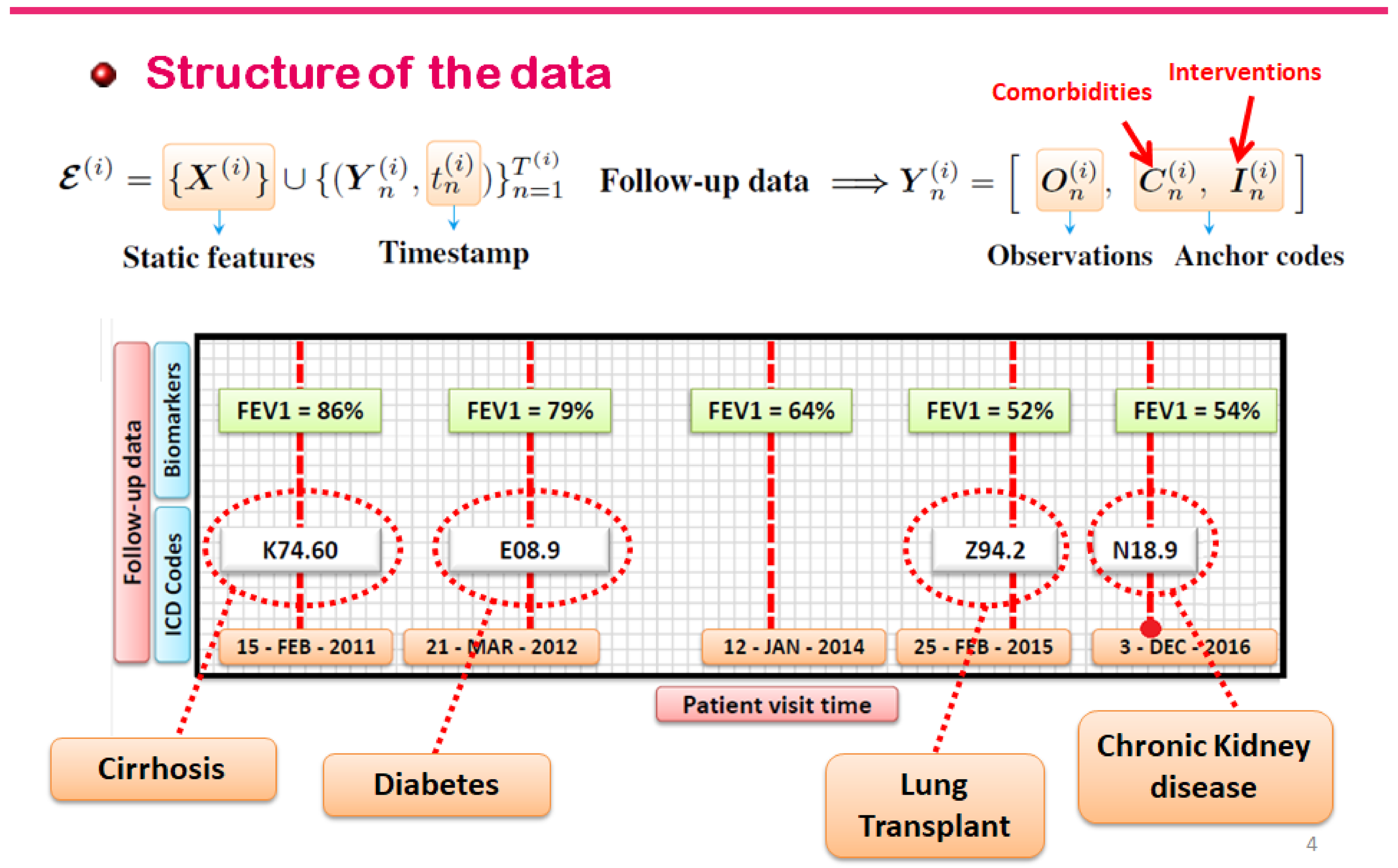
Forecasting Disease Trajectories in Cystic Fibrosis with Deep Learning

