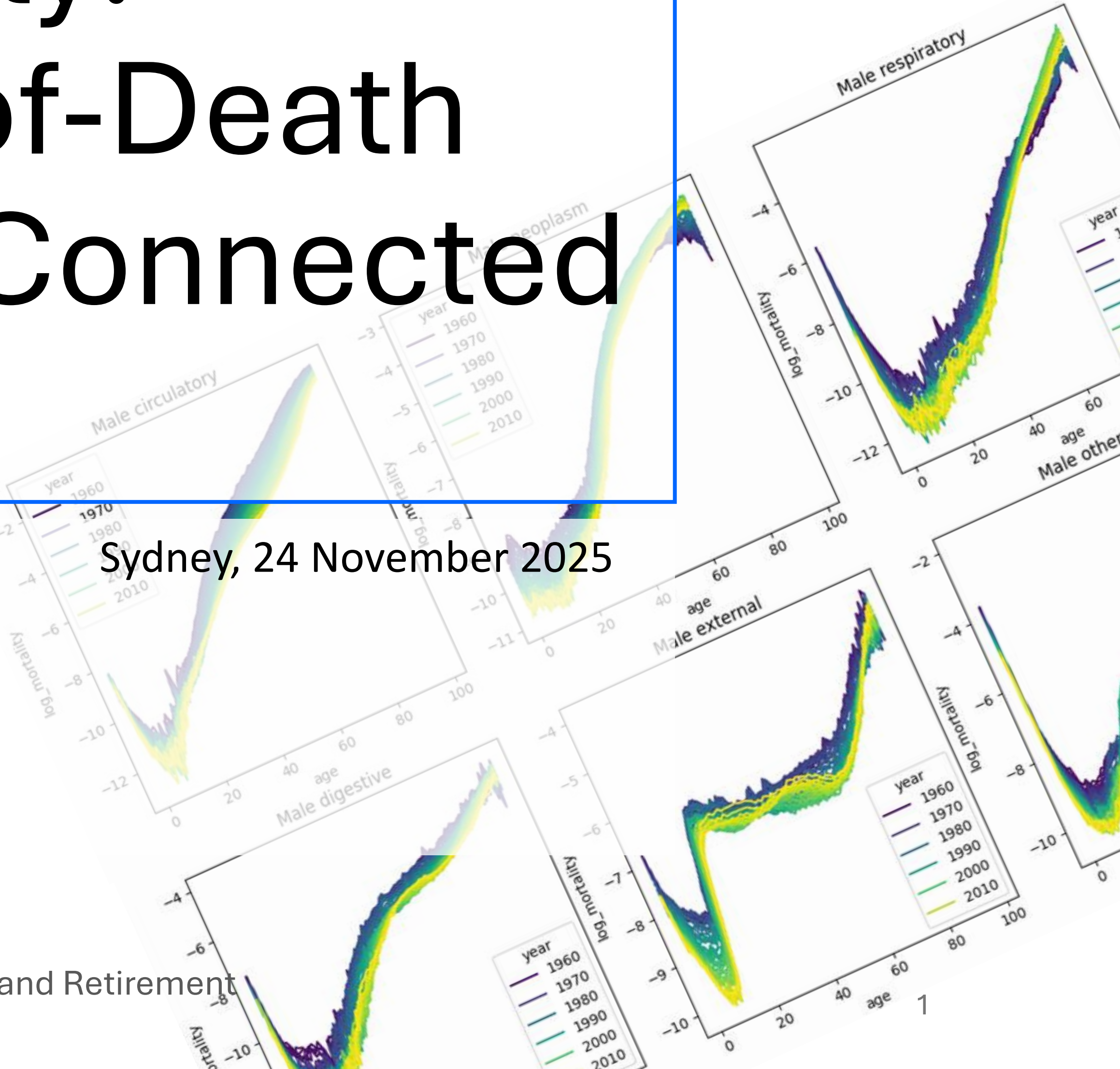
Forecasting Mortality: Leveraging Cause-of-Death Data through Fully Connected Neural Network

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

Sydney, 24 November 2025



Unil.

Agenda

- Background and Motivation
- Literature Reviews
- Cause-of-Death Datasets
- Methodology and Findings
- Final Remarks

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?

Literature Reviews

- Approaches to modelling cause-of-death mortality rates:
  - Multinomial logistic regression (Alai et al., 2014)
  - Independent LC model (Boumezoued et al., 2019)
  - Independent APC model (Villegas et al., 2021)
  - (Nested) Li-Lee model (Lyu et al., 2021)
- Approaches to modelling mortality data with NN:
  - NN to extract features to be fed to classical time series approach (Richman & Wüthrich, 2021, 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; J. Wang et al., 2023; Perla et al., 2024; Euthum et al., 2024; De Mori et al., 2025)

Literature Reviews

- Approaches to modelling cause-of-death mortality rates:
  - Multinomial logistic regression (Alai et al., 2014)
  - Independent LC model (Boumezoued et al., 2019)
  - Independent APC model (Villegas et al., 2021)
  - (Nested) Li-Lee model (Lyu et al., 2021)
- Approaches to modelling mortality data with NN:
  - NN to extract features to be fed to classical time series approach (Richman & Wüthrich, 2021, 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; J. Wang et al., 2023; Perla et al., 2024; Euthum et al., 2024; De 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 and Findings

- Cause-of-Death modelling
  - Which network architecture should be used to model cause-of-death log-mortality rate?
- Transfer learning (pretraining) based on simulated data
  - How to optimize (or improve) the FCNN performance in modelling cause-of-death log-mortality rates?
- Network attributions
  - What did the network learn?

Methodology: Cause-of-Death Modelling

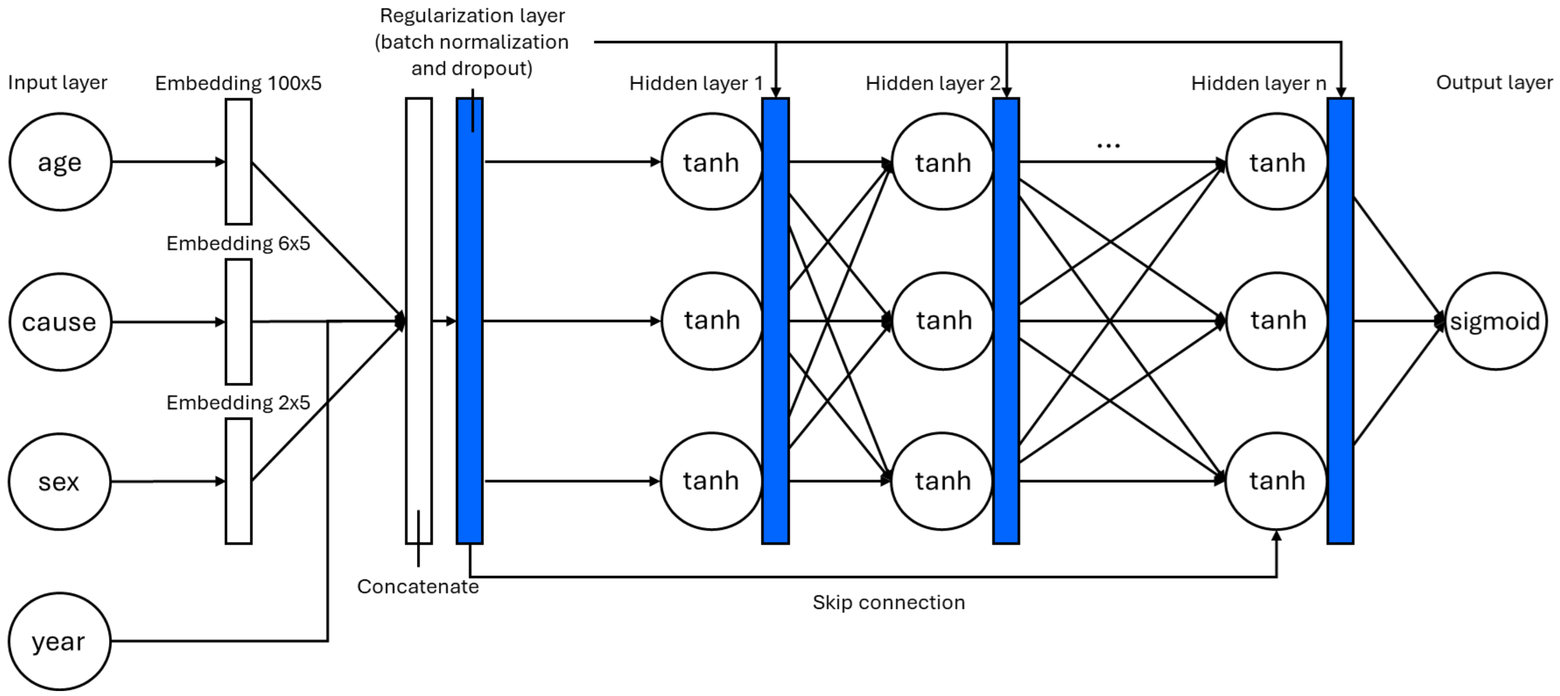


Figure 2. FFNN architecture

Findings: Cause-of-Death Modelling

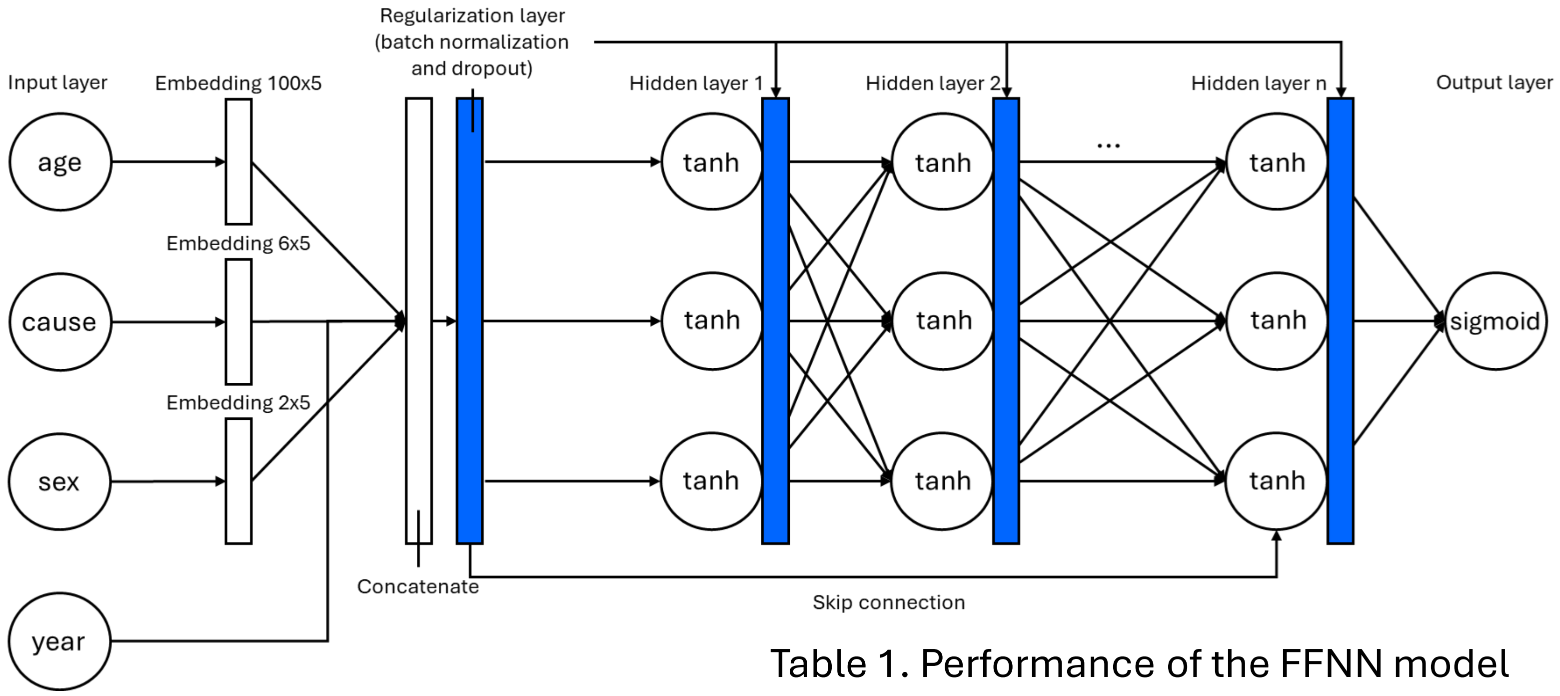


Figure 2. FFNN architecture

Table 1. Performance of the FFNN model

Model	Mean MSE of mortality rates (10^4)	Median MSE of mortality rates (10^4)	Wins over LC/APC/RH (12 subpopulations)
FFNN	0.44	0.013	9 | 7 | 7

