

eAppendix for “Changes in body mass index and rates of death and transplant in hemodialysis patients: a latent class joint modeling approach”

Samuel L. Brilleman, MSc ^{1,2*}

Margarita Moreno-Betancur, PhD ^{1,2,3,4}

Kevan R. Polkinghorne, PhD ^{1,5,6}

Stephen P. McDonald, PhD ^{7,8}

Michael J. Crowther, PhD ⁹

Jim Thomson, MSc ¹⁰

Rory Wolfe, PhD ^{1,2}

¹Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia; ²Victorian Centre for Biostatistics (ViCBiostat), Melbourne, Australia; ³Clinical Epidemiology and Biostatistics Unit, Murdoch Childrens Research Institute, Melbourne, Australia; ⁴Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia; ⁵Department of Nephrology, Monash Medical Centre, Melbourne, Australia; ⁶Department of Medicine, Monash University, Melbourne, Australia; ⁷ANZDATA Registry, SA Health and Medical Research Institute, Adelaide, Australia; ⁸Department of Medicine, University of Adelaide, Adelaide, Australia; ⁹Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, UK; ¹⁰Arthur Rylah Institute for Environmental Research, Department of Environment, Land, Water and Planning, Victoria, Australia.

*Corresponding author: sam.brilleman@monash.edu

Contents

1. Details related to patient exclusions	3
2. Observed BMI trajectories	3
3. Numbers of patients with BMI measurements in the 1 st , 2 nd , 3 rd , 4 th and 5 th years	3
4. Computational details for the latent class joint model.....	4
4.1. Overview.....	4
4.2. Example code.....	4
4.3. Initial values.....	5
4.4. Computation times.....	7
5. Choice of individual-level random effects structure	8
6. Results for the six-class joint model	8
7. Cumulative incidence functions	9
8. Goodness of fit of the final model (observed vs predicted longitudinal trajectories).....	10
9. GRoLTS checklist.....	11

1. Details related to patient exclusions

eFigure 1 provides a flowchart illustrating the derivation of the analysis sample. Of the 19,264 patients who initiated hemodialysis during the study period, we excluded the following: 536 patients who recovered kidney function, 879 patients with unknown height, race, or comorbidity status, 21 patients with extreme (<130 cm) baseline height, 1376 patients with an extreme (≥ 45 or <17.5 kg/m²) BMI measurement during follow up. Since there were some overlaps for these exclusion categories (e.g. a patient might have had a missing height measurement *and* have recovered kidney function) these exclusions meant there were 16,585 patients remaining, of which 16,414 patients had at least one BMI measurement recorded prior to death, transplant or censoring and were therefore included in the main analyses.

2. Observed BMI trajectories

Observed longitudinal BMI trajectories for a random sample of 25 patients are shown in eFigure 2.

3. Numbers of patients with BMI measurements in the 1st, 2nd, 3rd, 4th and 5th years

In the main manuscript we report that the percentage of patients with 1, 2, 3, 4, and 5 BMI measurements respectively (prior to death, transplant or censoring) was 18%, 21%, 17%, 14% and 30%. In Table 1 of the main manuscript we also report the specific number of patients with 1, 2, 3, 4, and 5 BMI measurements respectively. Here we provide some additional information about the frequency of the BMI measurements:

- Of the 12,449 patients who were still at risk of an event after 1 year of follow up, we found that 12,325 (99.0%) patients had a BMI measurement during their first year.
- Of the 9,196 patients who were still at risk of an event after 2 years of follow up, we found that 9,140 (99.4%) patients had a BMI measurement during their second year.
- Of the 6,588 patients who were still at risk of an event after 3 years of follow up, we found that 6,540 (99.3%) patients had a BMI measurement during their third year.

- Of the 4,511 patients who were still at risk of an event after 4 years of follow up, we found that 4,484 (99.4%) patients had a BMI measurement during their fourth year.

Of the 3,049 patients who were still at risk of an event after 5 years of follow up (i.e. those who were censored at the maximum follow up time), we found that 3,019 (99.0%) patients had a BMI measurement during their fifth year.

