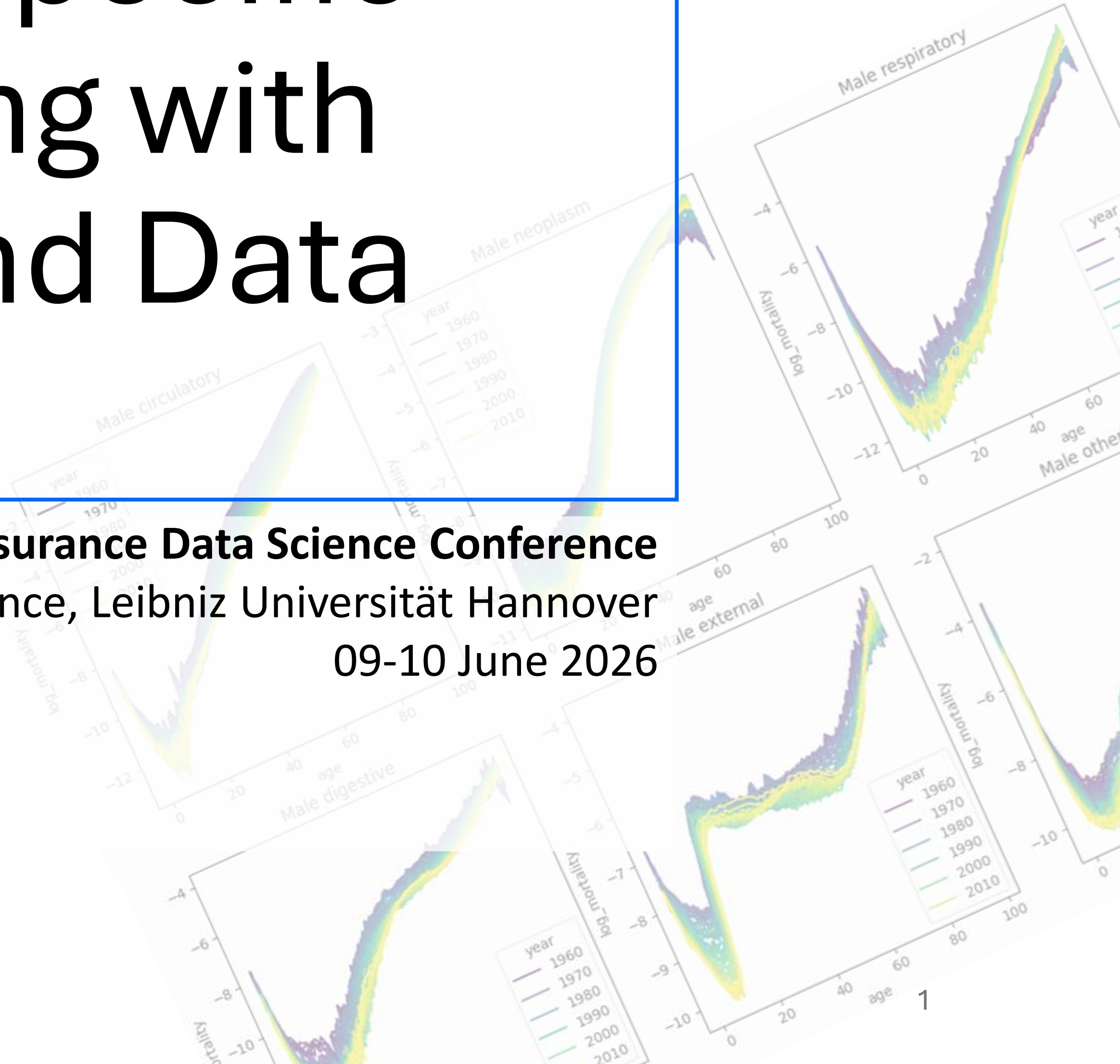
Advancing Cause-Specific Mortality Forecasting with Neural Networks and Data Augmentation

Séverine Arnold, HEC Lausanne
Michael Rockinger, HEC Lausanne
Dion Krisnadi, HEC Lausanne
Andrés Villegas, UNSW Sydney
Katja Hanewald, UNSW Sydney

Insurance Data Science Conference
House of Insurance, Leibniz Universität Hannover
09-10 June 2026



Unil.

Main Question

Can we leverage underlying cause-of-death data in order to improve aggregate mortality forecasts?

How well can neural networks model cause-of-death mortality rates?

Main Question

Can we leverage underlying cause-of-death data in order to improve aggregate mortality forecasts?

How well can neural networks model cause-of-death mortality rates?

- Feed-forward neural network (FFNN)
- Data augmentation within transfer learning framework

Literature Reviews

- Modelling cause-of-death mortality:
  - Independent assumption
    Villegas et al., 2021;
  - Multinomial logistic regression
    Alai et al., 2014
  - Co-integration
    Arnold and Glushko, 2022
  - Multi-population LC family
    Varga, 2025
  - Aggregation constraints
    Dong et al., 2025; Graziani and Nigri, 2026
  - Neural Network (as feature extractor)
    Tanaka and Matsuyama, 2025

- Modelling mortality with NN:
  - NN to extract features to be fed to classical time series approach
    La Rocca et al., 2023;
    Perla & Scognamiglio, 2023
  - NN to handle end-to-end mortality forecast
    Richman & Wüthrich, 2021;
    C. W. Wang et al., 2021;
    Perla et al., 2021;
    Euthum et al., 2024;
    Mori et al., 2025

Literature Reviews

- Modelling cause-of-death mortality:
  - Independent assumption
    Villegas et al., 2021;
  - Multinomial logistic regression
    Alai et al., 2014
  - Co-integration
    Arnold and Glushko, 2022
  - Multi-population LC family
    Varga, 2025
  - Aggregation constraints
    Dong et al., 2025; Graziani and Nigri, 2026
  - Neural Network (as feature extractor)
    Tanaka and Matsuyama, 2025

- Modelling mortality with NN:
  - NN to extract features to be fed to classical time series approach
    La Rocca et al., 2023;
    Perla & Scognamiglio, 2023
  - NN to handle end-to-end mortality forecast
    Richman & Wüthrich, 2021;
    C. W. Wang et al., 2021;
    Perla et al., 2021;
    Euthum et al., 2024;
    Mori et al., 2025

U.S. Cause-of-Death Dataset

(Early version of the HMD Human Cause-of-Death Data)

Year	Age	Cohort	Group	ICD	Gender	Deaths	Diseases	Level 1	Level 2
1959	0	1959	51	7	Female	63	Cerebrovascular disease	Circulatory diseases	CVD and stroke
1959	1	1958	30	7	Female	59	Leukemia	Neoplasms	Other cancers
1959	2	1957	58	7	Female	250	Pneumonia	Respiratory diseases	Influenza and pneumonia
1959	3	1956	63	7	Female	2	Gastric and duodenal ulcer	Digestive system	Gastric and duodenal ulcer
1959	4	1955	85	7	Female	15	Accidental falls	External causes	Other external causes
1959	5	1954	3	7	Female	6	Infectious Gastro-enteritis	Other	Rest of causes
							...		
2017	94	1923	50	10	Male	2043	Other forms of heart disease	Circulatory diseases	Other circulatory system diseases
2017	95	1922	18	10	Male	183	Trachea, bronchus, and lung	Neoplasms	Lung cancer
2017	96	1921	60	10	Male	3	Asthma	Respiratory diseases	Chronic lower respiratory disease
2017	97	1920	64	10	Male	9	Gastro-enteritis (non-infectious)	Digestive system	Other digestive system diseases
2017	98	1919	88	10	Male	22	Other accidents	External causes	Other external causes
2017	99	1918	74	10	Male	35	Other diseases of the genitourinary system	Other	Rest of causes

Figure 1. Excerpt of cause-of-death dataset

Methodology: Cause-of-Death modelling
(Baseline FFNN Architecture)

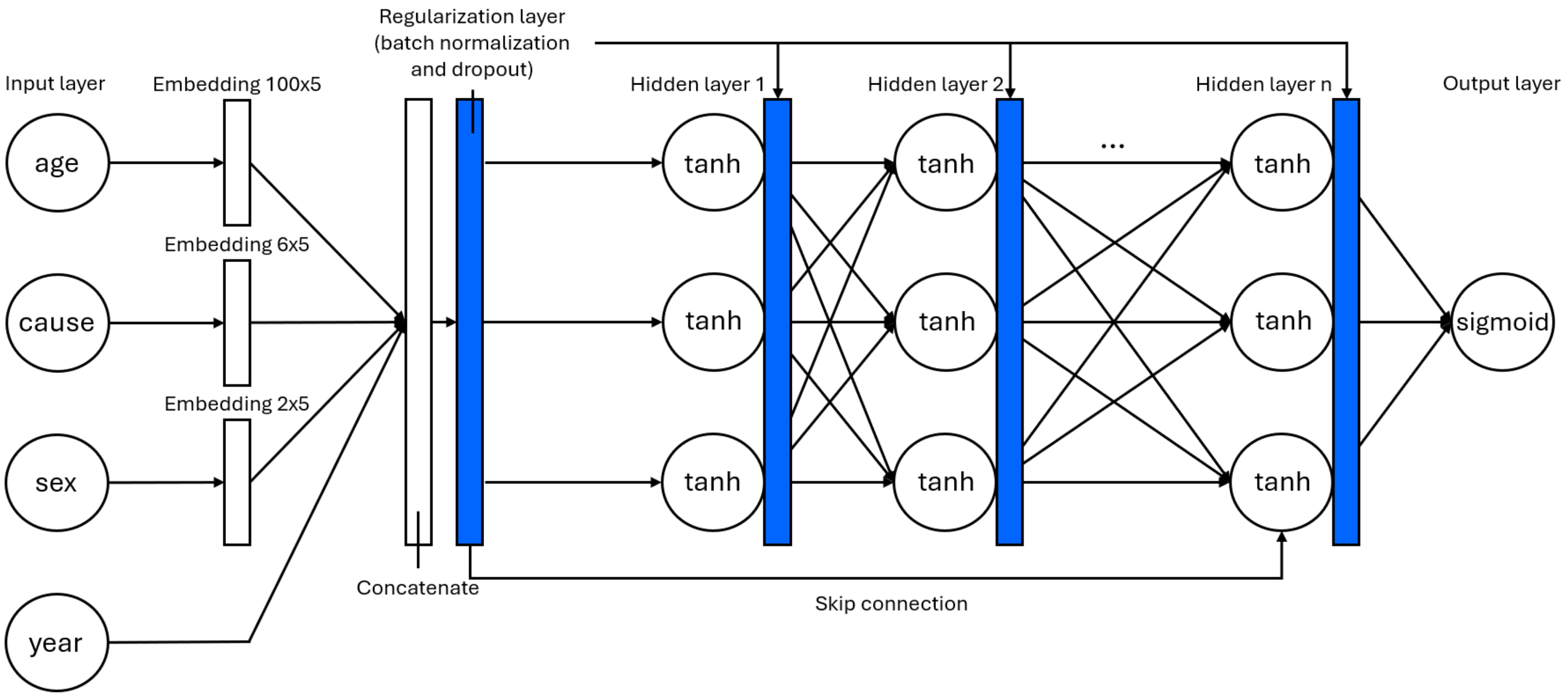


Figure 2. Baseline FFNN architecture, adapted from [Richman & Wüthrich, 2021](#) (72,253 parameters)