Methodology: Transfer Learning
(Simulation)

- One rate for each year, gender, age and cause (49,200 observations)
  - For each cause-gender: 4,100 samples (100 ages and 41 years)
- Simulate the number of deaths:
  - Take the last training year.
  - Calculate cause-specific death probability for **each age and gender**.
  - Add normal random perturbation:
    - with mean 0, and
    - variance from the last 20 years.
  - Repeat for a lot of time-steps (5,000 time steps).

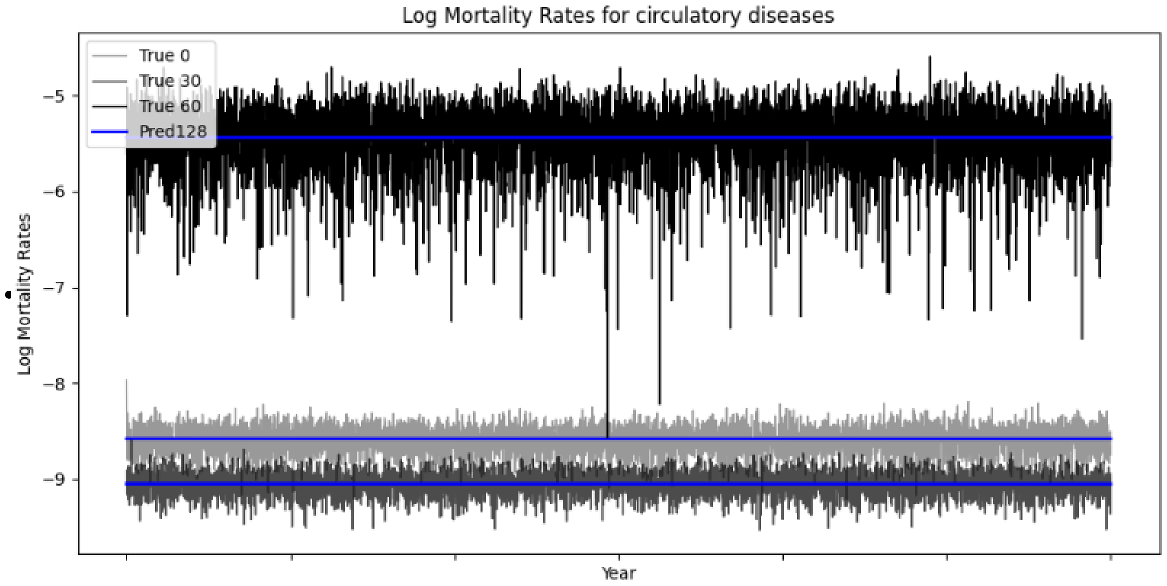


Figure 3. Model FFNN_ckpt predictions on simulated dataset

Methodology: Transfer Learning
(Hyperparameters)

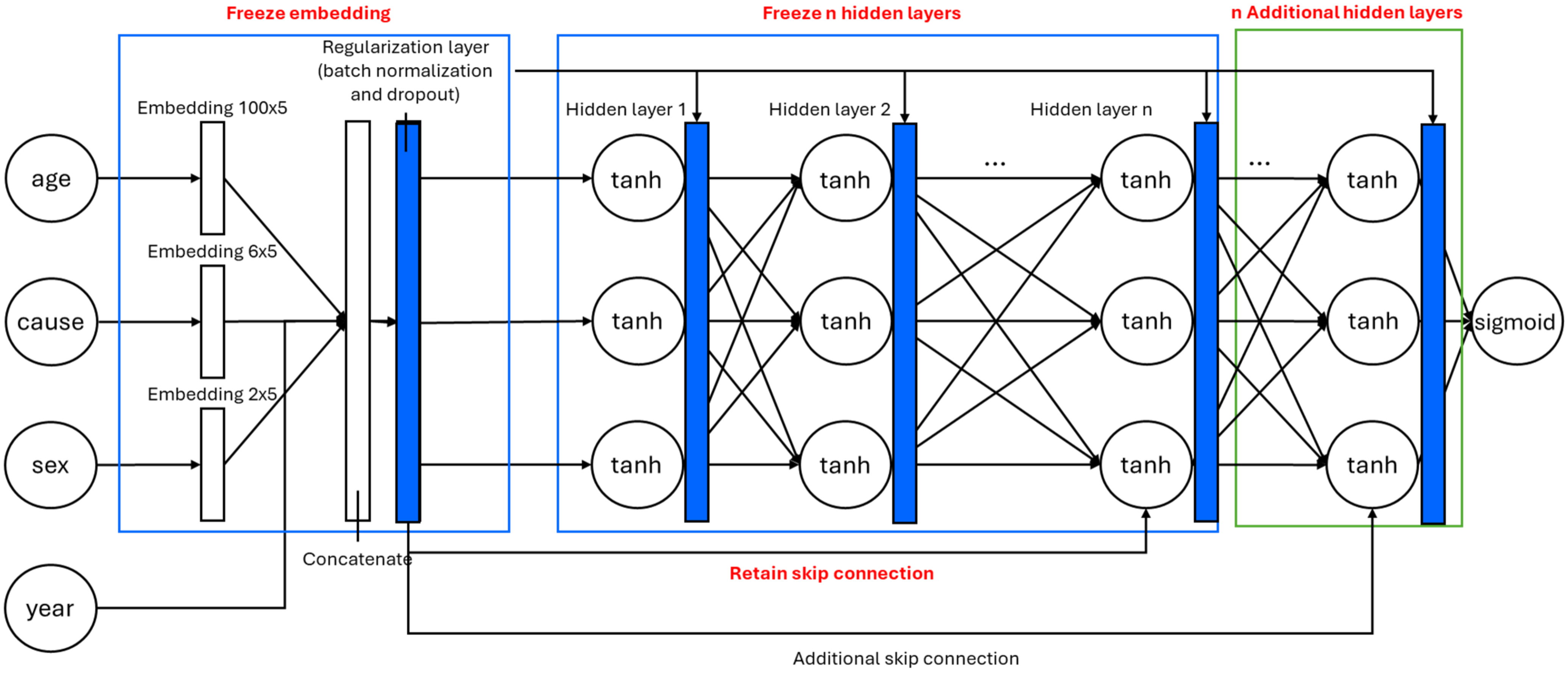


Figure 4. Pretrained FFNN architecture

Findings: Transfer Learning

(Transfer Learning Based on Simulated Data)

Table 2. Performance of the pretrained network and ones without pretraining

Model	Mean MSE of mortality rates (10^4)	Median MSE of mortality rates (10^4)	Win over LC/APC/RH (12 Subpopulations)
FFNN	0.44	0.013	9 | 7 | 7
FFNN with 400 epochs	0.44	0.011	9 | 8 | 7
FFNN_0002	0.33	0.017	10 | 8 | 8
FFNN_0002 with 200 epochs (without pretraining)	0.49	0.013	8 | 7 | 7
FFNN_0002 with 400 epochs (without pretraining)	0.41	0.014	9 | 8 | 7

Findings: Transfer Learning
(Log-Mortality Curves Male 22 and 65)

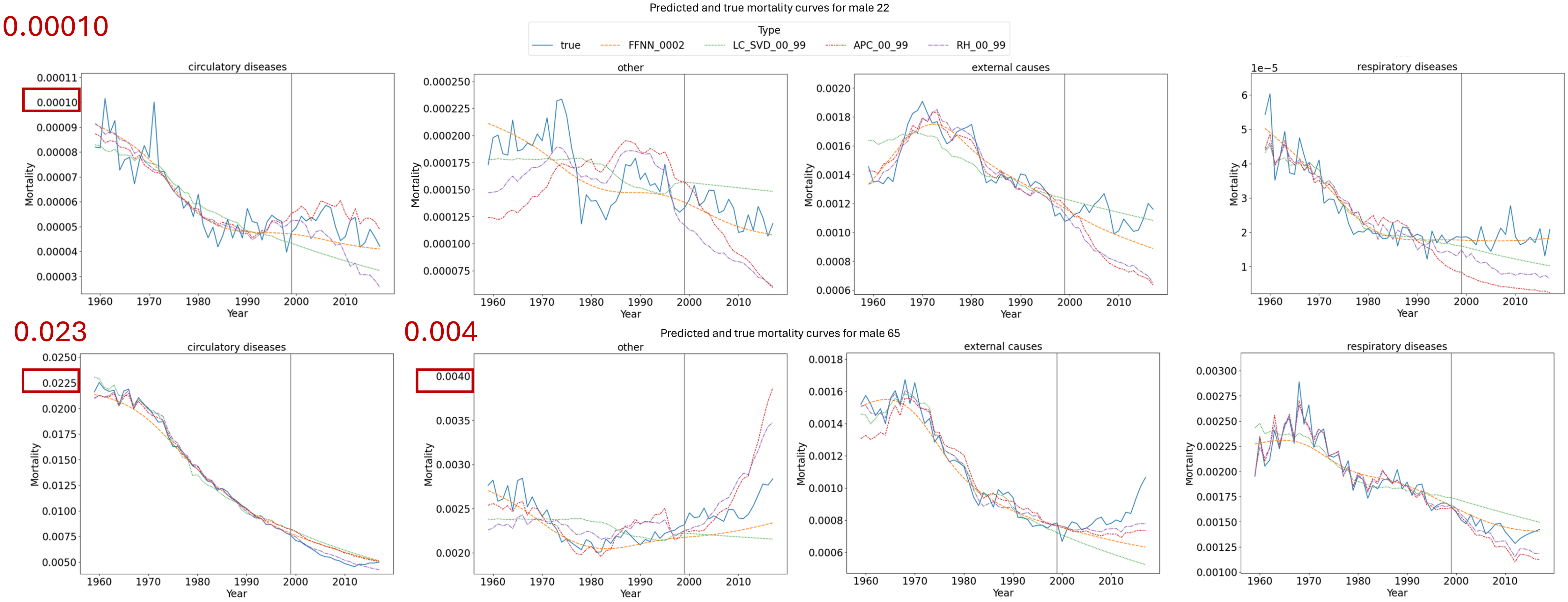


Figure 5. Predicted and true log-mortality curves across the year for male age 22 and 65

