other comorbidities. If desired, the model can also be extended to comorbidities comprising of three (or more) diseases by expanding the underlying latent state space.

**Implications:** Our model offers insights for clinical practice and research. By inferring spill-over effects, one can assess to what extent treatments for disease A also change the course (disease state) of disease B. Hence, when seeking to stabilize the course of disease B, clinical practitioners can thus decide upon a treatment with a direct effect on B or, by contrast, choose a treatment plan for A that also has an indirect effect such that disease B is stabilized. Here our model offers statistical insights, allowing for evidence-based decisions regarding treatment planning.

Our model was tested based on diabetes and chronic liver disease. However, our model is designed in a generic manner in order to ensure widespread applicability. The clinical co-author from our work foresees several use cases: Here our model could analyze comorbidities that are poorly understood or where there are mixed findings in the clinical literature. Given the vast number of comorbidities, this gives rich opportunities for follow-up research.

### 7 CONCLUSION

Comorbidities, defined as the presence of multiple co-occurring diseases, are widespread among patients, and yet comprehensive statistical frameworks for modeling the longitudinal dynamics of comorbidities are rare. Such models would allow for novel insights into the co-evolution of comorbidities over time. Specifically, this would allow to answer the question: \textit{how does treating disease A change the state of disease B?} To address this, we developed a probabilistic model for longitudinal data analysis: a coupled hidden Markov model with personalized, non-homogeneous transitions (Comorbidity-HMM). To the best of our knowledge, our Comorbidity-HMM is the first statistical model that is specifically tailored to capture longitudinal dynamics among comorbidities. Our model further ensures interpretability by offering clinically-relevant inferences. In particular, we presented a spill-over effect that measures the indirect effect of a treatment on disease states through the co-development of comorbidities. We evaluated our model based on a longitudinal dataset from 675 patients. We found that Comorbidity-HMM outperforms alternative model specifications in terms of model fit.

### ACKNOWLEDGMENTS

Financial supported from the Swiss National Science Foundation (SNSF) through an Eccellenza grant (186932) is acknowledged.

### REFERENCES

[1] Ardavan Afshar, Ioakeim Perros, Haesun Park, Christopher deFilippi, Xiaowei Yan, Walter Stewart, Joyce Ho, and Jimeng Sun. 2020. TASTE: Temporal and Static Tensor Factorization for Phenotyping Electronic Health Records. In Proceedings of the ACM Conference on Health, Inference, and Learning (Toronto, Ontario, Canada) (CHIL’20). Association for Computing Machinery, New York, NY, USA, 193–203.

[2] Rachel MacKay Altman. 2007. Mixed Hidden Markov Models. J. Amer. Statist. Assoc. 102, 477 (2007), 201–210.

[3] Clifford J. Bailey and Robert C. Turner. 1996. Metformin. New England Journal of Medicine 334, 9 (1996), 574–579. PMID: 8569826.

[4] Matthew J. Baur, Rebecca L. Robinson, Wayne Katon, and Kurt Kroenke. 2003. Depression and Pain Comorbidity: A Literature Review. Archives of Internal Medicine 163, 20 (2003), 2433–2445.

[5] Karen Barnett, Stewart W Mercer, Michael Norbury, Graham Watt, Sally Wyke, and Bruce Guthrie. 2012. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 380, 9836 (2012), 37–43.

[6] Nicola Bartolomeo, Paolo Trovatioli, and Gabriella Serio. 2011. Progression of liver cirrhosis to HCC: an application of hidden Markov model. BMC medical research methodology 11, 1 (2011), 38–38.

### Table 5: Estimated spill-over effects.

| Metformin prescribed ($f_1(z_1)$) | 5%  | 25%  | 50%  | 75%  | 95%  |
|-----------------------------------|-----|------|------|------|------|
|                                  | 0.000 % | 0.002 % | 0.017 % | 0.089 % | 0.627 % |
| No Metformin prescribed ($f_0(z’_1)$) | 0.000 % | 0.002 % | 0.006 % | 0.016 % | 0.044 % |
| Difference | -0.019 % | -0.002 % | 0.008 % | 0.077 % | 0.619 % |
| Quotient | 0.022 | 0.463 | 2.898 | 13.695 | 91.276 |

Note: The quantities of the probabilities of the transition (2, 2) \rightarrow (1, 2) \rightarrow (1, 1) are reported. Metformin prescribed assumes that the centered Metformin variables is equal to 0.5 and the remaining covariates are equal to the sample mean. In contrast, no Metformin prescribed corresponds to a centered Metformin variable equal to 0 and the remaining covariates also set to the sample mean. Additionally, the quantiles of the difference of these transition probabilities are given.
[51] Michael Mueller-Peltzer, Stefan Feuerriegel, Anne Molgaard Nielsen, Alice Kongsved, Werner Vach, and Dirk Neumann. 2020. Longitudinal healthcare analytics for disease management: Empirical demonstration for low back pain. Decision Support Systems (2020), Article 113271.