Methodology: Data Augmentation
(Poisson Bootstrap Simulation)

- **Base data:**
  - 49,200 total observations across year t, gender g, age x and cause c.
- **Data augmentation:**
  1. Model death counts using a Poisson distribution, $D^{c}_{(x,g,t)} \sim Poisson\left(\lambda^{c}_{(x,g,t)}\right)$
  2. Set $\lambda^{c}_{(x,g,t)}$ equals the observed deaths.
  3. Generate 100 independent Poisson random samples for each cell.

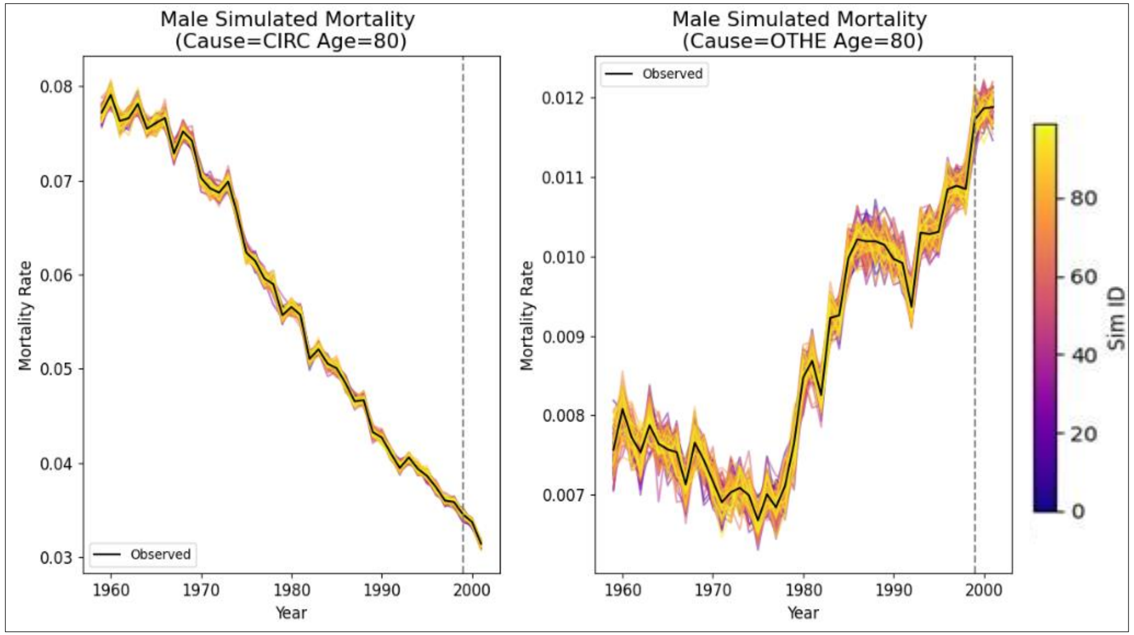


Figure 3. Examples of augmented data for circulatory and 'other' male age 80 across training years, black lines represent observed rates while different colors represent 100 different realizations.

Methodology: Transfer Learning
(Hyperparameters)

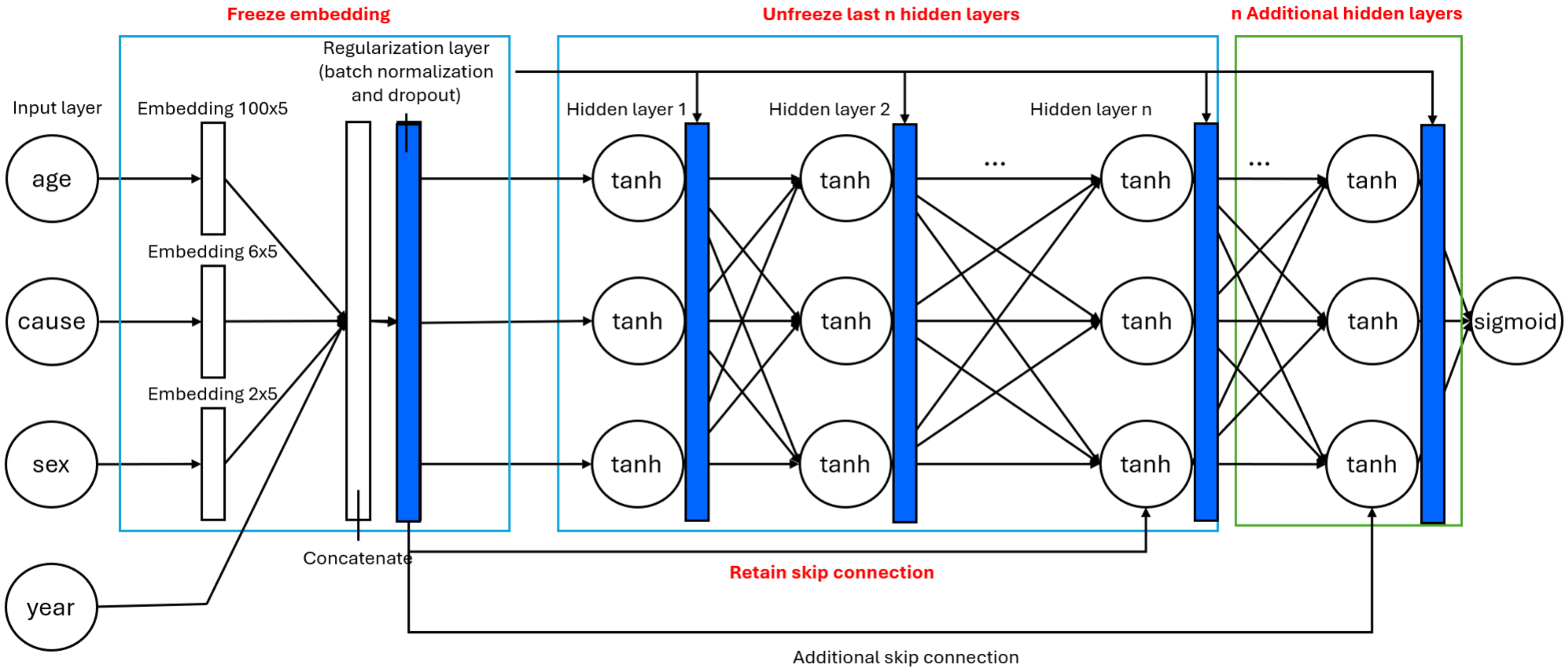


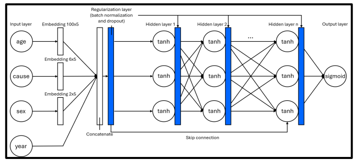
Figure 4. Fine-tuned FFNN architecture within transfer learning framework

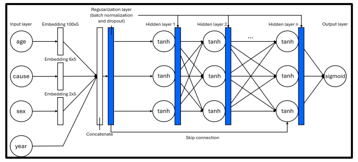
Methodology: Transfer Learning
(Robust Ensemble Strategy)

Pre-trained Checkpoints
(augmented data)

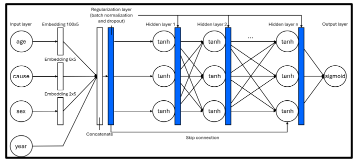
Fine-tuned Models
(observed data)