4. Computational details for the latent class joint model

4.1. Overview

The latent class joint model, defined in the main manuscript, was estimated using the 'Jointlcmm' command in the 'lcmm' R package [1–3]. Maximum likelihood estimates are obtained using a modified Marquardt algorithm (see the package documentation for details [2]). The package's default setting enforces strict convergence criteria, which we did not alter (tolerance of $1E-4$ for each of: parameter stability, log-likelihood stability, and stability of the first derivatives). We chose to use the default square transformation of the baseline hazard parameters (argument 'logscale = FALSE') to ensure positivity of the baseline hazard throughout the estimation process.

Several possible extensions to the model were also explored, but often led to difficulties achieving convergence. Attempts were made to use a baseline hazard modelled with cubic M-splines (argument 'hazard = "splines"'), however, difficulties with convergence were encountered (even when using a small number of internal knots for the splines). Therefore, a simpler Weibull baseline hazard specification was used for the models in the main manuscript. We also encountered difficulties achieving convergence when the variance of the individual-level random effects was allowed to differ across latent classes.

4.2. Example code

In several supplementary files, we provide an example of the R code used to fit the latent class joint

model. These files are named:

- “example_code_default.R”
- “example_code_gridsearch.R”
- “example_code_randominits.R”

The three files differ in their approach to starting / initial values for the parameters. The three approaches used for initial values are described in the next section.

4.3. Initial values

We found that we encountered difficulties with convergence unless reasonable initial values were used. In particular, our solution reported in the main manuscript was found using two strategies, described as follows.

Strategy A: example code shown in the supplementary file “example_code_default.R”.

A1. We first fit a model with the same structure as our intended latent class joint model, but with only one class (i.e. specifying the argument ‘ng = 1’ to the ‘Jointlcmm’ function in the lcmm R package). We denote this model: ‘initmod’.

A2. We then use the parameter estimates from the one class model (‘initmod’) to generate initial values for the M-class model we wish to fit (where M is the number of desired classes). This is achieved by specifying the arguments ‘ng = M’ and ‘B = initmod’ to the ‘Jointlcmm’ function. We use a maximum of 200 iterations. We denote this model: ‘mod’.

A3. If the M-class model (‘mod’) converged, then we stop. Otherwise, if the M-class model (‘mod’) did not converge after 200 iterations then we do the following:

A3.1. We extract the final parameter estimates from the M-class joint model that did not converge.

A3.2. We then fit an M-class model for the longitudinal BMI data only (ignoring the death and

transplant data) and we extract the final parameter estimates from that model.

A3.3. We substitute the estimates from Step A3.2 into the vector of parameter estimates from Step A3.1 and then refit the model using those parameter estimates as starting values for a new attempt at fitting the M-class joint model.

Strategy B: example code shown in the supplementary file “example_code_gridsearch.R”. Note that Strategy B is the same as Strategy A, except that a grid search is used in the second step.

B1. Same as Step A1.

B2. We then use the parameter estimates from the one class model ('initmod') to randomly generate 3 sets of initial values for M-class model, we then run each of these 3 models for a maximum of 50 iterations. This is achieved by using the 'gridsearch' function in the lcmm package, with arguments 'rep = 3' and 'maxiter = 50'. We then take the model with the highest log-likelihood, and run that model until convergence or until the maximum number of iterations (200) is reached. Let us denote this model: 'mod'.

B3. Same as A3.

We then confirmed that Strategy A and Strategy B led to the same solution. Note that in most cases the model in Step A2/B2 did not converge, and the parameter estimates from an M-class model with just the longitudinal BMI data was required to produce starting values for the M-class joint model (i.e. Steps A3.1 through A3.3, or Steps B3.1 through B3.3, were required).

Additional Strategy: In response to a reviewer's suggestion, we then also ran the five-class joint model (i.e. our final model) with a variety of random initial values. Specifically, we used 27 randomly generated sets of initial values. The function to generate the initial values can be found in the supplementary file “example_code_randominits.R”. From these 27 models, we found the following:

- 11 stopped abnormally (with no additional information provided by the lcmm package)
- 13 models did not converge after 200 iterations

- 3 models converged

Of the models that converged, one provided a seemingly nonsensical solution. The two other models converged to a common solution that was very similar, but not identical, to the solution found using Strategy A and Strategy B.