Findings: Transfer Learning
(Selected Log-Mortality Curves Male 22 and 65)

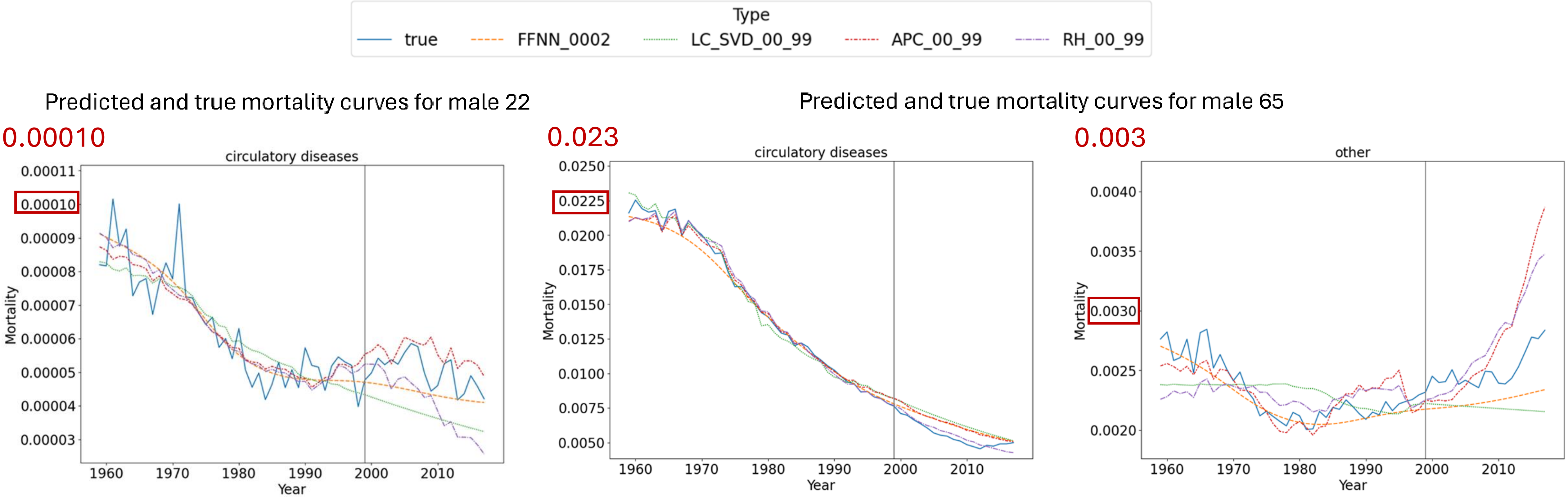
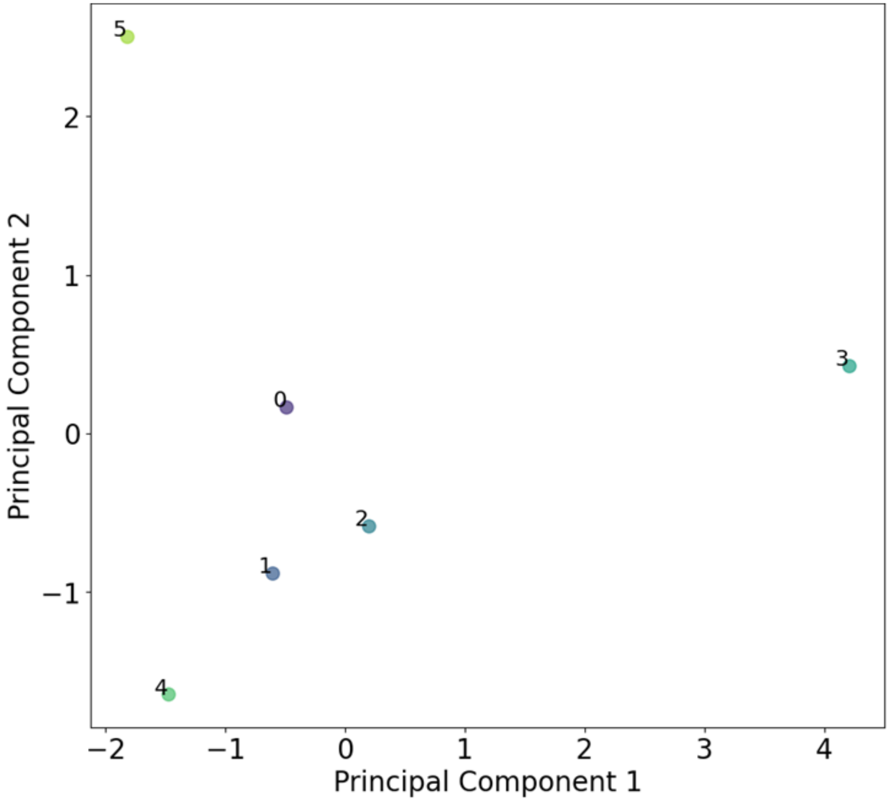


Figure 6. Predicted and true log-mortality curves across the year for male age 22 (circulatory) and 65 (circulatory and other)

Findings: Network Interpretations
(PCA on Embedding Weights)

PCA of embedding weights for the cause feature
from a single FFNN_0002 model with the lowest MSE
mean MSE: 0.29 median MSE: 0.03



PCA of embedding weights for the cause feature
from a single FFNN_0002 model with the highest MSE
mean MSE: 0.82 median MSE: 0.03

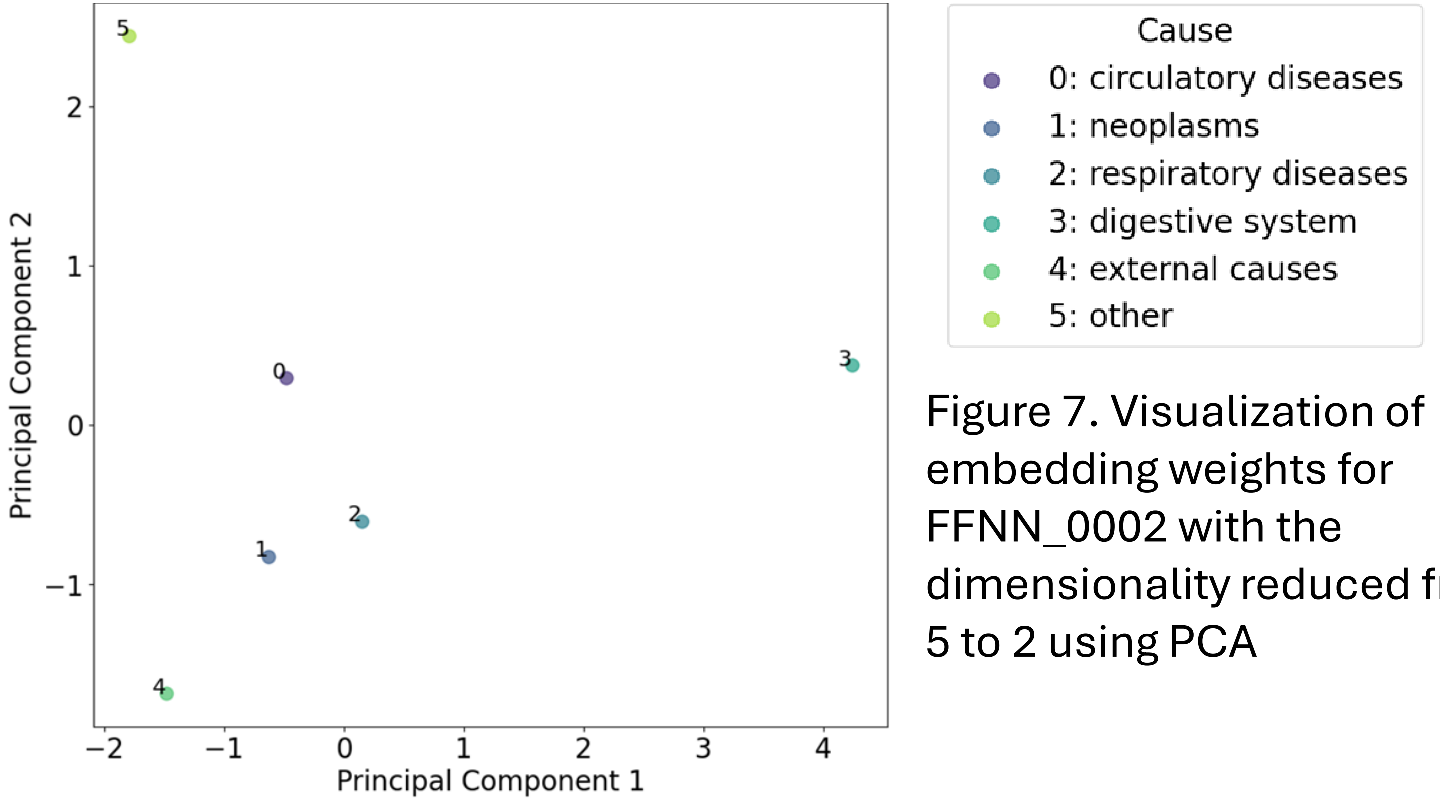


Figure 7. Visualization of embedding weights for FFNN_0002 with the dimensionality reduced from 5 to 2 using PCA

Findings: Network Interpretations
(Feature Importance)

Permutation Feature Importance averages the increase in error when permutating each specific feature.

SHAP Value averages how much a particular feature change the prediction from the 'expected' prediction to each instance's prediction.

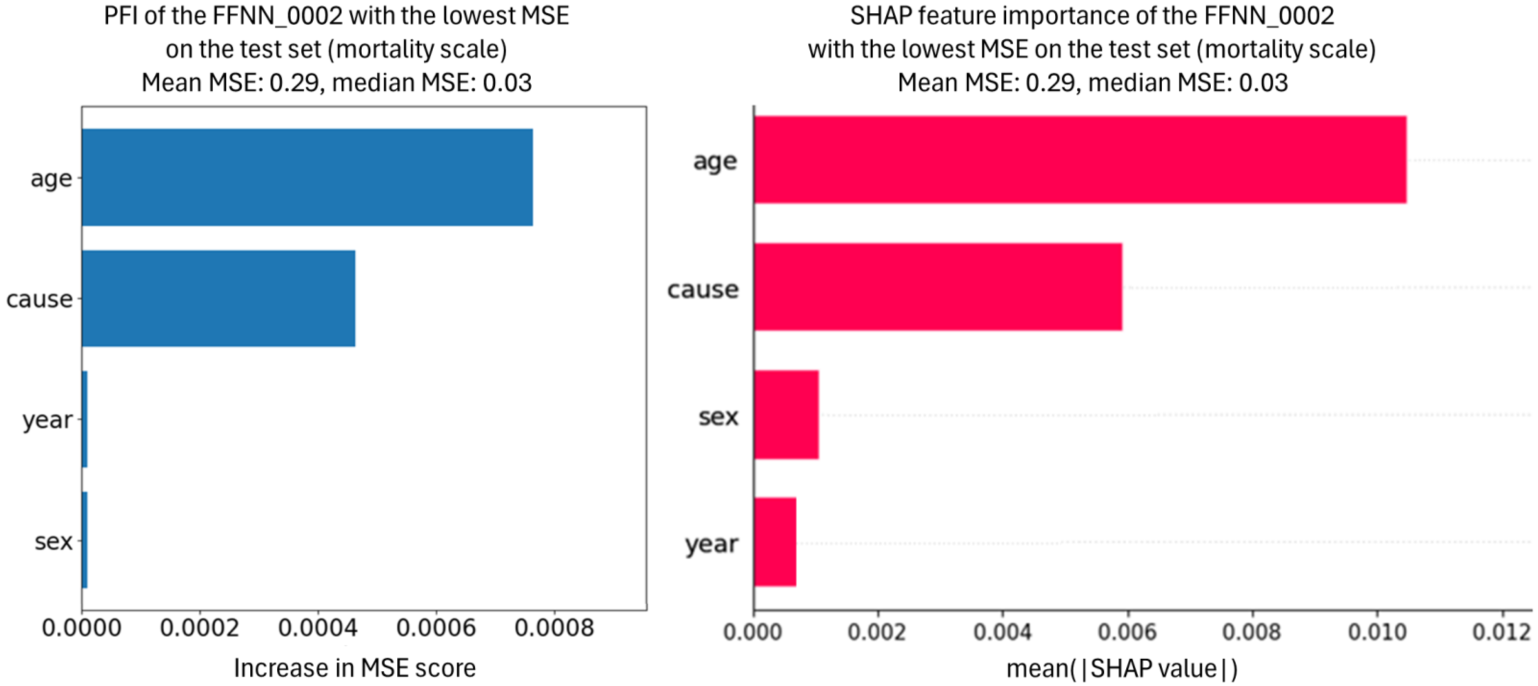


Figure 8. Feature importance of a single FFNN_0002 with the lowest MSE in predicting mortality rates on test dataset