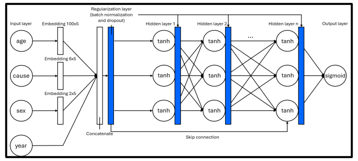
First-30 Ensemble

1


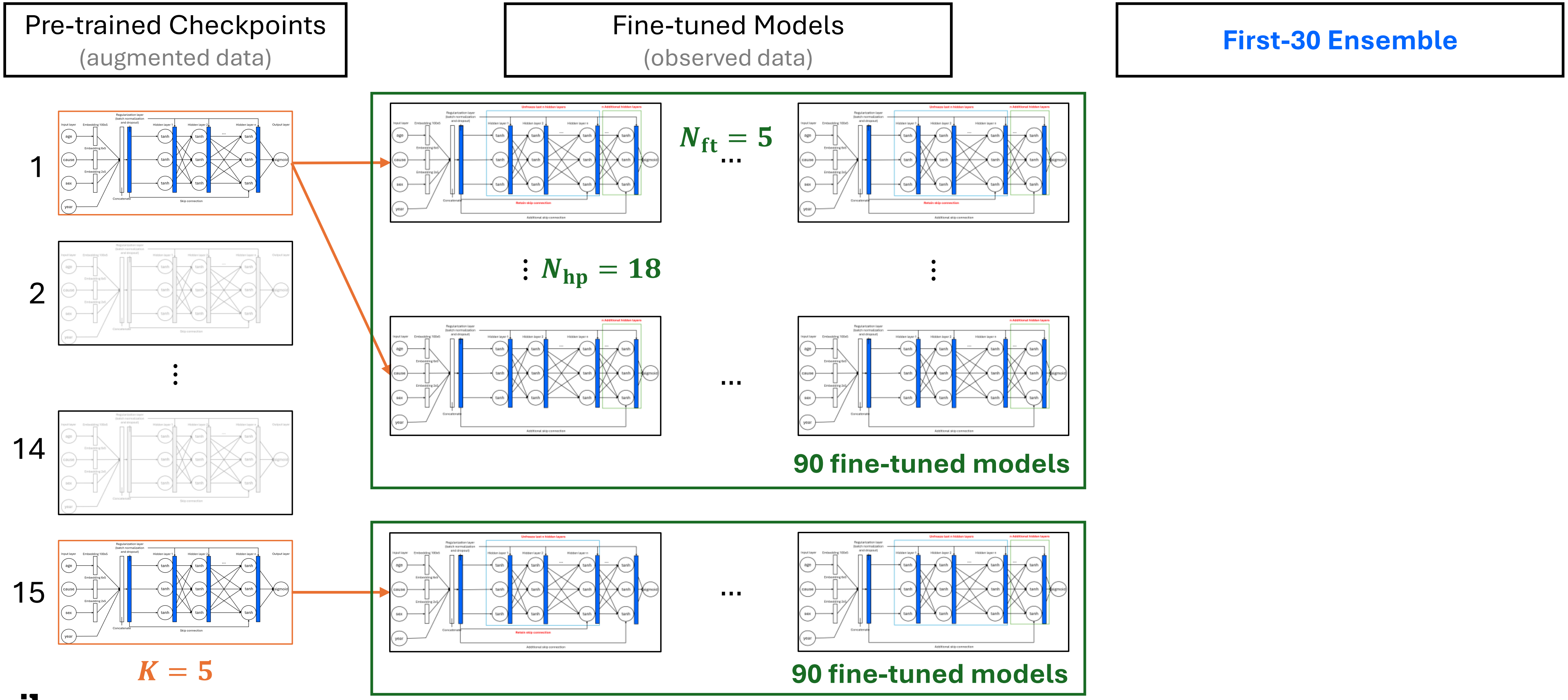
2


⋮

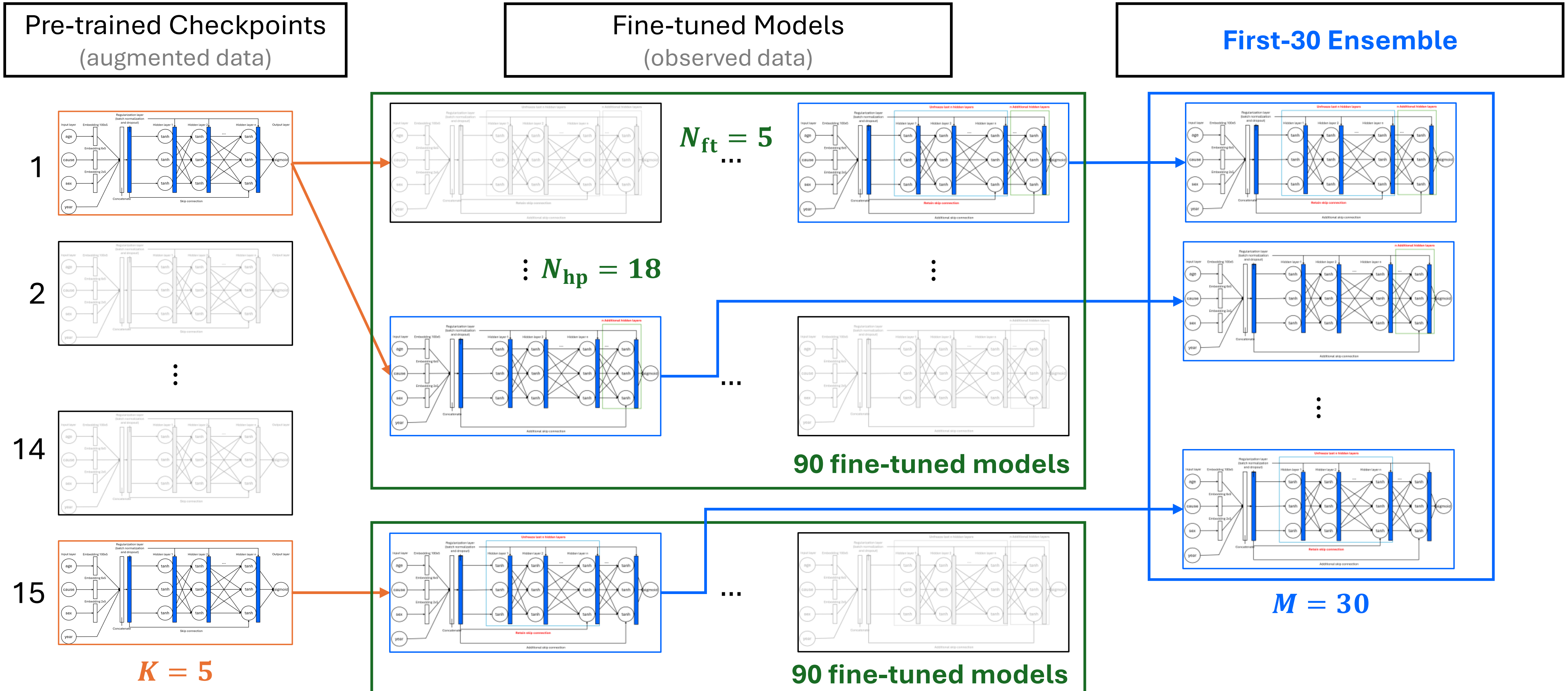
14


15


Methodology: Transfer Learning
(Robust Ensemble Strategy)



Methodology: Transfer Learning
(Robust Ensemble Strategy)



Findings: FFNN with Data Augmentation

Table 1. Performance of baseline FFNN and pre-trained with the proposed methodology (First-30 Ensemble) on the US cause-of-death test set (2000-2017), along with the three classical benchmarks.

Model	Mean cause-sex MSE (mortality x 10^4)	Median cause-sex MSE (mortality x 10^4)	Win counts vs LC|APC|RH (12 subpopulations)	All-cause MSE (mortality x 10^4)
FFNN	0.295	0.017	8 | 8 | 8	1.493
First-30 Ensemble	0.221	0.021	10 | 7 | 9	0.363
LC_MLE	0.187	0.116		0.429*
APC	0.282	0.089		2.283*
RH	0.618	0.095		0.477*

*For LC, APC, and RH, all-cause predictions are obtained by fitting a separate model directly to the aggregated all-cause mortality. For neural networks, it is obtained by summing the six cause-specific predicted mortality rates.

Findings: Network Understanding
(Cause Embedding Geometry)

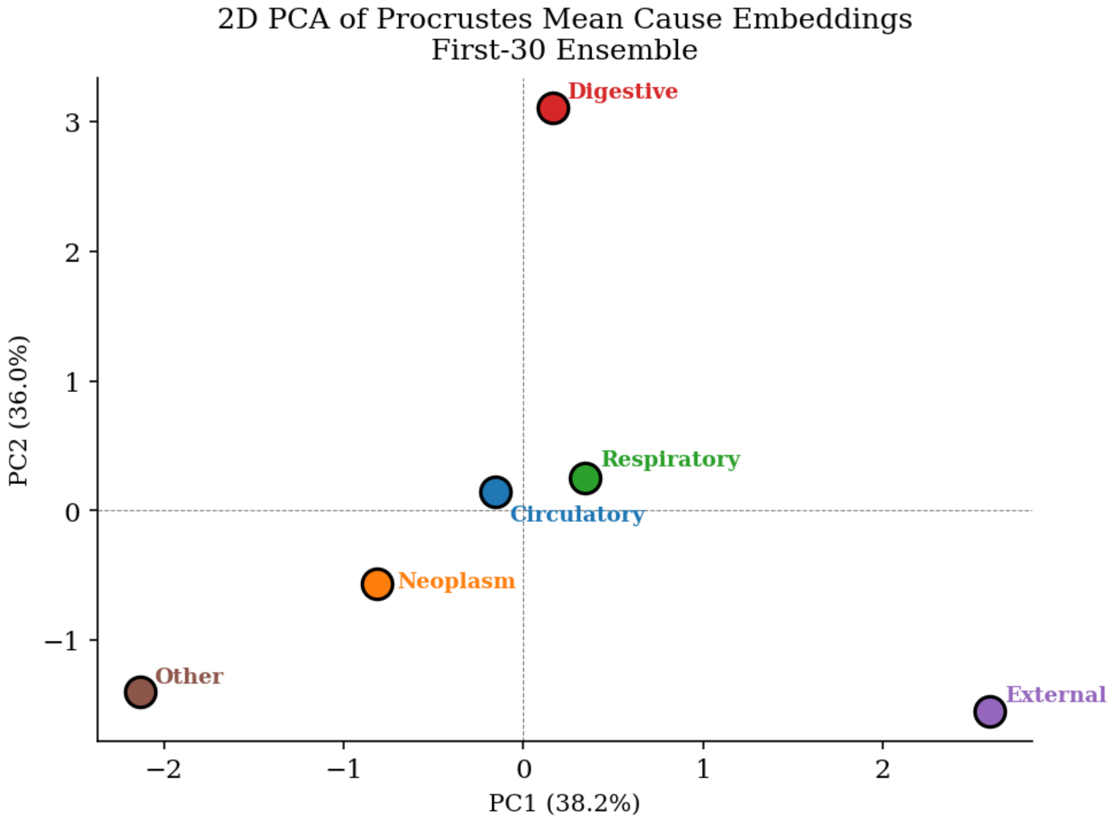


Figure 5. Two-dimensional PCA projection of the mean cause embeddings across members of the First-30 Ensemble, after Procrustes alignment.