Bryan Lim[1], Thomas Daniels[2], Andres Floto[3,4] and Mihaela van der Schaar[1,5]
[1] Dept. of Engineering, University of Oxford, [2] University Hospital Southampton,
[3] Dept. of Medicine, University of Cambridge, [4] Royal Papworth Hospital, Cambridge, [5] Alan Turing Institute

Motivations

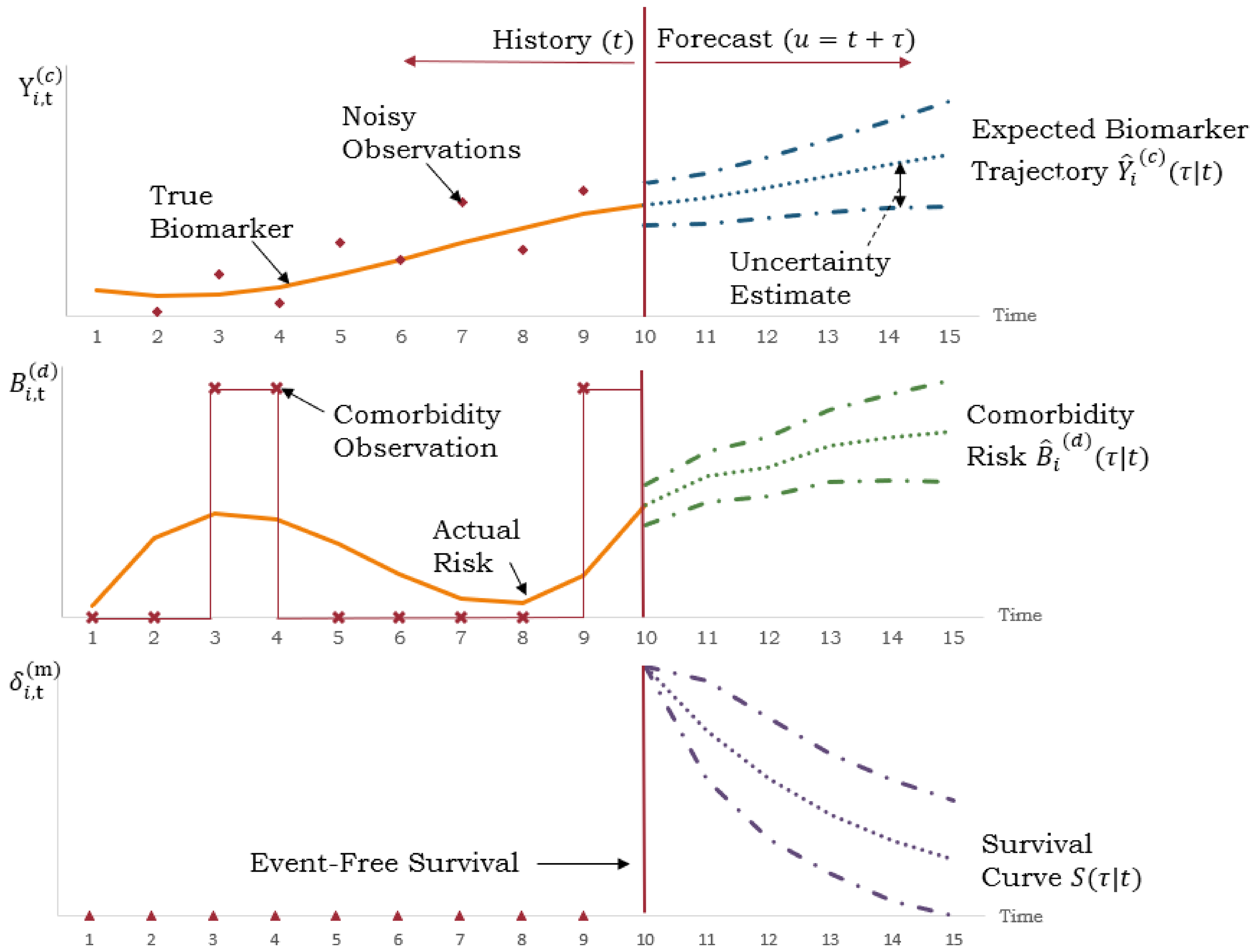
Chronic disease management is challenging!