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

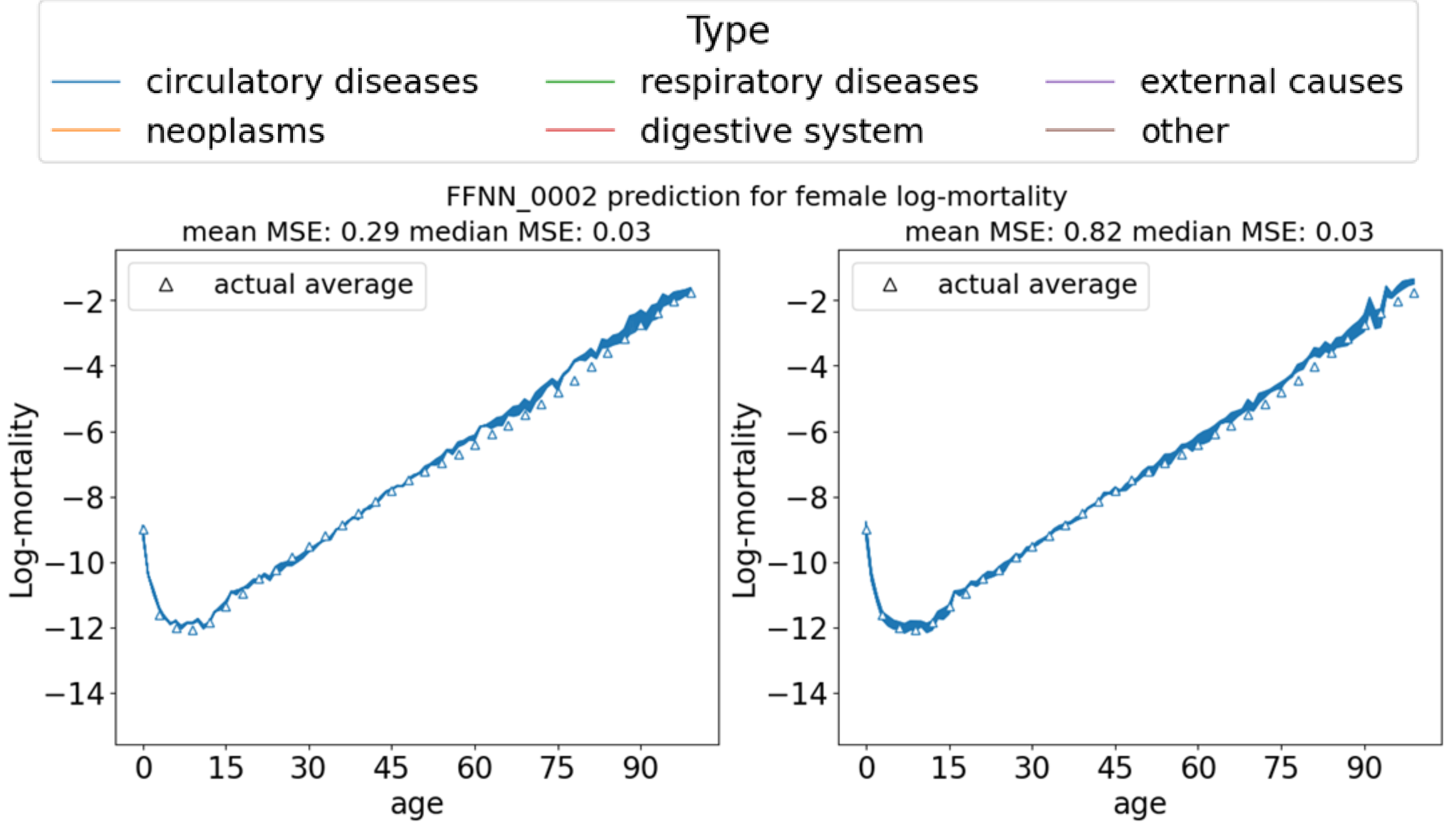


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

Findings: Network Interpretations

(Predicted Log-Mortality Coloured by Cause)

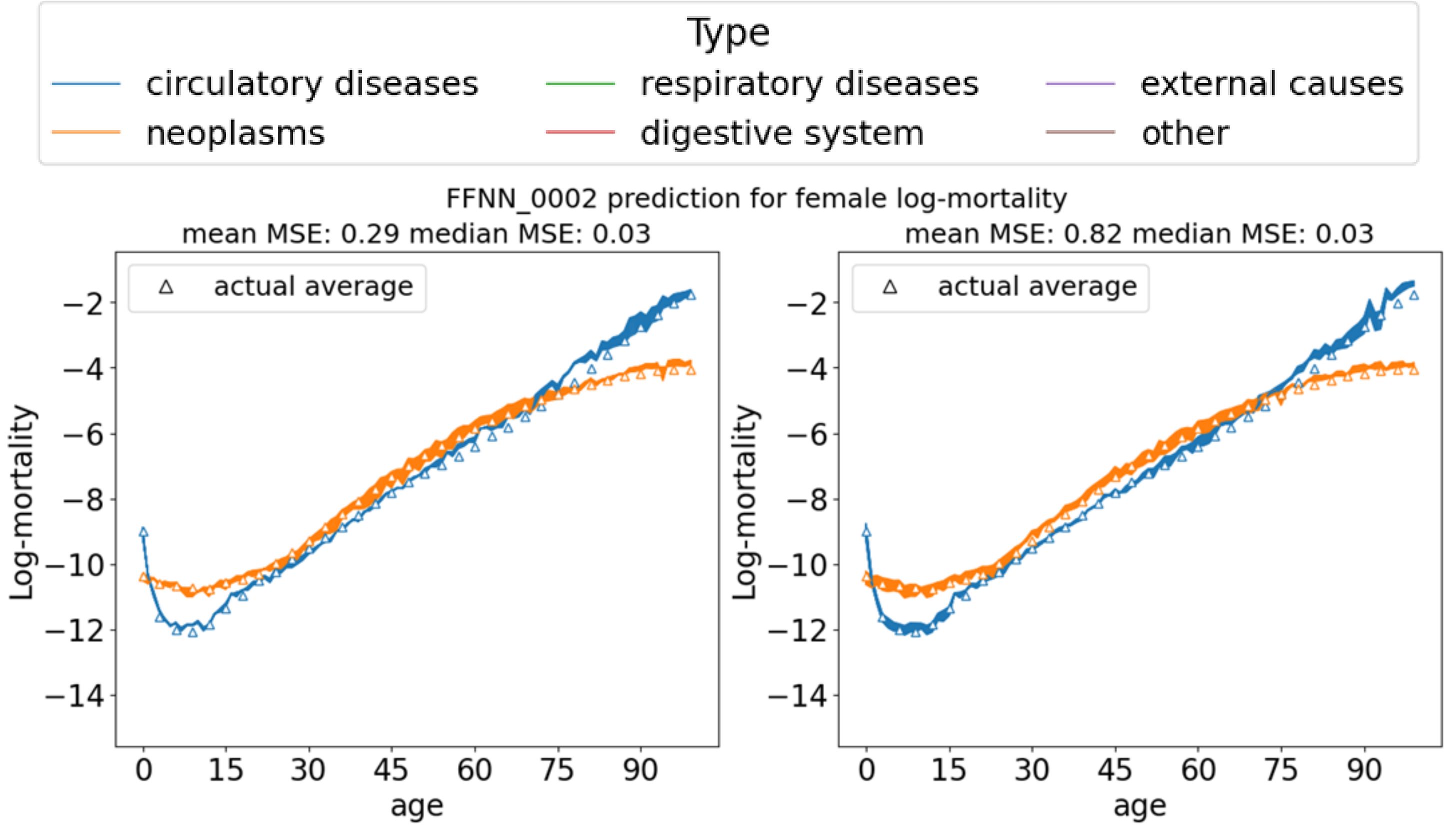


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

Findings: Network Interpretations

(Predicted Log-Mortality Coloured by Cause)