Findings: Data Augmentation

(Selected Predicted Mortality Trend, Male 35 and 80)

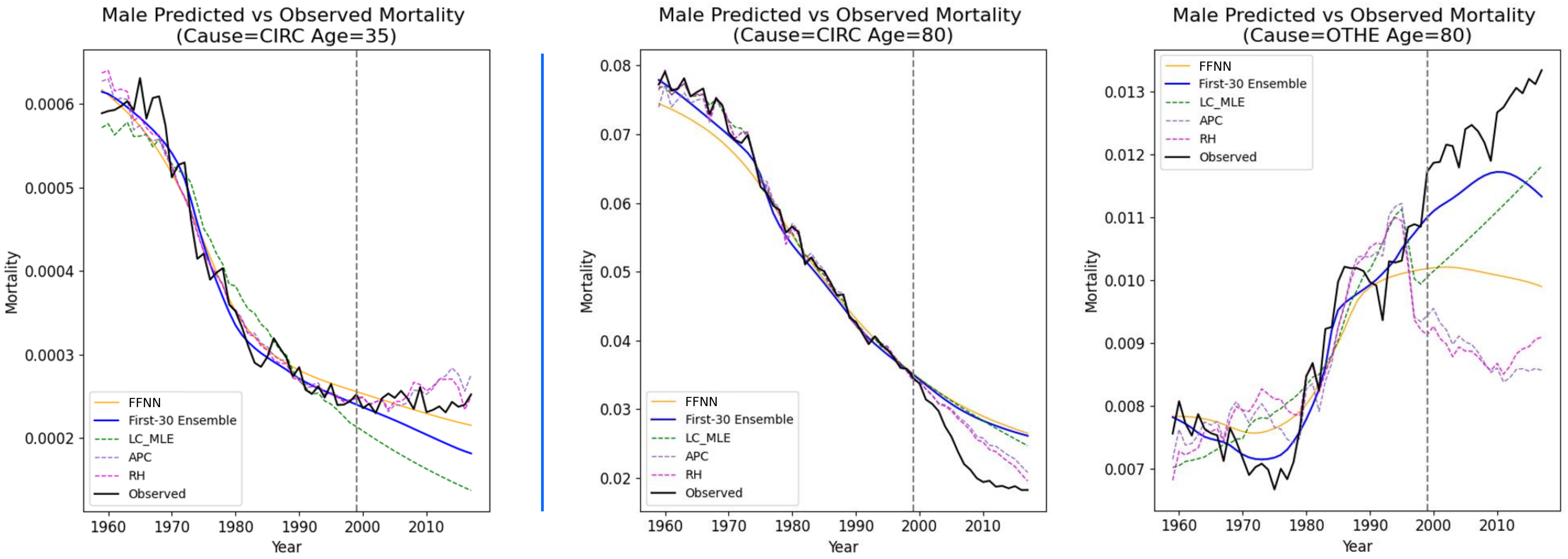
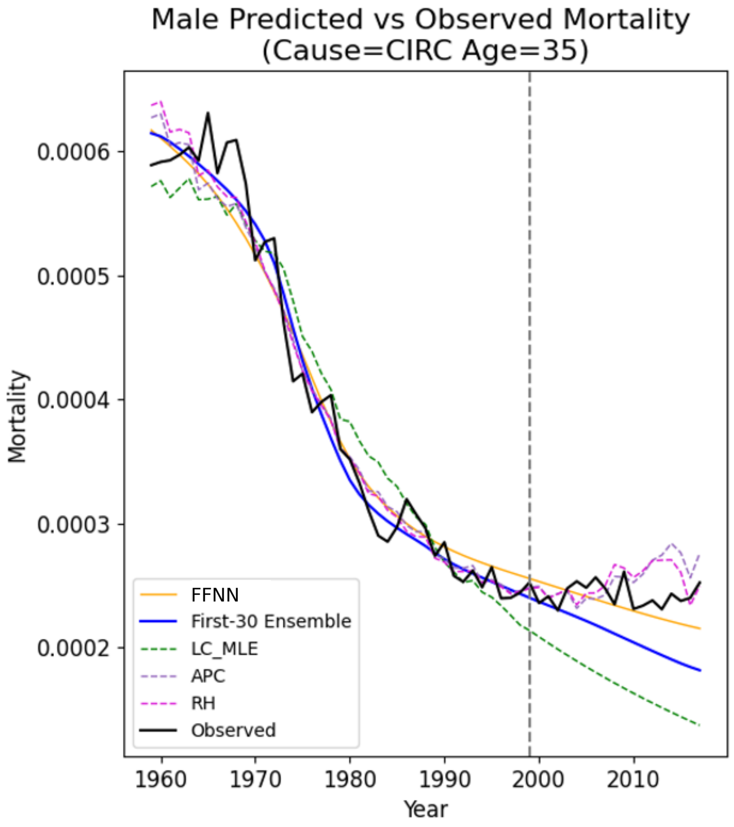
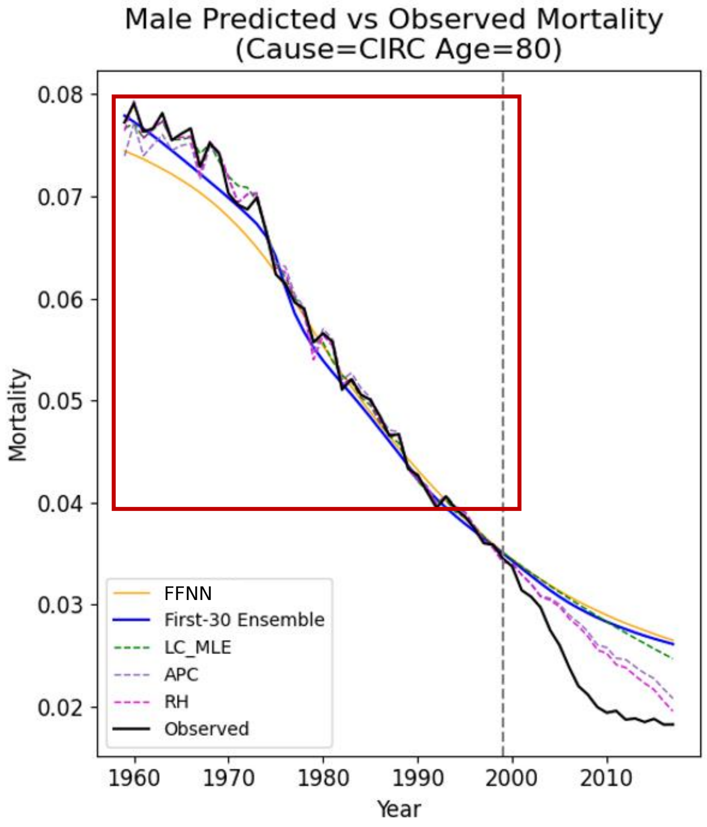


Figure 6. Observed (black) and predicted mortality male age 35 (circulatory) and 80 (circulatory and other) for the baseline FFNN (orange), the First-30 Ensemble (blue), and the three classical benchmarks (dashed lines) across the years for

Findings: Data Augmentation
(Selected Predicted Mortality Trend, Male 35 and 80)





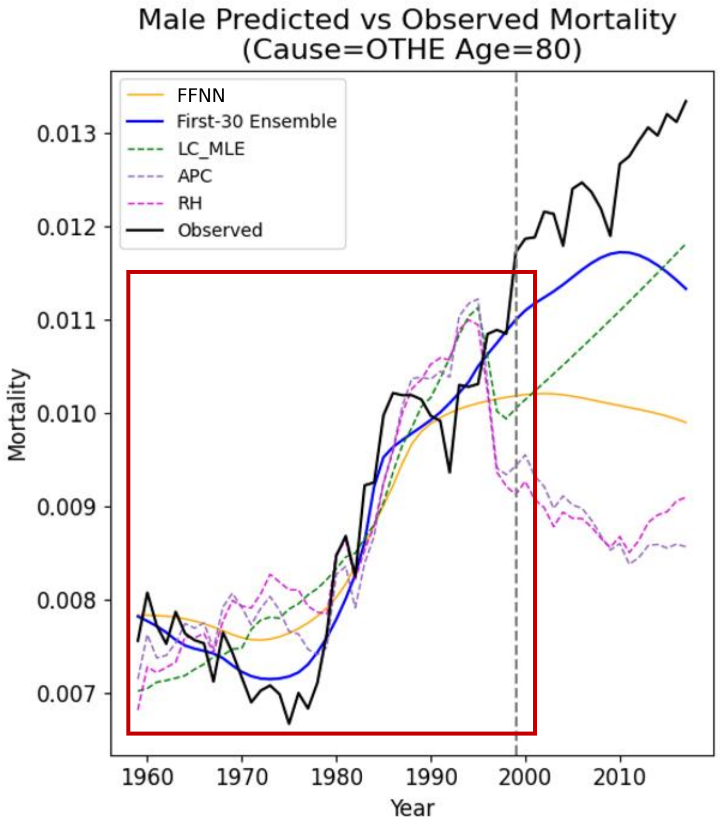


Figure 6. Observed (black) and predicted mortality male age 35 (circulatory) and 80 (circulatory and other) for the baseline FFNN (orange), the First-30 Ensemble (blue), and the three classical benchmarks (dashed lines) across the years for

Findings: Data Augmentation
(Selected Predicted Mortality Trend, Male 35 and 80)

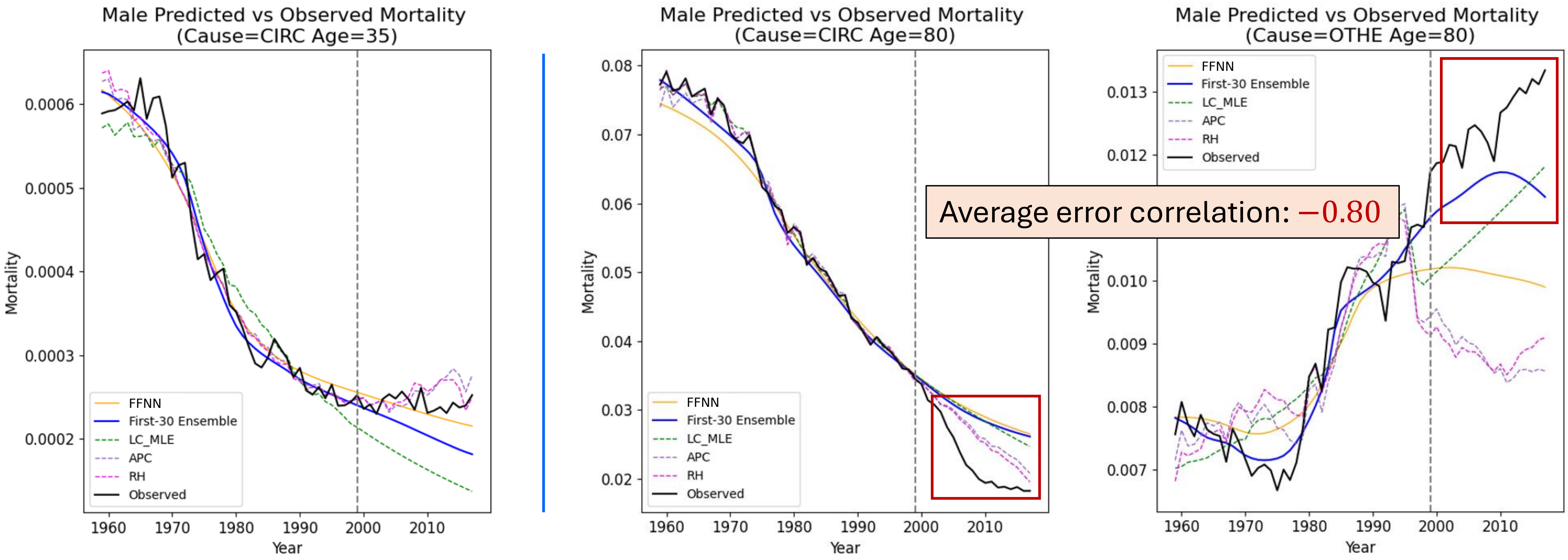


Figure 6. Observed (black) and predicted mortality male age 35 (circulatory) and 80 (circulatory and other) for the baseline FFNN (orange), the First-30 Ensemble (blue), and the three classical benchmarks (dashed lines) across the years for