[52] Katherine L. Musliner, Trine Munk-Olsen, Thomas M. Laursen, William W. Eaton, Peter P. Zandi, and Preben B. Mortensen. 2016. Heterogeneity in 10-Year Course Trajectories of Moderate to Severe Major Depressive Disorder: A Danish National Register-Based Study. JAMA Psychiatry, 73, 4 (04 2016), 346–355.

[53] A. V. Nefian, L. Liang, Xi P. Li, Xiaoxiang C. Mao, and K. Murphy. 2002. A coupled HMM for audio-visual speech recognition. In 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing. Vol. 2. –II–2013–II–2016.

[54] Yilmazcan Oezyurt, Mathias Kraus, Tobias Hatt, and Stefan Feuerriegel. 2021. A primer on coupled state-switching models for multiple interacting time series. Statistical Modelling (2020).

[55] Zhaoqi Qian, Mahmoud A. Alaa, Alexis Bellot, Jem Rashbash, and Mihaela van der Schaar. 2020. Learning Dynamic and Personalized Comorbidity Networks from Event Data using Deep Diffusion Processes. arXiv:2001.02585 [cs.LG]

[56] L. R. Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77, 2 (Feb 1989), 257–286.

[57] I. Rezek and S. J. Roberts. 2000. Estimation of coupled hidden Markov models with application to biosignal interaction modelling. In Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No.00TH8501). Vol. 2. 804–813 vol.2.

[58] Lisa K. Richardson, Leonard E. Egede, Martina Mueller, Carrae L. Echols, and L. R. Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. In 1989 IEEE Workshop on Neural Networks for Signal Processing. Vol. 2.–2002–II–2020.

[59] Steven L Scott, Gareth M James, and Catherine A Sugar. 2005. Hidden Markov Models for Longitudinal Comparisons. J. Amer. Statist. Assoc. 100, 420 (2005), 359–369.

[60] Chris Sherlock, Tatiana Xifara, Sandra Teller, and Mike Begon. 2013. A coupled hidden Markov model for disease interactions. Journal of the Royal Statistical Society: Series C (Applied Statistics) 62, 4 (2013), 669–627.

[61] Jennifer L. Wolff, Barbara Starfield, and Gerard Anderson. 2002. Prevalence, Expenditures, and Complications of Multiple Chronic Conditions – A Two-State Hidden Markov Regression Approach. Biometrical Journal 43, 8 (2001), 941–962.

[62] Ryan Zabora, Karlynn BrintzenhofeSzoc, Barbara Curbow, Craig Hooker, and Steven Piantadosi. 2001. The prevalence of psychological distress by cancer site. Psycho-Oncology 10, 1 (2001), 19–28.

[63] Wei Zhang, Zhaobin Kuang, Peggy Peissig, and David Page. 2020. Adverse Drug Reaction Discovery from Electronic Health Records with Deep Neural Networks. In Proceedings of the ACM Conference on Health, Inference, and Learning, Toronto, Ontario, Canada (CHIL ’20). Association for Computing Machinery, New York, NY, USA, 3363–3369.

[64] Hamed Majidi Zolbanin, Dursun Delen, and Amir Hassan Zadeh. 2015. Predicting overall survivability in comorbidity of cancers: A data mining approach. Decision Support Systems 74 (2015), 150 – 161.

[65] Walter Zucchini, Iain L MacDonald, and Roland Langrock. 2017. Hidden Markov models for time series: an introduction using R. CRC press.

[66] Basil Maag et al. 2015. Distinct HaA1c trajectories in a type 2 diabetes cohort. Acta Diabetologica 52, 2 (2015), 267–275.

[67] Lijuan (Peggy) Wang and Scott E. Maxwell. 2015. On Disaggregating Between-Person and Within-Person Effects With Longitudinal Data Using Multivariate Hidden Markov Models. Psychological Methods, 20, 1 (2015), 63–83.

[68] Peiming Wang and Martin L. Puterman. 2001. Analysis of Longitudinal Data of Epgistic Seizure Counts – A Two-State Hidden Markov Regression Approach. The Annals of Family Medicine 30, 4 (2008), 599–514.

[69] Petra J. M. Elders, Annette C. Moll, Bettine C. P. Polak, Jacqueline M. Dekker, and Giel Nijpels. 2015. Distinct HbA1c trajectories in a type 2 diabetes cohort. Acta Diabetologica 52, 2 (2015), 267–275.

[70] Worl Health Organization et al. 2016. Global report on diabetes. (2016).

[71] Peter P. Zandi, and Preben B. Mortensen. 2016. Heterogeneity in 10-Year Course Trajectories of Moderate to Severe Major Depressive Disorder: A Danish National Register-Based Study. JAMA Psychiatry, 73, 4 (04 2016), 346–355.

[72] Jose M. Valderas, Barbara Starfield, Bonnie Sibbald, Chris Salisbury, and Martin Roland. 2009. Defining Comorbidity: Implications for Understanding Health and Health Services. The Annals of Family Medicine 7, 4 (2009), 357–363.