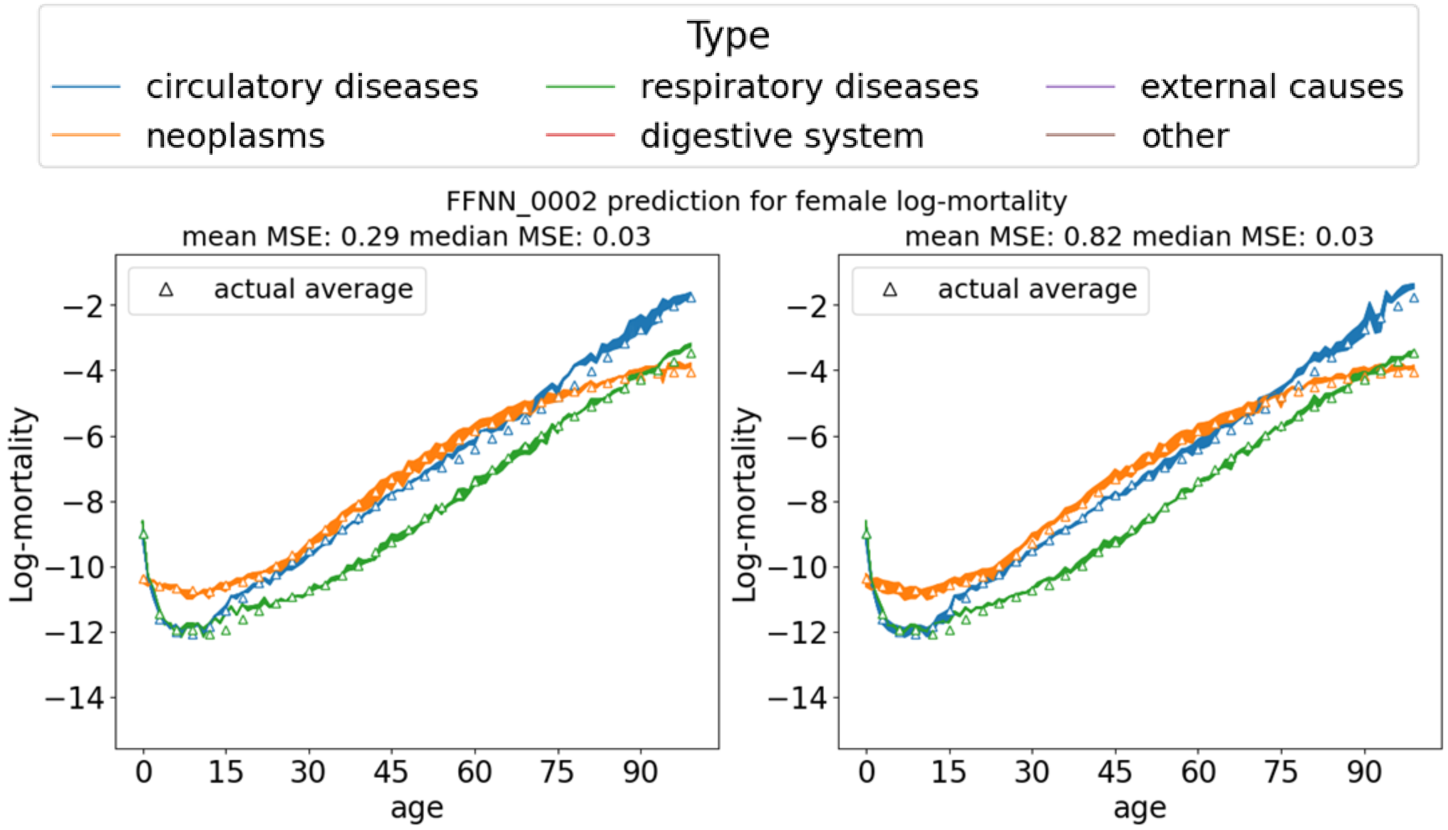


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

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

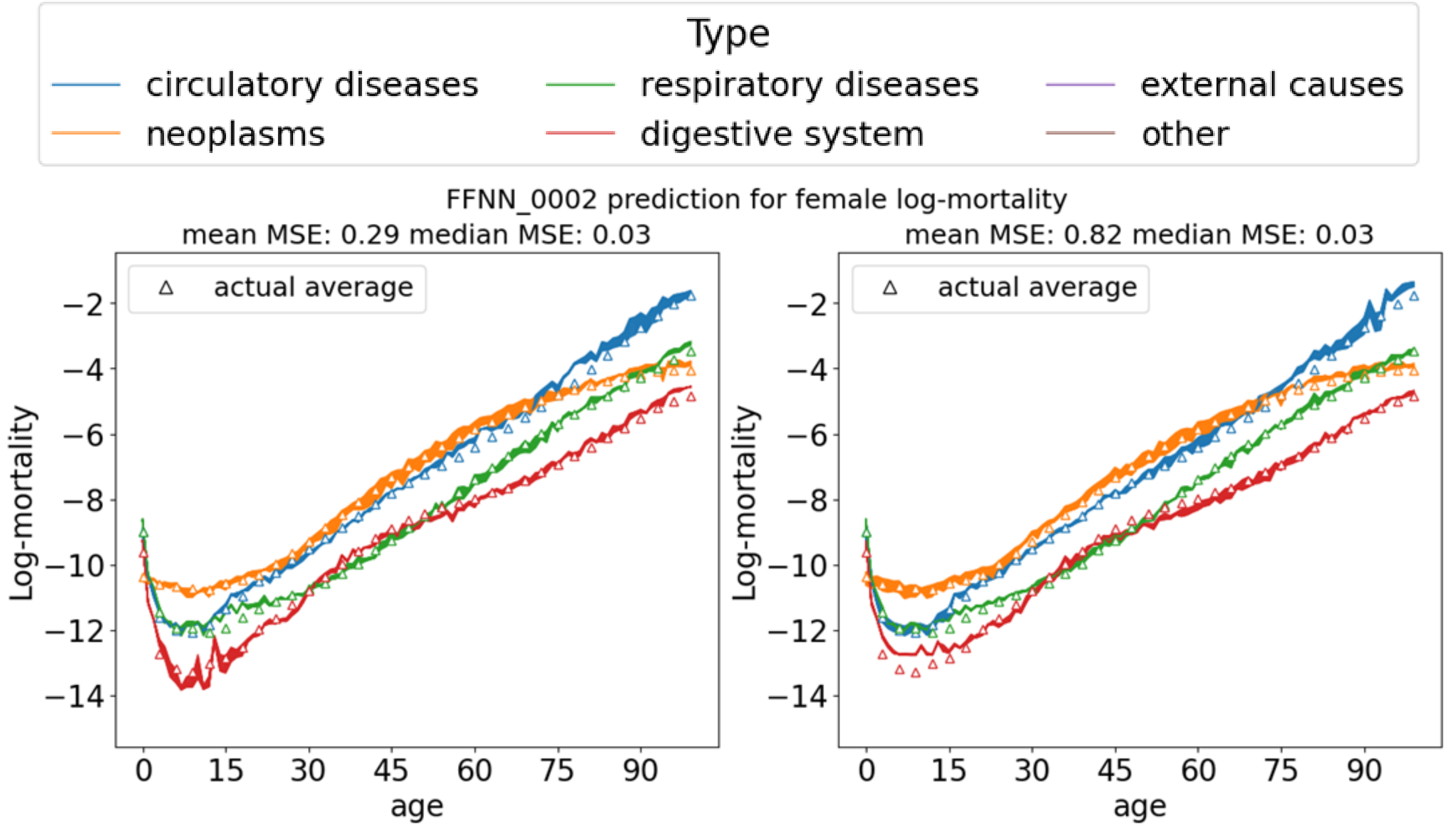


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

Findings: Network Interpretations

(Predicted Log-Mortality Coloured by Cause)