Findings: Data Augmentation
(Selected Predicted Mortality Trend, Male 35 and 80)

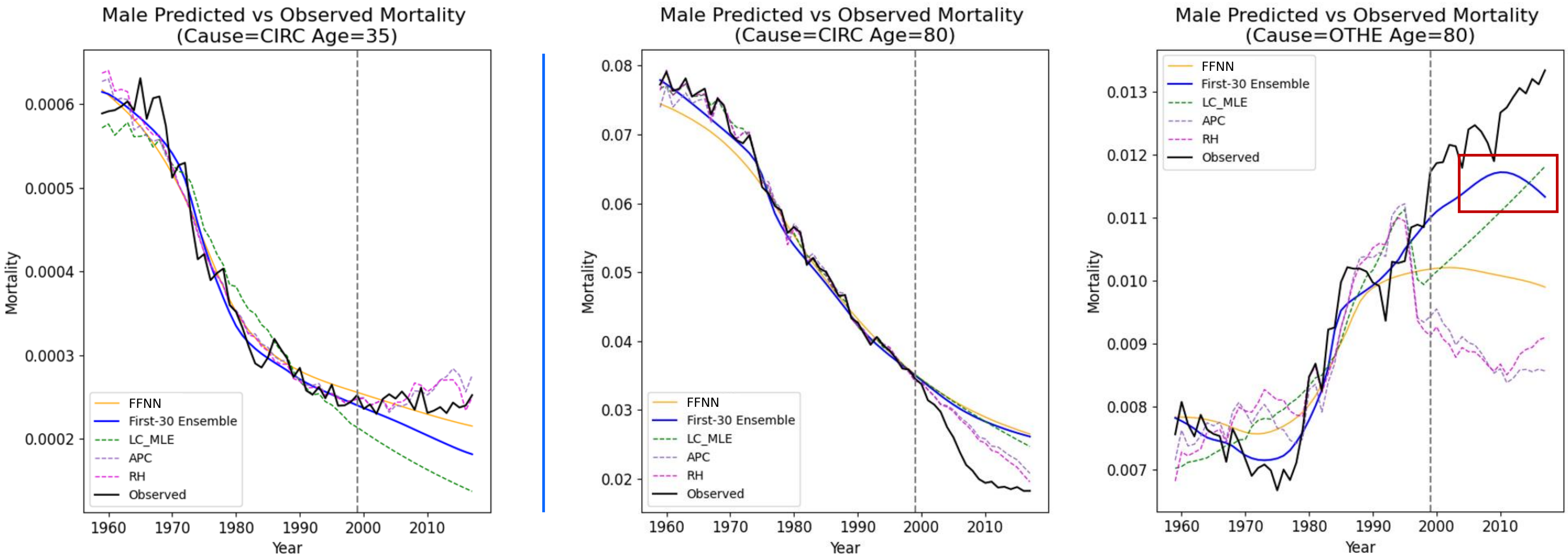


Figure 6. Observed (black) and predicted mortality male age 35 (circulatory) and 80 (circulatory and other) for the baseline FFNN (orange), the First-30 Ensemble (blue), and the three classical benchmarks (dashed lines) across the years for

Findings: Data Augmentation
(Selected Predicted Mortality Trend, Male 35 and 80)

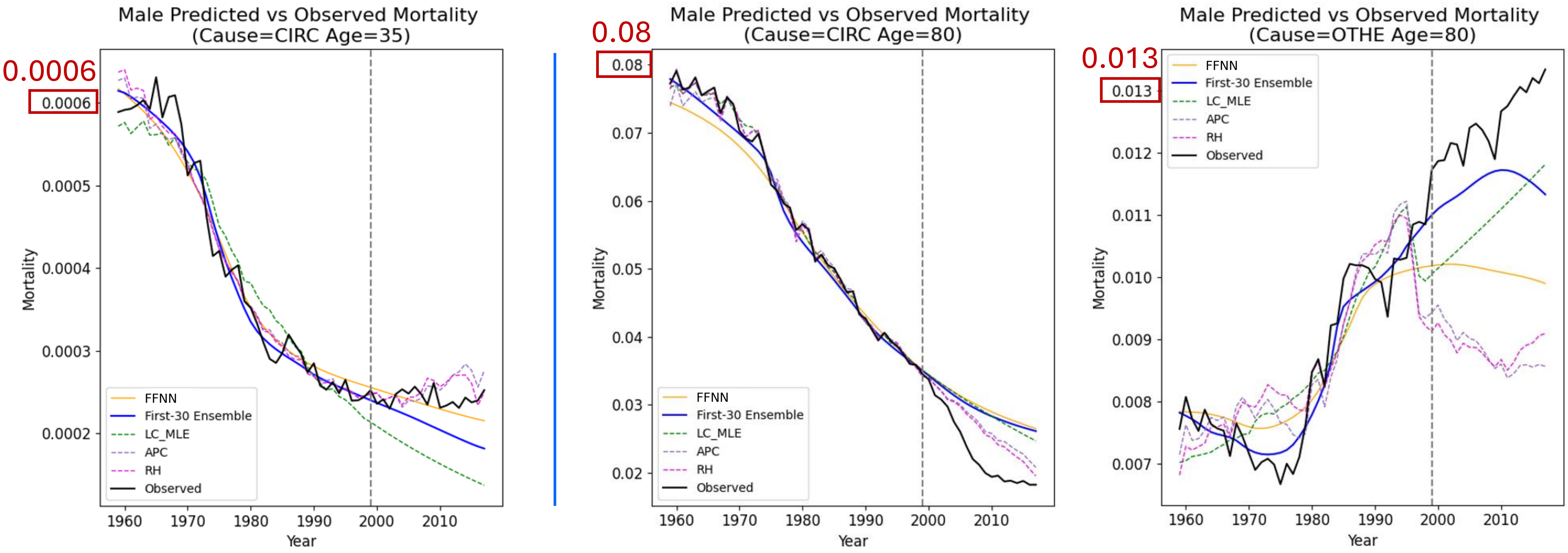


Figure 6. Observed (black) and predicted mortality male age 35 (circulatory) and 80 (circulatory and other) for the baseline FFNN (orange), the First-30 Ensemble (blue), and the three classical benchmarks (dashed lines) across the years for