[73] Marjan van den Akker, Frank Buntinx, and J André Knottnerus. 1996. Comorbidity. Expenditures, and Complications of Multiple Chronic Conditions in the Elderly. The World Wide Web Conference 16879.

[74] I. Rezek and S. J. Roberts. 2000. Estimation of coupled hidden Markov models for disease interactions. Journal of the Royal Statistical Society: Series C (Applied Statistics) 62, 4 (2013), 669–627.

[75] Jennifer L. Wolff, Barbara Starfield, and Gerard Anderson. 2002. Prevalence, Expenditures, and Complications of Multiple Chronic Conditions – A Two-State Hidden Markov Regression Approach. Biometrical Journal 43, 8 (2001), 941–962.

[76] Ryan Zabora, Karlynn BrintzenhofeSzoc, Barbara Curbow, Craig Hooker, and Steven Piantadosi. 2001. The prevalence of psychological distress by cancer site. Psycho-Oncology 10, 1 (2001), 19–28.

[77] Wei Zhang, Zhaobin Kuang, Peggy Peissig, and David Page. 2020. Adverse Drug Reaction Discovery from Electronic Health Records with Deep Neural Networks. In Proceedings of the ACM Conference on Health, Inference, and Learning, Toronto, Ontario, Canada (CHIL ’20). Association for Computing Machinery, New York, NY, USA, 3363–3369.

[78] Hamed Majidi Zolbanin, Dursun Delen, and Amir Hassan Zadeh. 2015. Predicting overall survivability in comorbidity of cancers: A data mining approach. Decision Support Systems 74 (2015), 150 – 161.

[79] Walter Zucchini, Iain L MacDonald, and Roland Langrock. 2017. Hidden Markov models for time series: an introduction using R. CRC press.
A ADDITIONAL TABLES

Table 6: Transition matrix (with corresponding $\alpha$).

| Latent states from | to | 1  | 2  | 3  | 4  |
|-------------------|----|----|----|----|----|
| (1,1)             |    | $\alpha_{12}$ | $\alpha_{13}$ | $\alpha_{14}$          |
| (1,2)             | $\alpha_{21}$ |    | $\alpha_{23}$ | $\alpha_{24}$          |
| (2,1)             | $\alpha_{31}$ | $\alpha_{32}$ |    | $\alpha_{34}$          |
| (2,2)             | $\alpha_{41}$ | $\alpha_{42}$ | $\alpha_{43}$ |    |

Table 7: Summary statistics.

| Model variable                          | Overall ($N = 3253$) |
|-----------------------------------------|----------------------|
| Blood glucose level [in mg/dl]          | Mean (SD) 99.59 (11.68) |
|                                        | Range 62.00 – 211.00  |
| Alanine aminotransferase level [in u/l] | Mean (SD) 24.23 (11.43) |
|                                        | Range 4.05 – 143.40  |
| BMI                                     | Mean (SD) 30.55 (5.86) |
|                                        | Range 15.48 – 66.67  |
| Sex                                     | Male 1194 (36.7%)    |
|                                        | Female 2059 (63.3%)  |
| Age [in years]                          | Mean (SD) 55.34 (8.27) |
|                                        | Range 24.00 – 70.00  |
| Metformin treatment                     | Not prescribed (= 0) 3080 (94.7%) |
|                                        | Prescribed (= 1) 173 (5.3%) |

Note: Traceplot for 1500 iterations of each chain (warmup of MCMC algorithm omitted).

Figure 4: Traceplots of emission means $\mu_s^{(A)}$ and $\mu_s^{(B)}$.

We further assessed our model by conducting posterior predictive checks following recommendations on Bayesian modeling [25, 28]. For this, new observations were simulated based on the posterior draws of the model parameters as well as the most probable latent states which were estimated using the Viterbi algorithm. The results are in Fig. 5. Based on it, we compare the credible intervals of the generated observations against the actual observations. The results suggest that our model fits the data well. We further checked whether Comorbidity-HMM is able to recover the specified parameters from simulated data. Here we obtain confirmatory results, suggesting that it can successfully capture the underlying dynamics.

B MODEL DIAGNOSTICS

We follow common practice in Bayesian modeling [28] by performing the following model diagnostics. This is to ensure convergence of the MCMC algorithm and thus precise estimates. First, we inspected the effective sample size $n_{eff}$, indicating that the number of MCMC samples is sufficient. Second, we calculated the Gelman-Rubin convergence diagnostic $\hat{R}$ of all model parameters. The $\hat{R}$ is below the critical threshold of 1.1, suggesting convergence of the MCMC chain. Third, we retrieved traceplots (Fig. 4). The traceplots show the MCMC draws of the emission mean across the latent states and difference diseases (i.e., $\mu_s^{(A)}$ and $\mu_s^{(B)}$, $s \in \{1, 2\}$). The traceplots suggest that the chains have mixed well. Altogether, the previous diagnostics provide ample evidence that the MCMC algorithm has converged.
Figure 5: Posterior predictive checks (indicating the accuracy of the model on simulated data).