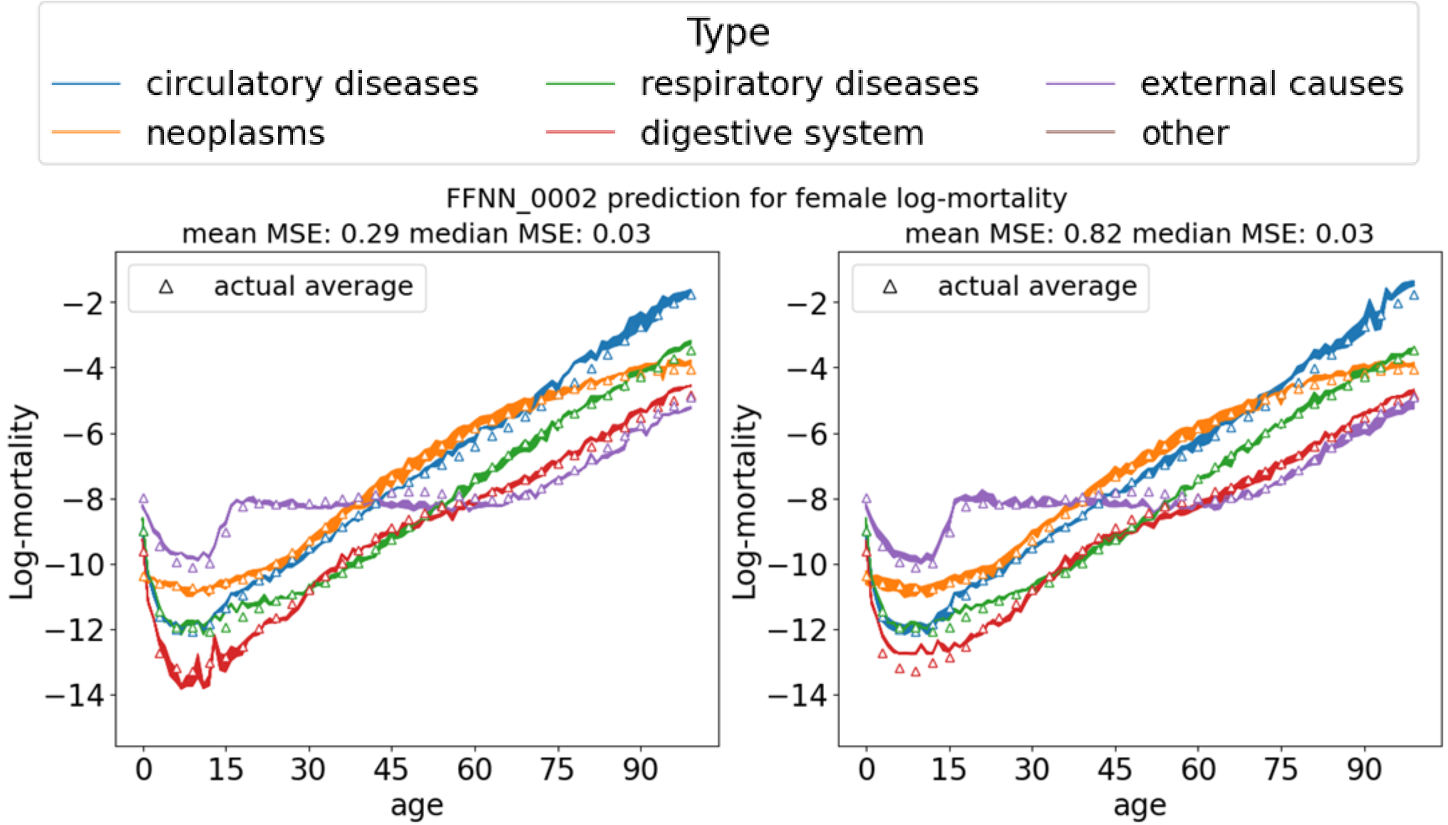


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

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

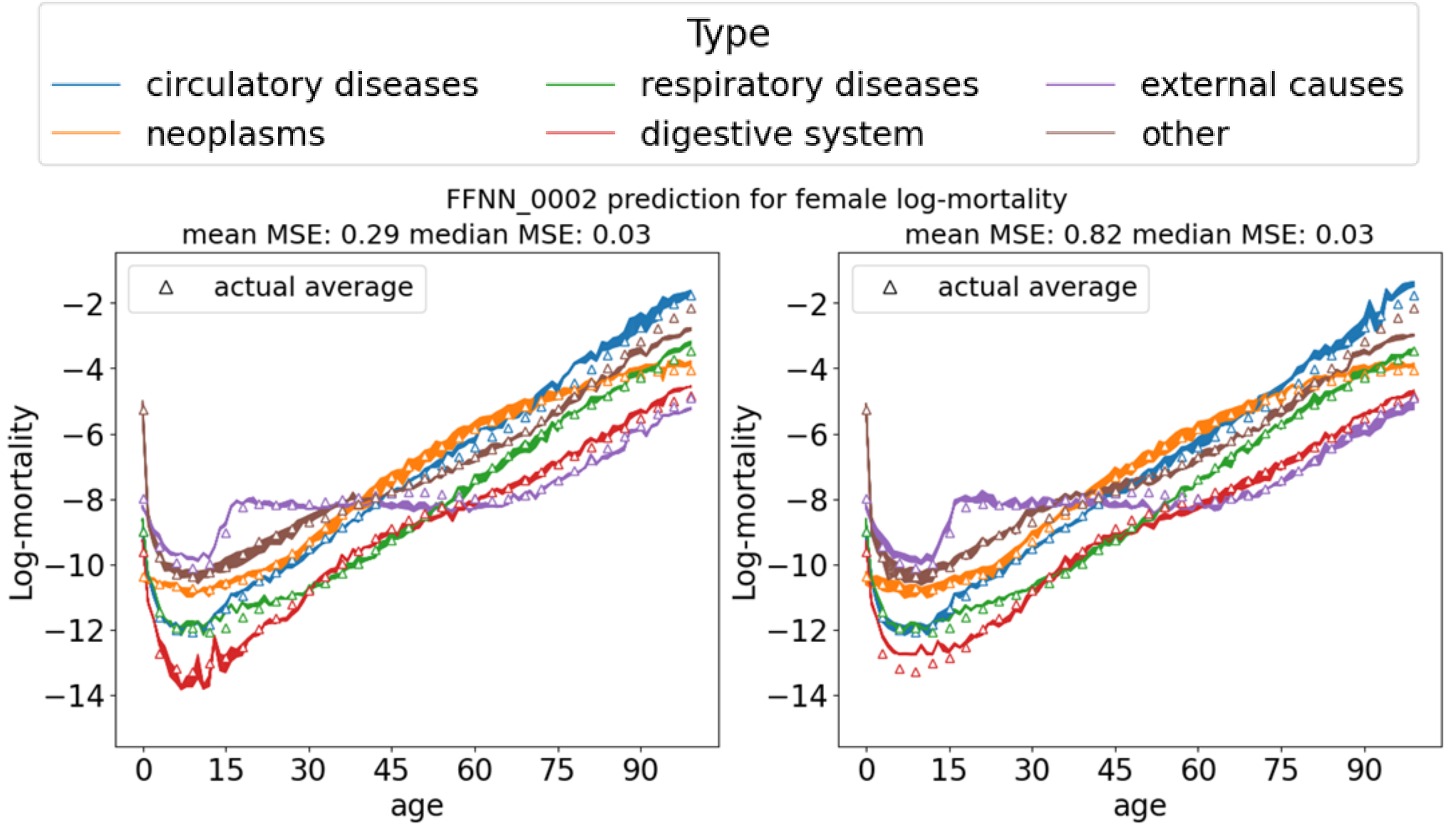


Figure 9. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

Findings: Network Interpretations

(Predicted Log-Mortality Coloured by Cause)

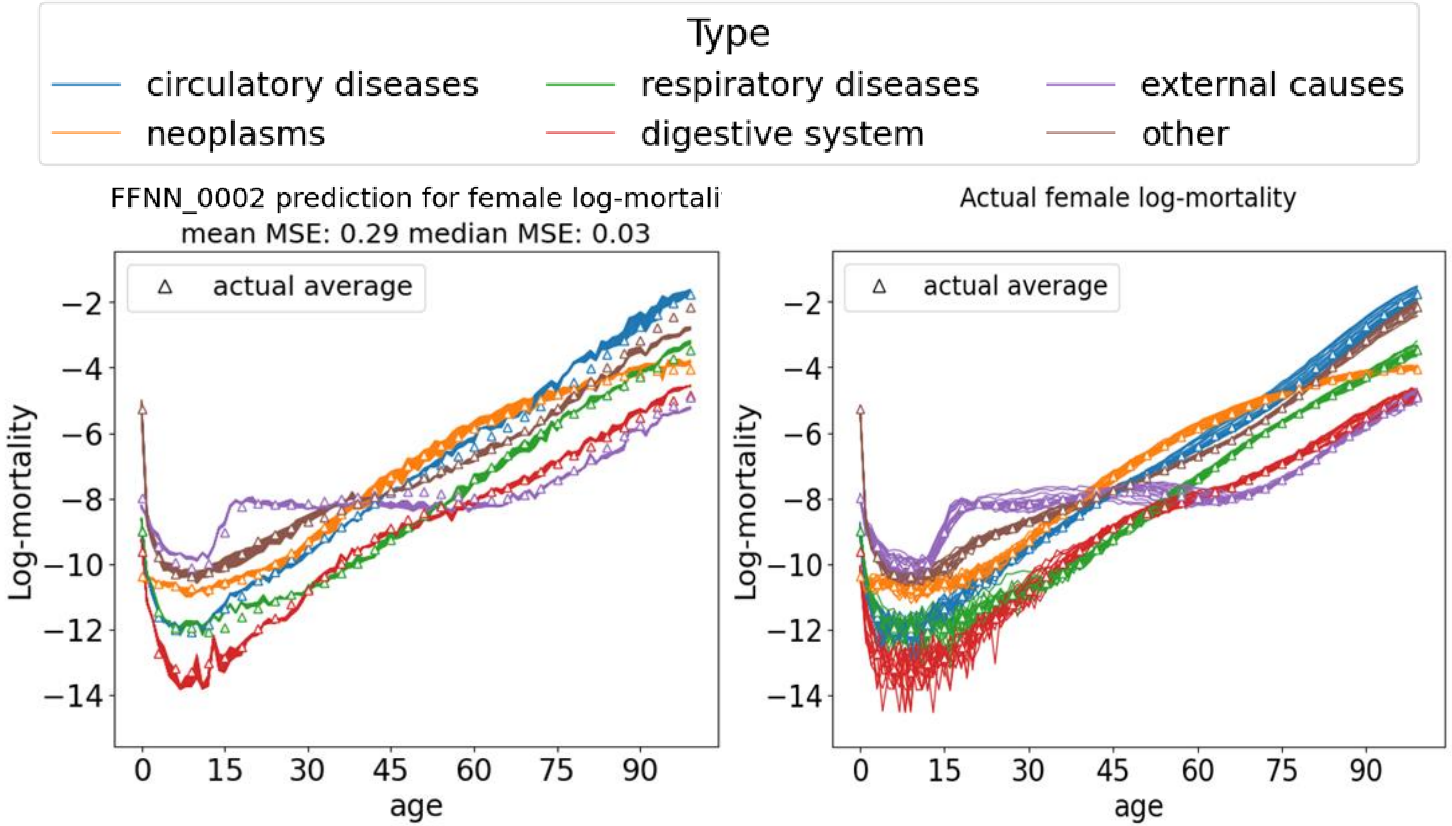


Figure 10. Predicted log-mortality from FFNN_0002 and actual log-mortality for male (left) and female (right) across all years in the testing dataset. Different color represents different cause, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