- Managing chronic diseases requires balancing **multiple clinical outcomes of interests** – taking into account both quality of life and treatment considerations.
- Multiple factors need to be considered during the decision making process to help inform decisions around **frequency of testing**, **intensity of treatment**, and **treatment burden**.
- Patients require monitoring at **infrequent intervals** but **over a very long duration** – e.g. annual follow ups over **multi-year** horizons.



Solution – Disease-Atlas: New deep learning model to **simultaneously forecast** multiple outcomes over time, incorporating uncertainty estimates to reflect model confidence.

Disease-Altas Output Types



Patient Data

Data Description

- Data obtained from UK Cystic Fibrosis Trust for a cohort of **10980** patients with annual follow ups between 2008 – 2015.
- Each patient associated with **87 variables**.
- Joint predictions of **death**, **2 lung function** scores, **9 comorbidities** and **11 infections**.

Accuracy of Forecasts

Performance Metrics

- **FEV** forecasts evaluated in terms of mean squared error (MSE).
- **Mortality, comorbidity and infection** prediction accuracy measured by the area under the precision-recall curve (AUPRC).
- Benchmarks used - **deep neural networks** (i.e. Long Short Term Memory Networks (LSTM)) and **traditional statistical methods** (i.e. landmarking & joint models (JM))

Mortality & FEV1 Predictions

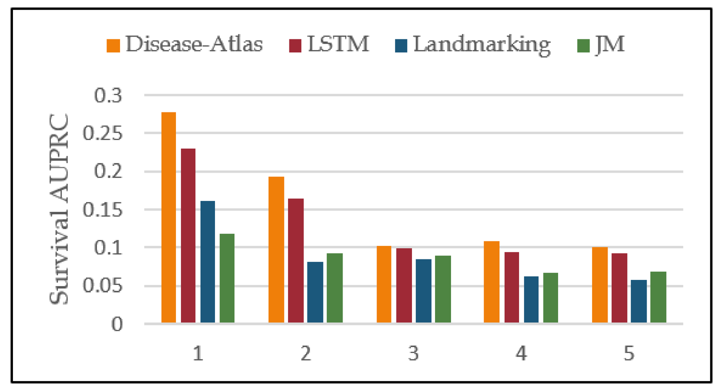


Figure 1: Mortality AUPRC

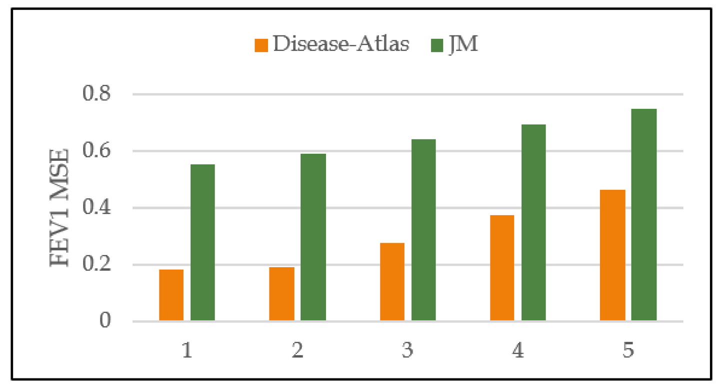


Figure 2: FEV1 MSE

Comorbidity Predictions (AUPRC)