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

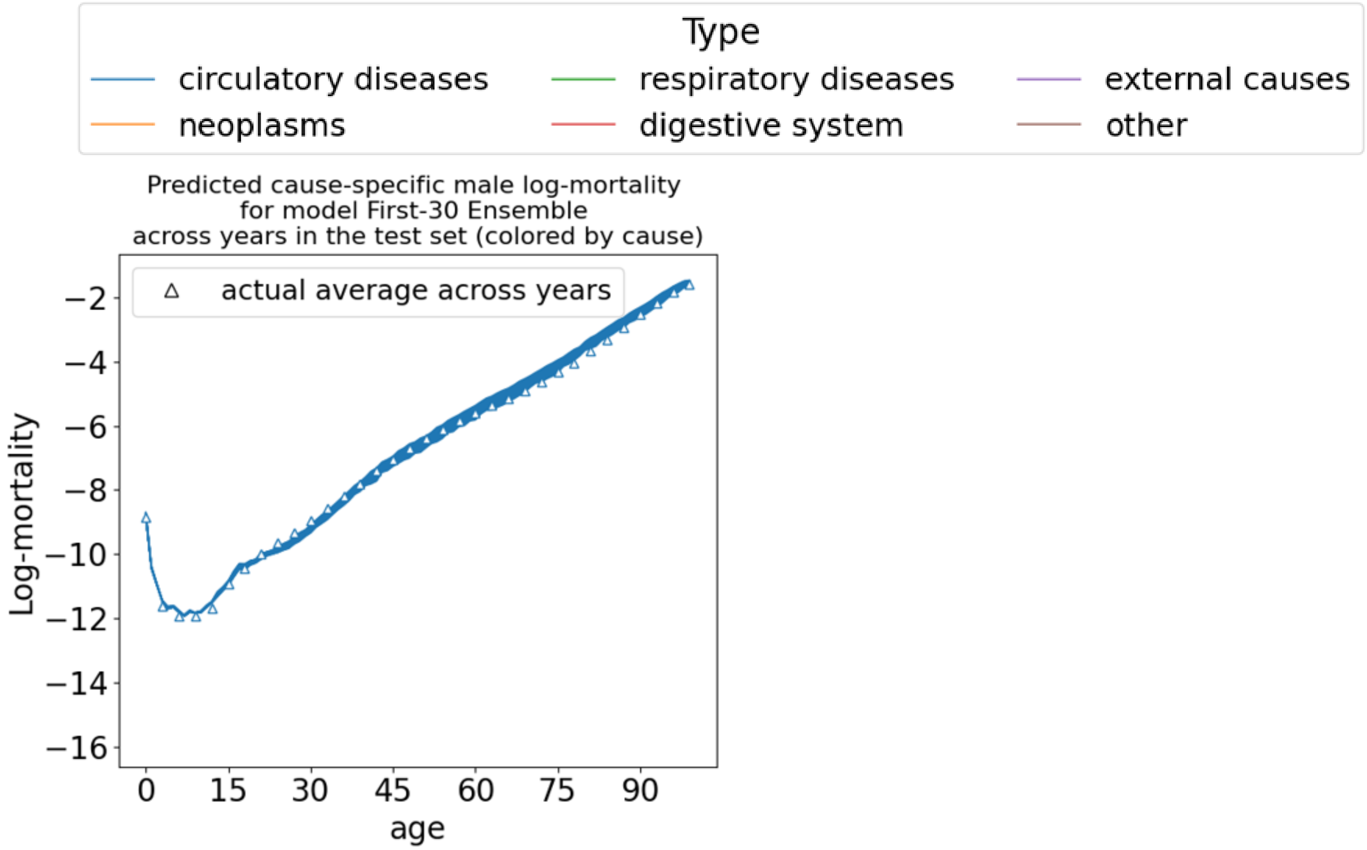


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

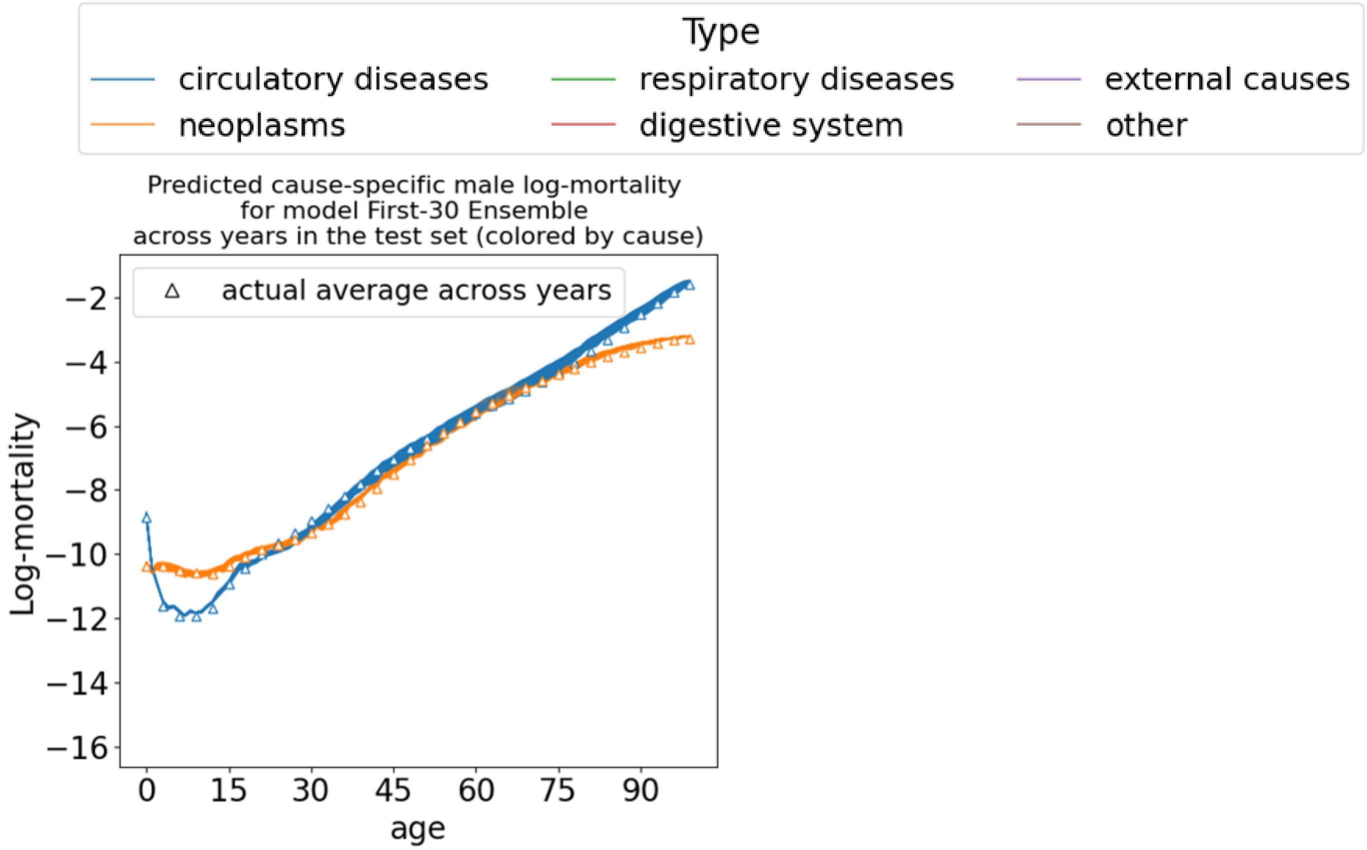


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

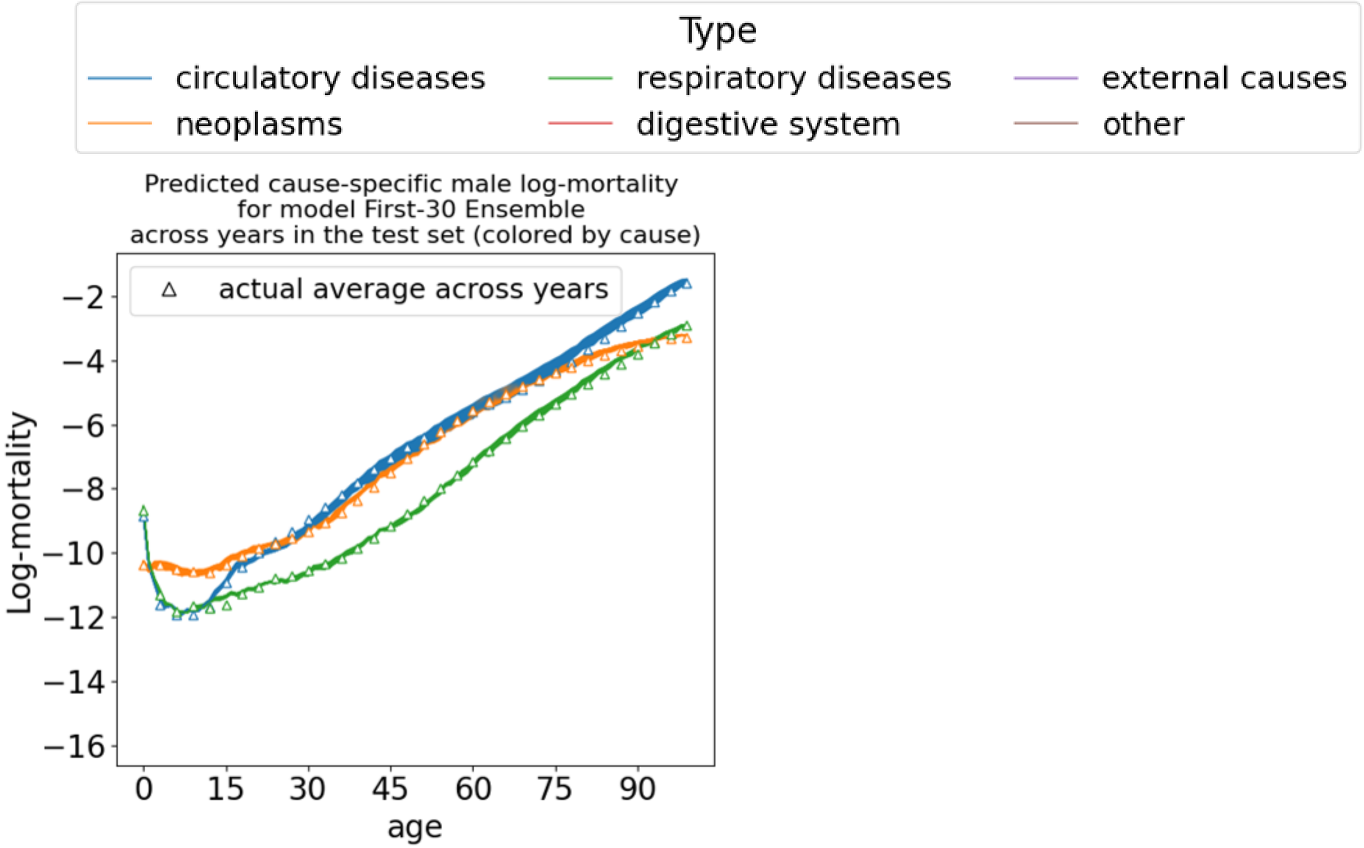


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

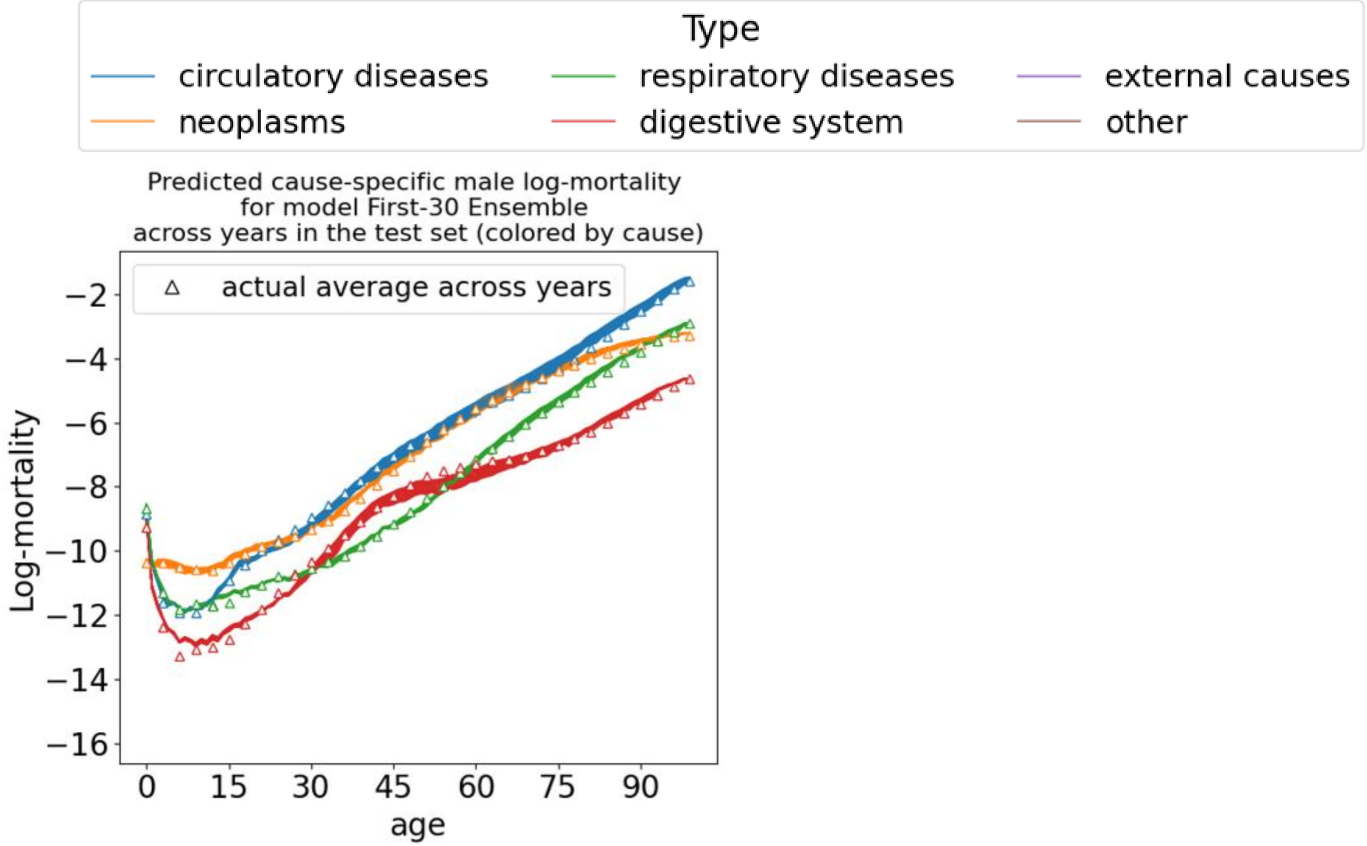


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