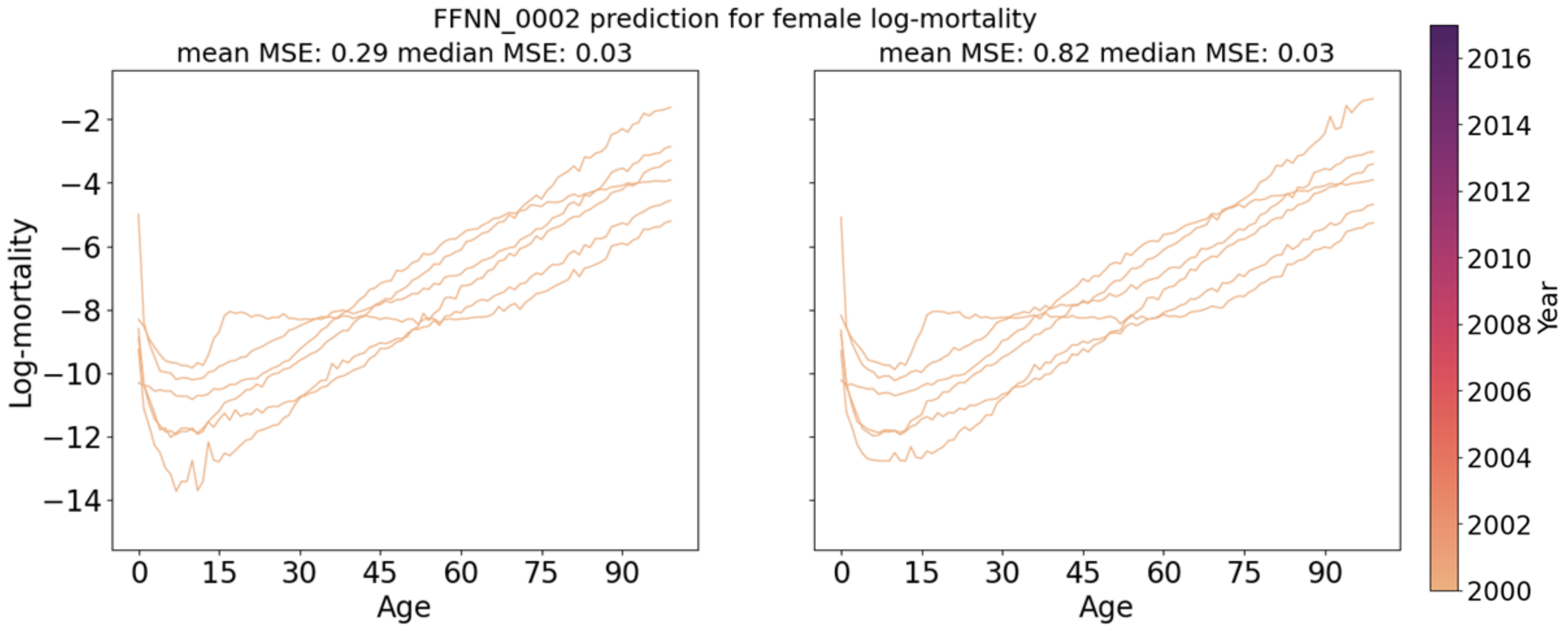


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

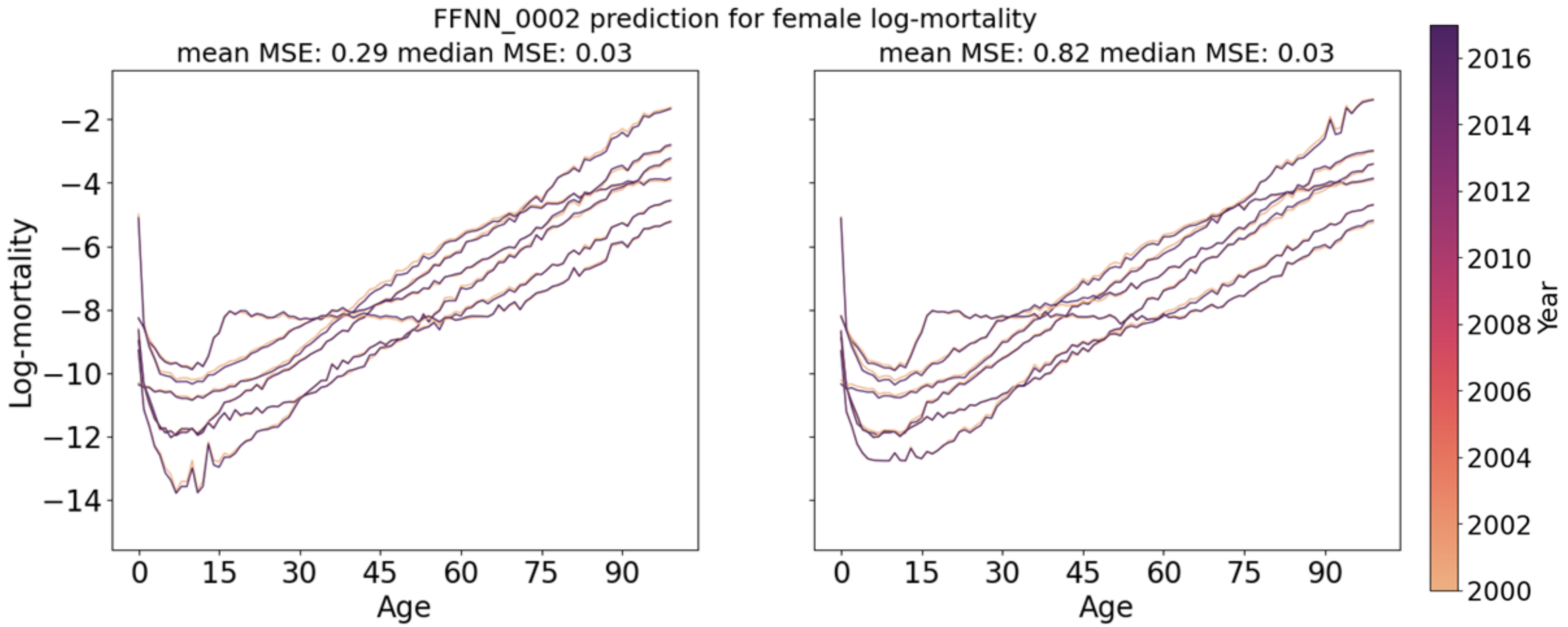


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

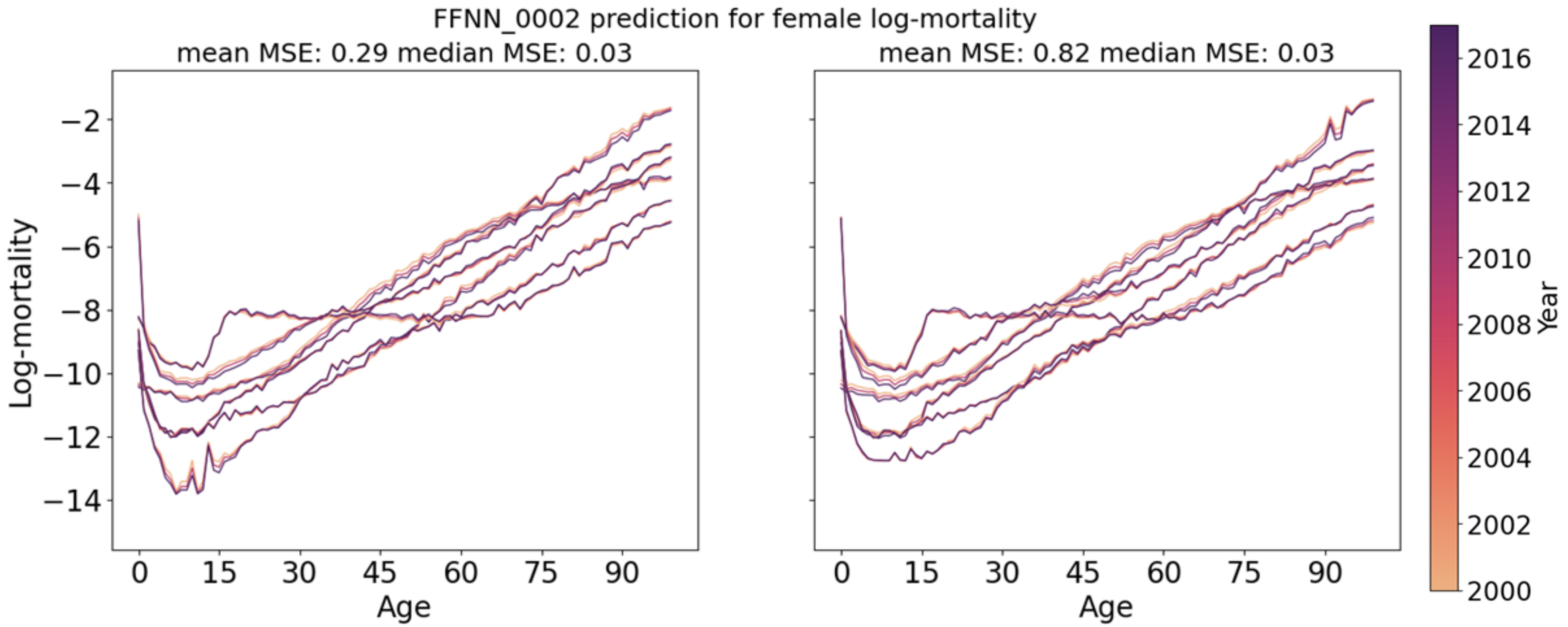


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

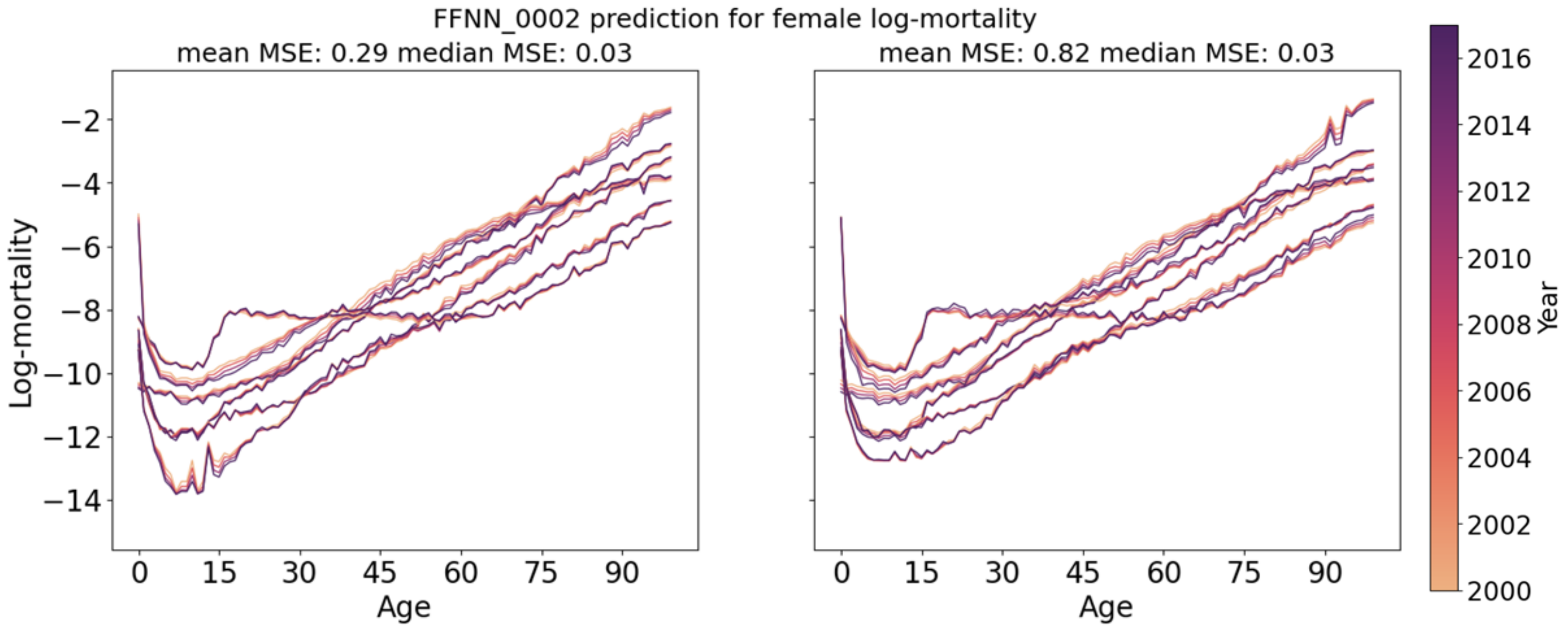


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

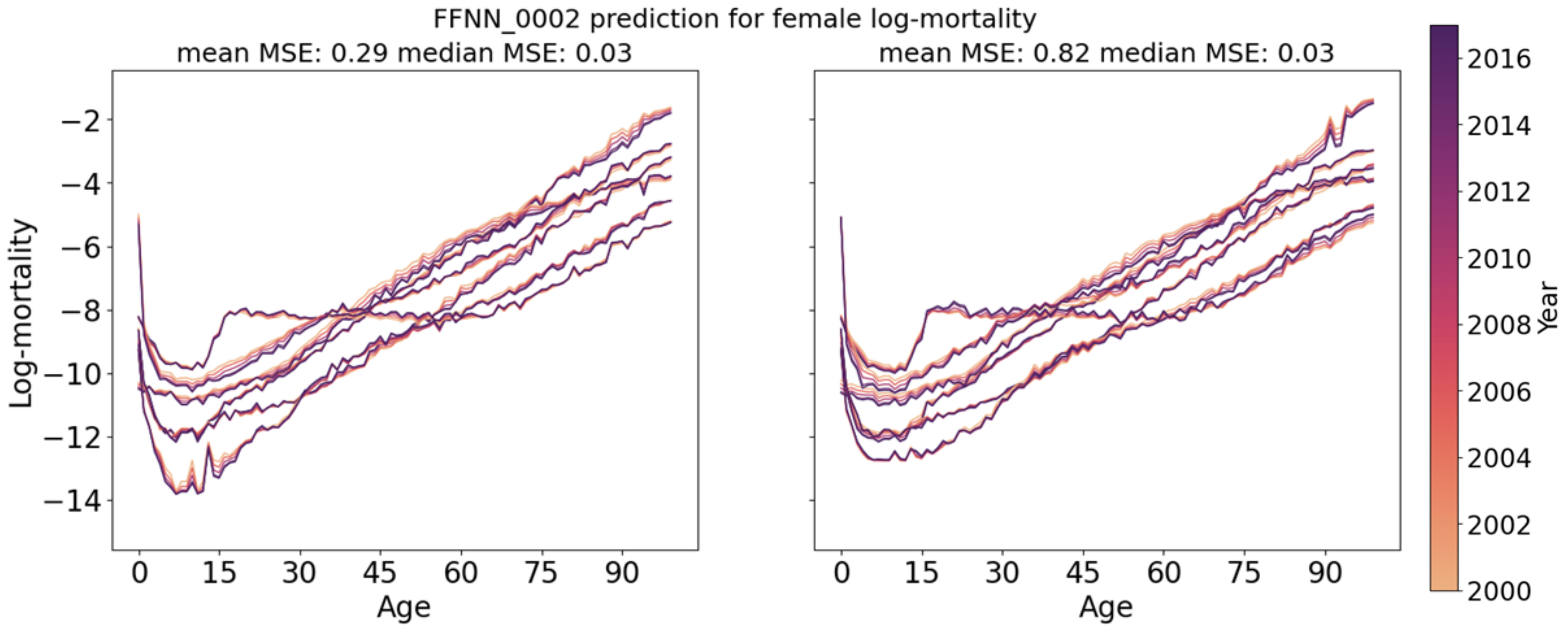


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

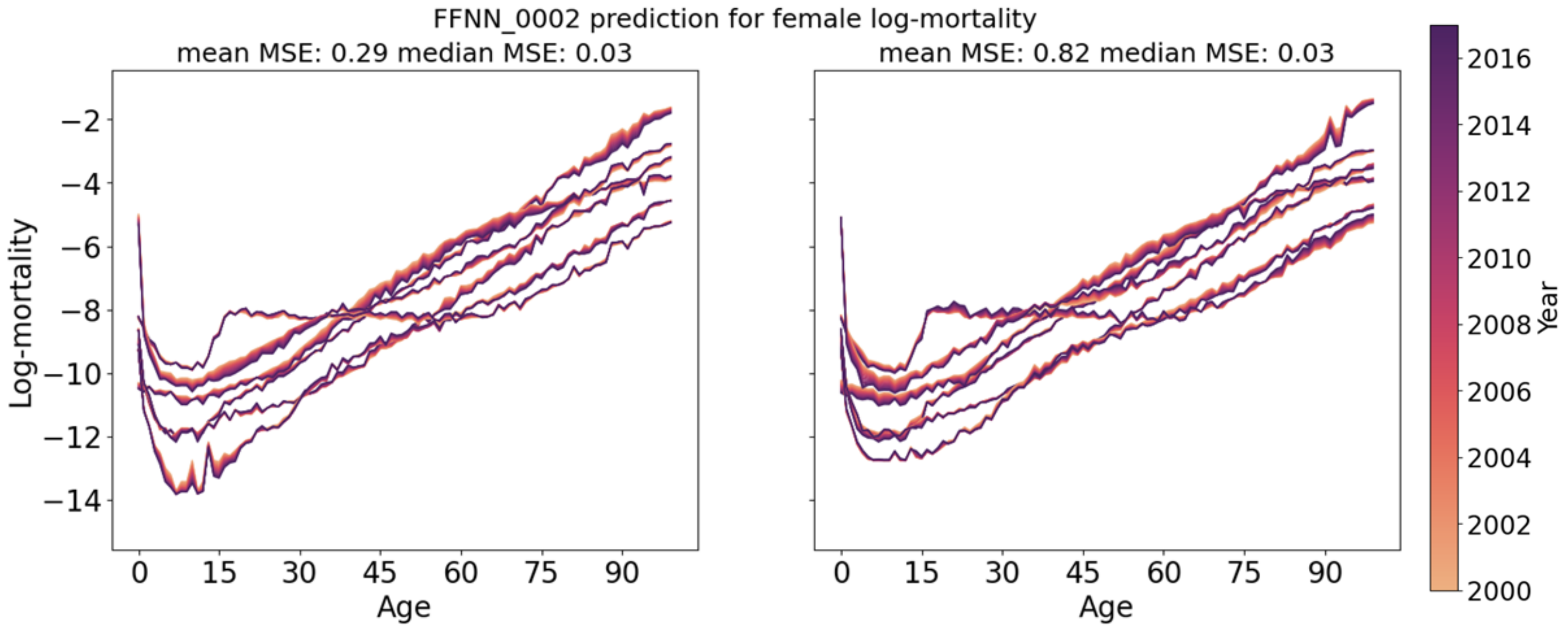


Figure 11. Predicted log-mortality from FFNN_0002 for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(Predicted Log-Mortality Coloured by Year)

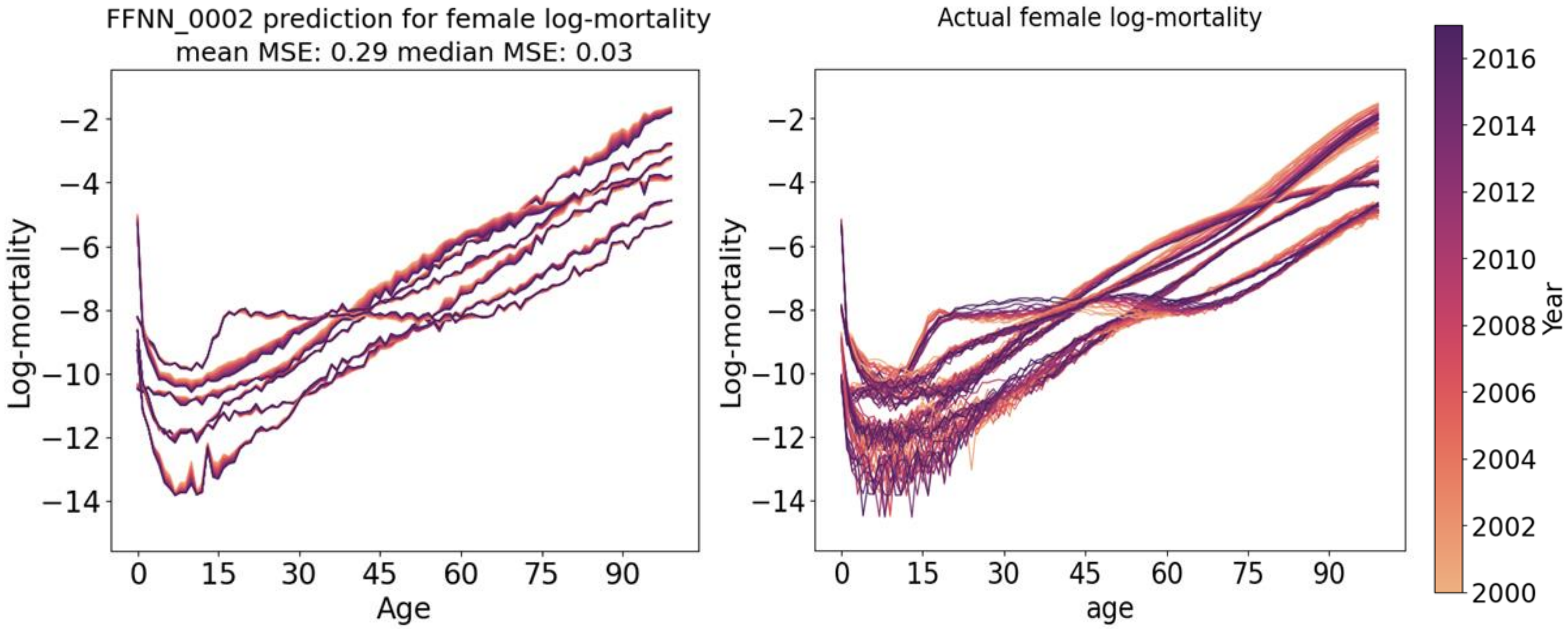


Figure 12. Predicted log-mortality from FFNN_0002 and actual log-mortality for male (left) and female (right) across all years in the testing dataset. Different color represents different year, while different columns represent different weight initialization.

Findings: Network Interpretations
(SHAP Scatter Plots)

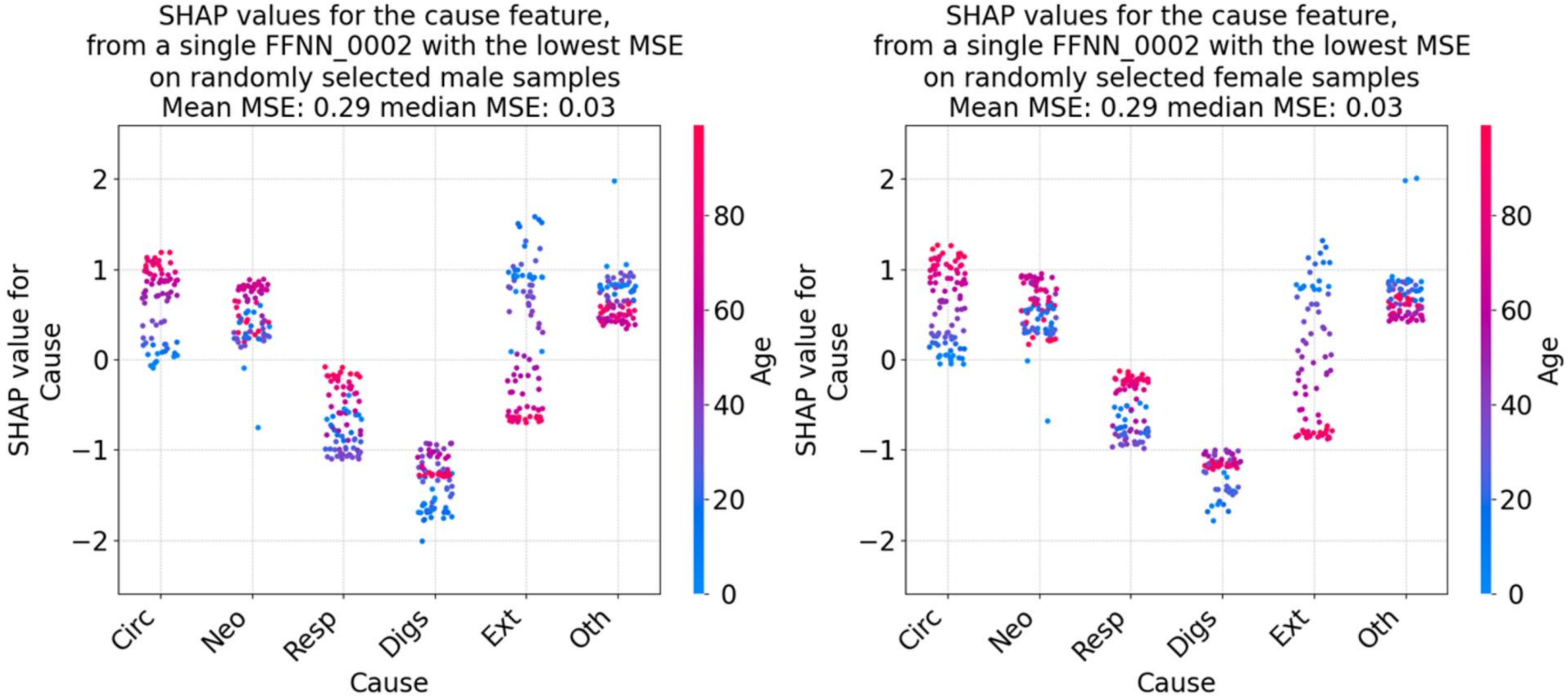


Figure 13. SHAP values for the cause feature from the FFNN_0002 model with the lowest mean MSE for male (left) and female (right) samples. Each marker represents the contribution of the cause feature for a specific sample, and its color indicates its age. All values are on a log-mortality scale and calculated with data from1999 as the baseline.

Final Remarks

- Conclusions:
  - FFNN can model the cause-of-death log-mortality rate (better than the LC, APC, and RH models in general).
    - Perceive circulatory, neoplasm, and respiratory as relatively similar.
    - Capture dependency structure between age and cause.