	Disease-Atlas					Joint Models				
Prediction Horizon (Yrs)	1	2	3	4	5	1	2	3	4	5
Liver Disease	0.862	0.825	0.709	0.616	0.513	0.181	0.186	0.197	0.2	0.207
Asthma	0.904	0.845	0.773	0.642	0.544	0.272	0.261	0.258	0.245	0.24
Arthropathy	0.799	0.76	0.621	0.5	0.347	0.134	0.142	0.148	0.155	0.154
Bone Fracture	0.064	0.043	0.052	0.032	0.031	0.006	0.007	0.007	0.009	0.01
Raise Liver Enzymes	0.784	0.726	0.536	0.409	0.338	0.163	0.16	0.156	0.157	0.172
Osteopenia	0.758	0.742	0.648	0.577	0.526	0.245	0.255	0.266	0.278	0.28
Osteoporosis	0.658	0.644	0.507	0.406	0.322	0.144	0.149	0.151	0.146	0.134
Hypertension	0.308	0.34	0.309	0.277	0.227	0.123	0.13	0.141	0.142	0.142
Diabetes	0.85	0.798	0.774	0.663	0.64	0.319	0.334	0.342	0.348	0.356
Average	**0.665**	**0.636**	**0.548**	**0.458**	**0.388**	**0.176**	**0.180**	**0.185**	**0.187**	**0.188**
Standard Deviation	**0.287**	**0.269**	**0.237**	**0.205**	**0.189**	**0.093**	**0.094**	**0.096**	**0.097**	**0.099**

Infection Predictions (AUPRC)

	Disease-Atlas					Joint Models				
Prediction Horizon (Years)	1	2	3	4	5	1	2	3	4	5
Burkholderia Cepacia	0.692	0.672	0.639	0.576	0.471	0.054	0.058	0.056	0.056	0.062
Pseudomonas Aeruginosa	0.84	0.828	0.815	0.8	0.794	0.636	0.641	0.65	0.655	0.649
Haemophilus Influenza	0.369	0.332	0.265	0.243	0.278	0.181	0.204	0.233	0.231	0.202
Aspergillus	0.38	0.315	0.337	0.27	0.293	0.22	0.22	0.218	0.212	0.216
NTM	0.237	0.073	0.181	0.133	0.138	0.076	0.068	0.072	0.062	0.041
Ecoli	0.506	0.242	0.089	0.036	0.008	0.098	0.037	0.025	0.011	0.005
Klebsiella Pneumoniae	0.299	0.146	0.06	0.01	0.015	0.051	0.037	0.026	0.025	0.027
Gram-Negative	0.028	0.038	0.022	0.027	0.022	0.009	0.01	0.012	0.012	0.015
Xanthomonas	0.298	0.202	0.218	0.18	0.128	0.079	0.079	0.087	0.092	0.098
Staphylococcus Aureus	0.771	0.706	0.612	0.537	0.497	0.336	0.337	0.344	0.347	0.345
ALCA	0.153	0.148	0.155	0.144	0.175	0.037	0.04	0.037	0.04	0.047
Average	**0.416**	**0.337**	**0.308**	**0.269**	**0.256**	**0.162**	**0.157**	**0.160**	**0.158**	**0.155**
Standard Deviation	**0.259**	**0.274**	**0.265**	**0.259**	**0.246**	**0.184**	**0.190**	**0.195**	**0.197**	**0.195**