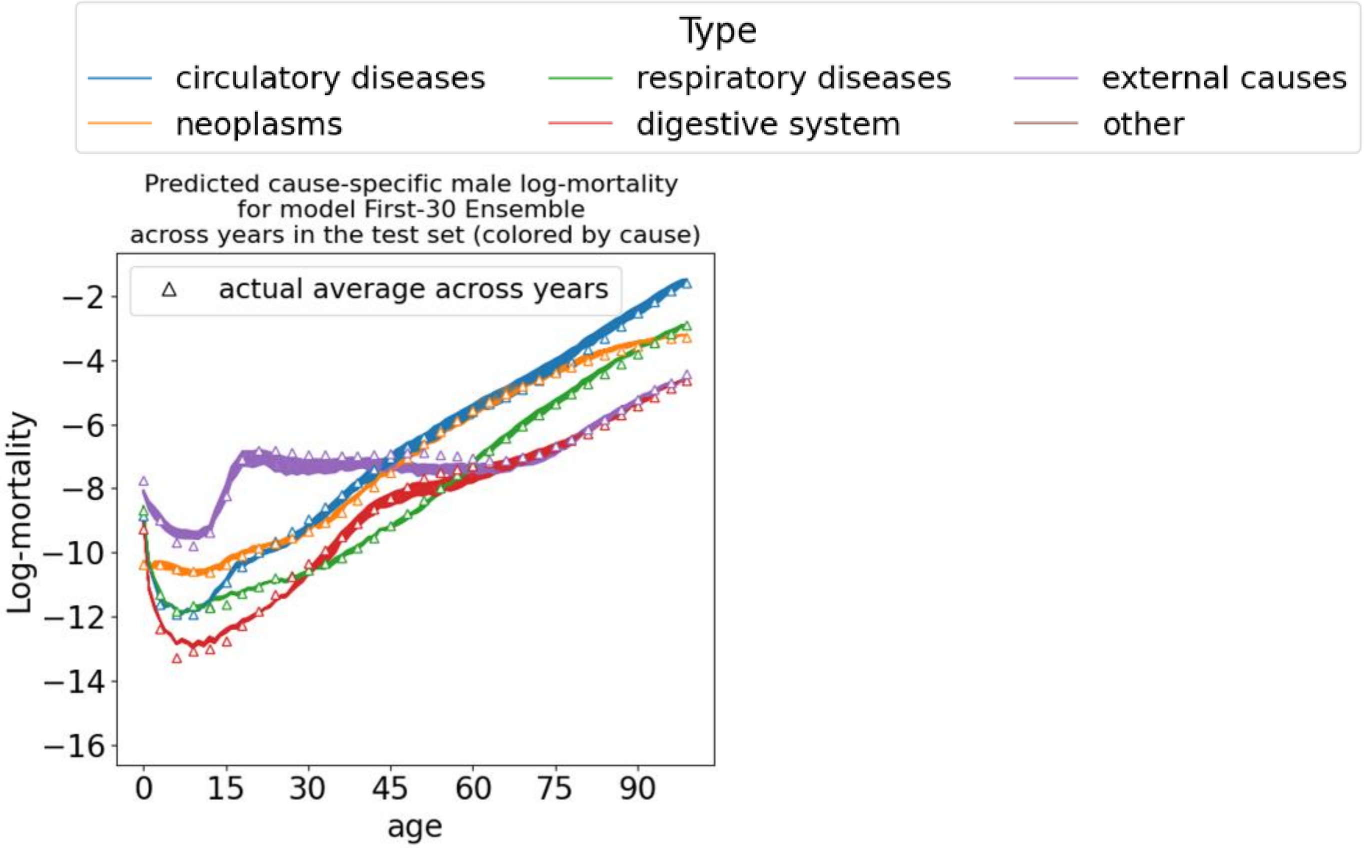


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

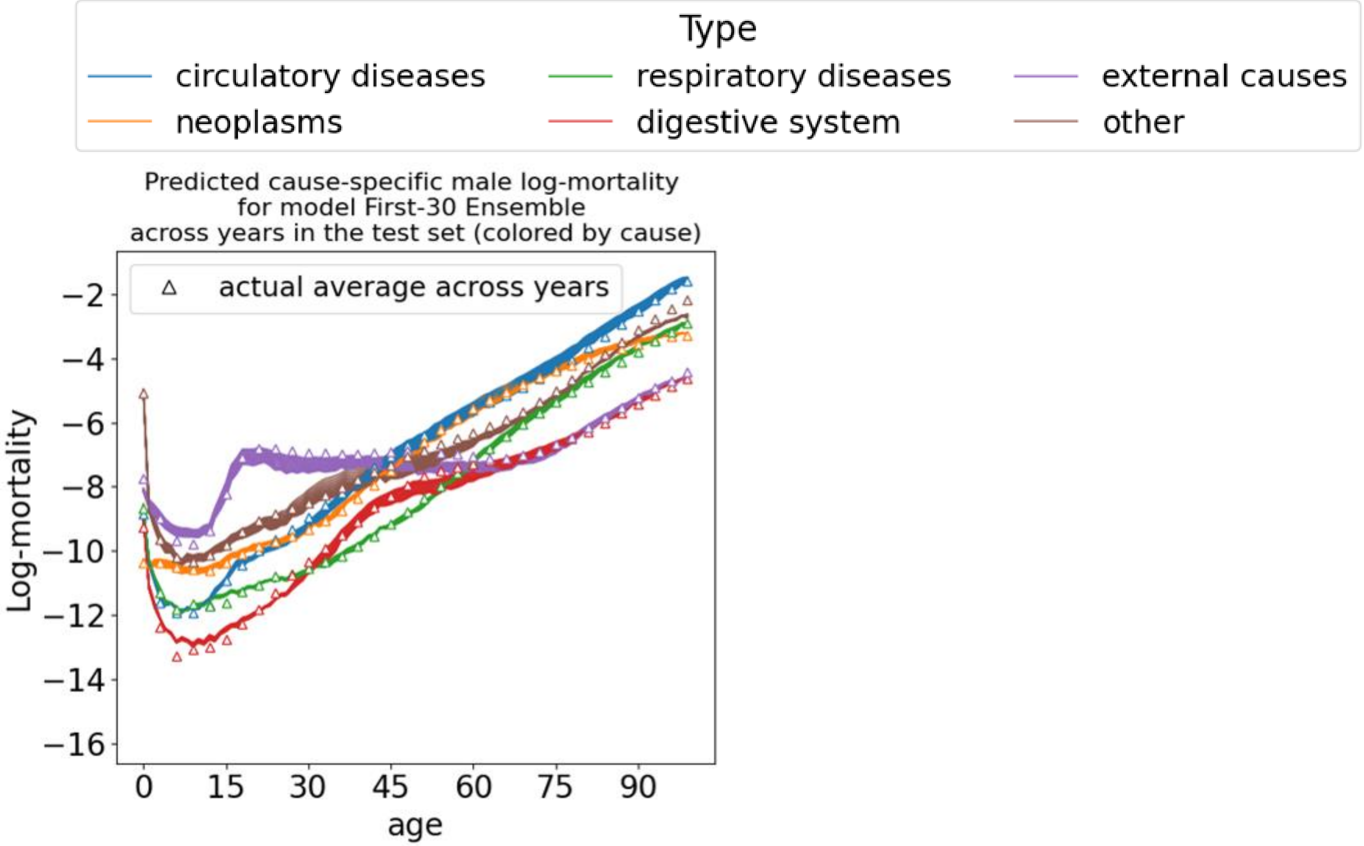


Figure 7. Predicted male cause-specific log-mortality over the test period (2000-2017). The close alignment between the predicted curves (whose width reflects year-to-year variation) and observed cause-age averages (open triangles) demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding
(Predicted Log-Mortality Coloured by Cause)