    - Has not utilized effectively the year features.
  - Transfer learning using simulated data improves network learning.
- Further research:
  - Experiment with network architecture (to incorporate sequence data, cause elimination, or reconciliation, and to improve interpretability).
  - Experiment with cause-of-death dataset from other countries.

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

Boumezoued, A., Coulomb, J.-B., Klein, A., Louvet, D., & Titon, E. (2019). Modeling and Forecasting Cause-of-Death Mortality. https://www.soa.org/4b140c/globalassets/assets/files/resources/research-report/2019/cod-mortality-forecasting.pdf

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

Hu, Q., Zhang, R., & Zhou, Y. (2016). Transfer learning for short-term wind speed prediction with deep neural networks. Renewable Energy, 85, 83–95. https://doi.org/10.1016/J.RENENE.2015.06.034

Lyu, P., De Waegenaere, A., & Melenberg, B. (2021). A Multi-population Approach to Forecasting All-Cause Mortality Using Cause-of-Death Mortality Data. North American Actuarial Journal, 25(S1), S421–S456. https://doi.org/10.1080/10920277.2019.1662316

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

Bibliography

Perla, F., Richman, R., Scognamiglio, S., & Wüthrich, M. V. (2024). Accurate and explainable mortality forecasting with the LocalGLMnet. Scandinavian Actuarial Journal. https://doi.org/10.1080/03461238.2024.2307620

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

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

Wang, J., Wen, L., Xiao, L., & Wang, C. (2023). Time-series forecasting of mortality rates using transformer. Scandinavian Actuarial Journal. https://doi.org/10.1080/03461238.2023.2218859

Thank you!