Decision Support With Disease-Atlas

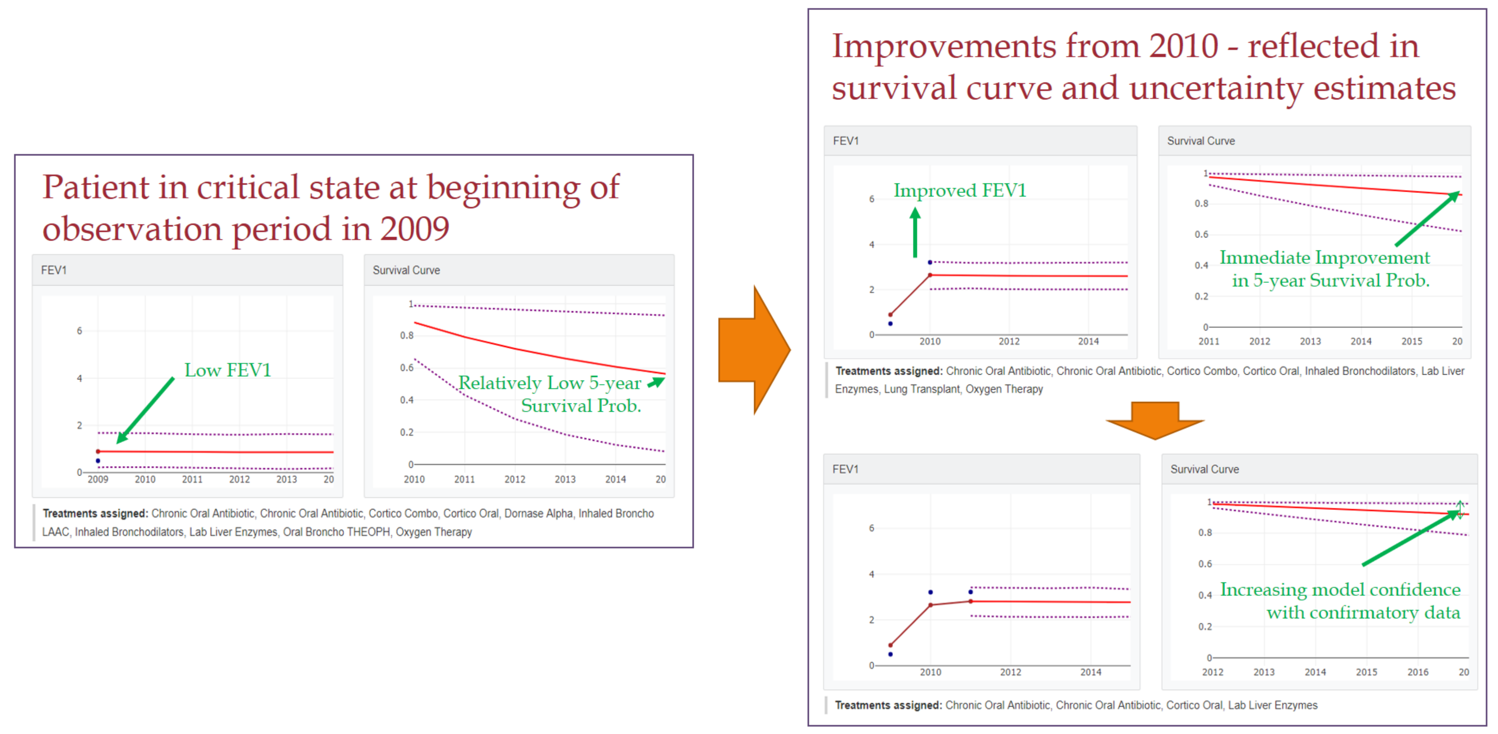
Web App Available!