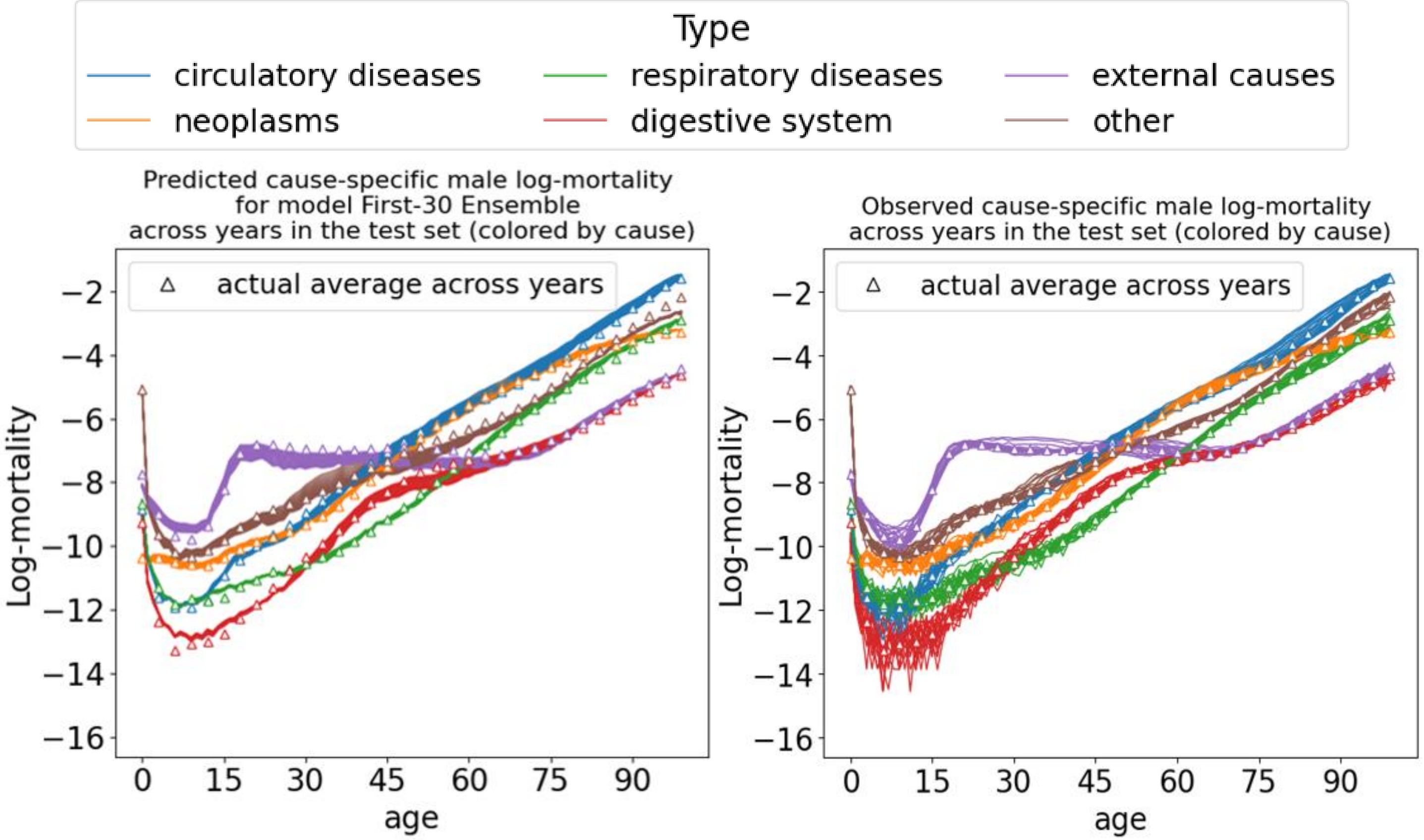


Figure 8. Predicted (left) and observed (right) male cause-specific log-mortality over the test period (2000-2017). The close alignment demonstrates **the network's ability to successfully capture the distinct age profile of each cause**.

Findings: Network Understanding

(Predicted Log-Mortality Coloured by Cause)

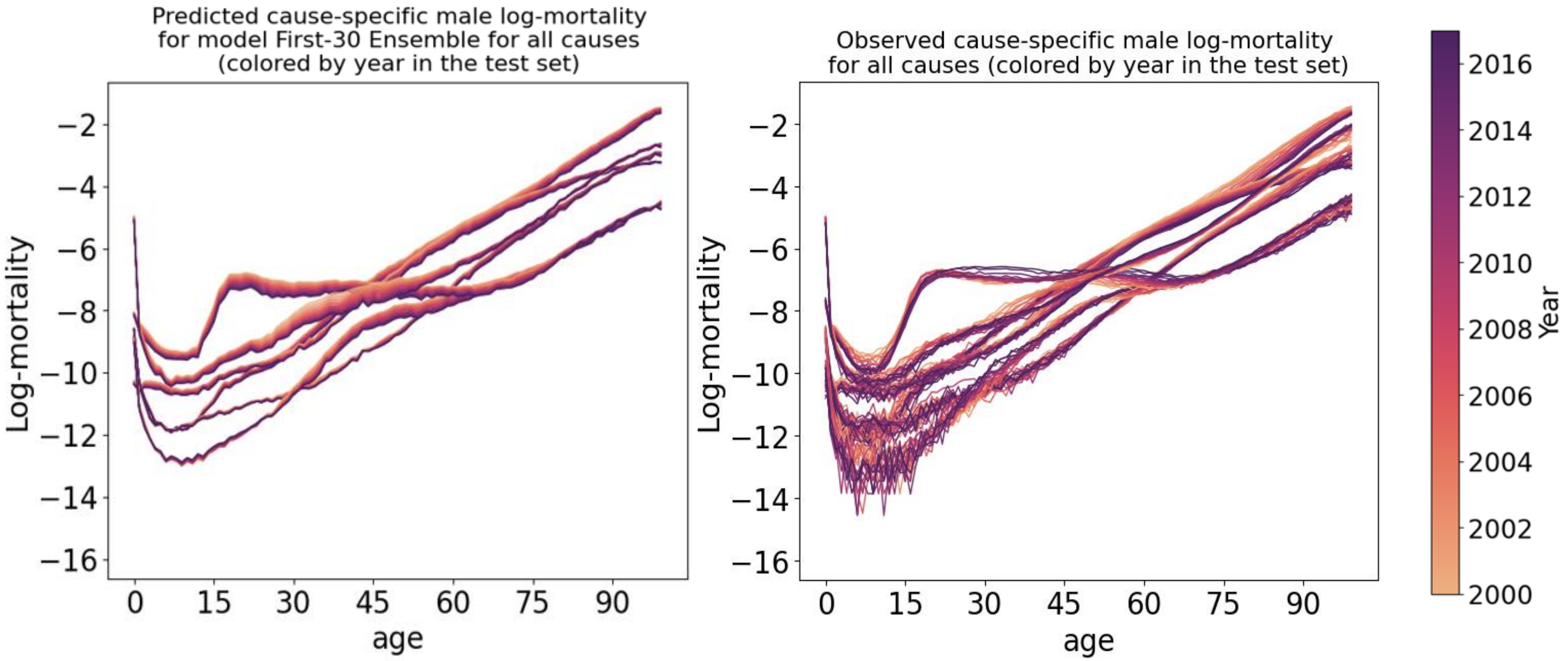


Figure 9. Predicted (left) and observed (right) male log-mortality over the test period (2000-2017). Different colors represent different years.

Conclusions & Future Directions

- **Core Takeaway**: Shared-parameter FFNNs + data augmentation effectively model cause-of-death rates.
  - **Outperforms** classical benchmarks (LC, APC, RH) in cause-specific forecasting.
  - **Achieves** superior all-cause performance implicitly without explicit training on aggregate data.
  - **Captures** age-cause interactions.
  - **Learns** meaningful cause embeddings aligned with known epidemiological groups.

- **Future Directions**
  - **Architecture**: Evaluate sequence models, cause elimination, and reconciliation tools.
  - **Cause Grouping**: Leverage embeddings to cluster causes for downstream models.
  - **Data**: Extend to multi-population or contributory cause-of-death datasets.

Bibliography

Alai, D. H., Gaille, S. A., & Sherris, M. (2014). Modelling cause-of-death mortality and the impact of cause-elimination. Annals of Actuarial Science, 9(1), 167–186. https://doi.org/10.1017/S174849951400027X

Arnold, S., & Glushko, V. (2022). Short- and long-term dynamics of cause-specific mortality rates using cointegration analysis.North American Actuarial Journal, 26, 161–183. https://doi.org/10.1080/10920277.2021.1874421

Dong, Z. M., Shang, H. L., Hui, F., & Bruhn, A. (2025). A compositional approach to modeling cause-specific mortality withzero counts. Annals of Actuarial Science, 19(3), 416–441. https://doi.org/10.1017/S1748499524000320

Euthum, M., Scherer, M., & Ungolo, F. (2024). A neural network approach for the mortality analysis of multiple populations:A case study on data of the Italian population. European Actuarial Journal, 1–30. https://doi.org/10.1007/S13385-024-00377-5

Graziani, R., & Nigri, A. (2026). An age–period–cohort model in a Dirichlet framework: A coherent causes of death estimation.Insurance: Mathematics and Economics, 126, 103194. https://doi.org/https://doi.org/10.1016/j.insmatheco.2025.103194

Mori, L. D., Haberman, S., Millossovich, P., & Zhu, R. (2025). Mortality forecasting via multi-task neural networks. ASTINBulletin: The Journal of the IAA, 1–19. https://doi.org/10.1017/ASB.2025.10

Bibliography

Richman, R., & Wüthrich, M. V. (2021). A neural network extension of the Lee-Carter model to multiple populations. Annals of Actuarial Science, 15(2), 346–366. https://doi.org/10.1017/S1748499519000071

Tanaka, S., & Matsuyama, N. (2025). An interpretable neural network approach to cause-of-death mortality forecasting.Annals of Actuarial Science, 1–20. https://doi.org/10.1017/S1748499524000319

Varga, L. (2025). Forecasting cause-of-death mortality with single- and multi-population models in Hungary. ScandinavianActuarial Journal, 0(0), 1–32. https://doi.org/10.1080/03461238.2025.2480846

Villegas, A. M., Bajekal, M., Haberman, S., & Zhou, L. (2021). Analysis of Historical U. S. Population Mortality Improvement Drivers 1959-2016 | SOA. https://www.soa.org/resources/research-reports/2021/analysis-historical-us-drivers/

Wang, C. W., Zhang, J., & Zhu, W. (2021). Neighbouring prediction for mortality. ASTIN Bulletin, 51(3), 689–718. https://doi.org/10.1017/asb.2021.13

Perla, F., Richman, R., Scognamiglio, S., & Wüthrich, M. V. (2021). Time-series forecasting of mortality rates using deep learning. Scandinavian Actuarial Journal, 2021(7), 572–598. https://doi.org/10.1080/03461238.2020.1867232

Thank you!