`https://disease-atlas-online.herokuapp.com`

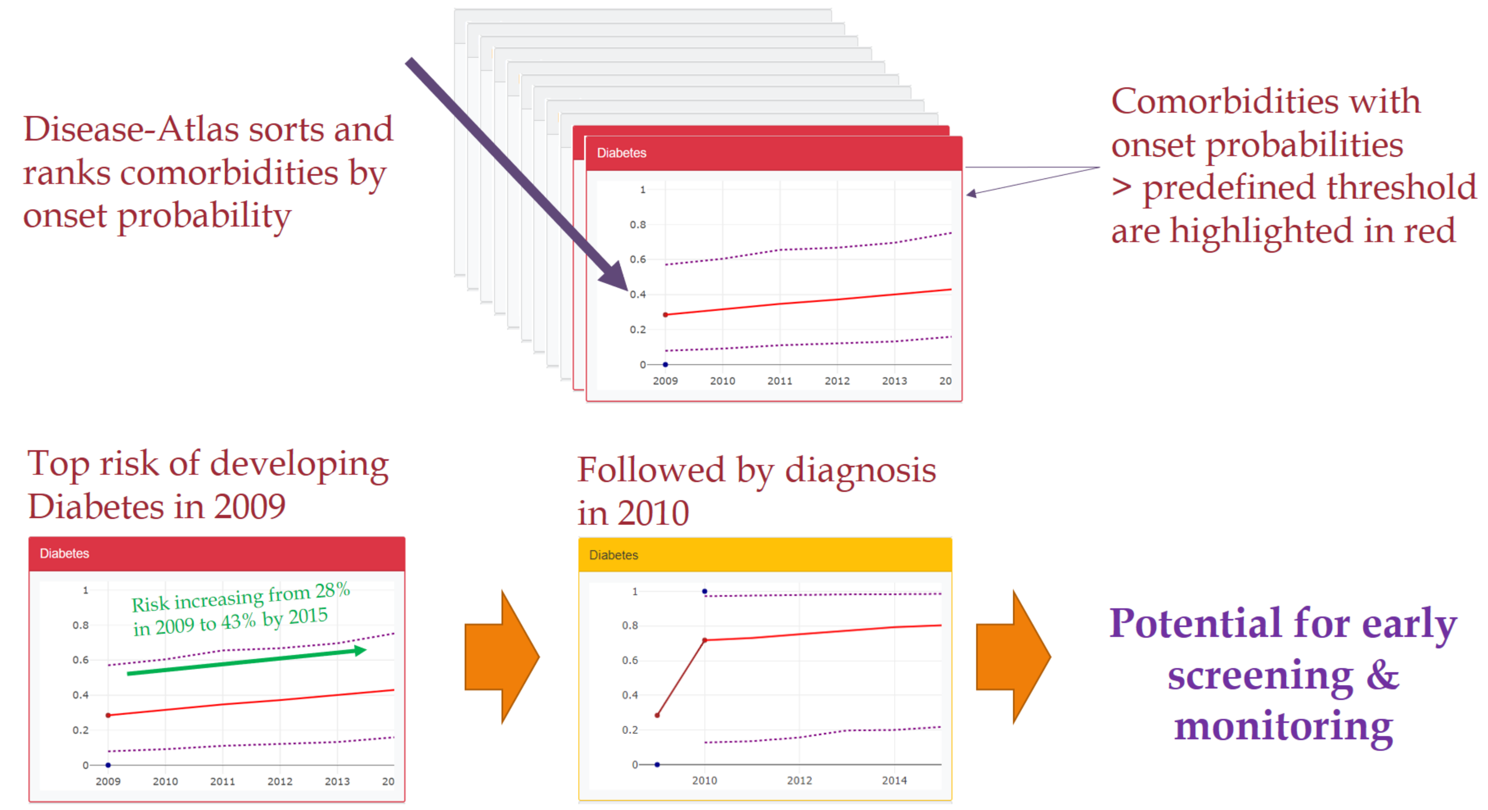
Test Patient

In the following use cases, we consider the Disease-Atlas forecasts for a **27-year old male patient** with follow ups between **2009 - 2015**.

Use Case 1 – Patient Monitoring



Use Case 2 – Prioritising Screening



Conclusions

- Large, high quality datasets provide opportunities for **machine learning methods to predict future health events** for people with CF.
- Translating predictive analytics into **real-time decision support tools** can facilitate **personalised data-driven decision making** in CF clinics – leveraging the individuals' own data and associated risks.
- The methodology should also be validated on a different cohort.