Compared with the final solution reported in the manuscript (found using Strategy A and Strategy B), the solution found with the random initial values had a slightly higher log-likelihood (-27324 vs -27434) and slightly lower BIC value (55367 vs 55586). However, these differences are small relative to the decrease in BIC observed with adding additional latent classes; for example, the six-class solution reported in the manuscript had a BIC value of 54878. Moreover, the findings related to the predicted longitudinal trajectories and hazard functions (see eFigure 3) were very similar to the results reported in the manuscript.

Since multimodality is such a critical issue in latent class models, one ideally wants to use many sets of randomly selected initial values. This can provide reassurance (but not certainty) that the final solution corresponds to a global maximum. Ideally, many sets of randomly selected initial values would also be used for the models estimated along the model building / selection / comparison process. However, we found that such an endeavour is somewhat hampered in latent class joint models by their computational complexity and - in particular – issues with convergence and long computation times (see the next section). Nonetheless, it is worth noting that the computation times would be shorter if fitting these models to a smaller dataset (for example less individuals and/or less longitudinal measurements), which would allow a greater number of models to be estimated in a given time.

4.4. Computation times

Fitting the latent class joint models was time consuming. For example, for the five-class joint model using Strategy A for specifying the initial values, it took approximately 30 hours to fit the final model using a single 2.5GHz core on the Monash University computing cluster. However, Steps A3.1

through A3.3 (see Strategy A in the previous section) of that process took only 3 hours, meaning that once reasonable starting values were used for the five class-specific longitudinal BMI trajectories the estimation time decreased dramatically.

5. Choice of individual-level random effects structure

We considered adding additional individual-level random effects to the model. Specifically, we considered including individual-level random effects for the coefficients of the cubic splines basis terms. However, this led to a model where the predicted BMI trajectory for each latent class was relatively stable/flat, with the latent classes primarily distinguished by different starting/average BMI values, or differences in the event rates, and not by differences in the *shapes* of the longitudinal BMI trajectories (see eFigure 4). That is, in a model that included individual-level random effects for the cubic spline terms, variation in the shapes of the longitudinal BMI trajectories appeared to be attributed to between-individual (i.e. within-class) heterogeneity.

Accordingly, we chose to simplify the individual-level random effects structure in our model (that is, only include an individual-level random intercept). By doing this, we believe that differences in the shapes of the longitudinal profiles are exhibited primarily through between-class differences, and less absorbed by within-class (between-individual) variation. We believe that this approach more closely aligns with our study objectives, specifically to explore differences in the shapes of the class-specific longitudinal BMI trajectories, and how those are associated with differences in the class-specific rates of the competing events.

6. Results for the six-class joint model

eFigure 5 shows the predicted BMI trajectories and cause-specific hazard functions, for each latent class, based on the six-class model. The covariate values used in the predictions are the same as used for the hazard functions in Figure 2 of the main manuscript. The difference here is that eFigure 5 is for the six-class model, whereas Figure 2 in the main manuscript is for the five-class model.

As discussed in the main manuscript, we calculated Bayesian information criterion (BIC) values for models with a varying number of latent classes. The BIC values suggested that higher numbers of latent classes consistently resulted in a better fitting model. However, as the number of latent classes increased, the groups became less distinguished from one another (accompanied by a decrease in relative entropy, see Table 1 of the manuscript) and therefore less useful in terms of drawing meaningful conclusions from a clinical perspective. That is, based on a purely statistical criterion (the BIC), there is a suggestion that increasing numbers of classes are better, but from an interpretational perspective an ever-increasing number of latent classes was not useful. We therefore chose the five-class model as our final model.

It can be seen in eFigure 5 that, for the six-class model, the main difference from the five-class model is that the “late BMI decline” class is split based on baseline BMI values and associated hazard rates for transplant. Since our primary interest is in the association between BMI and risk of death, we determined that increasing the number of latent classes to six was not warranted and we had similar conclusions from exploratory analysis with models having higher numbers of classes.

7. Cumulative incidence functions

The cause-specific hazard functions presented in Figure 2 of the main manuscript show the instantaneous rates (i.e. “hazard”) of each event at time t , given that the individual is still at risk of the event. Alternatively, we can present cumulative incidence functions for each of the competing events; these are shown in eFigures 6 (without 95% confidence limits) and 7 (with 95% confidence limits), for the same covariate profile as used for the hazard functions in Figure 2 of the main manuscript. The cumulative incidence functions show the cumulative risk (i.e. probability) of the event having occurred at any point up to time t .

Broadly speaking, the cause-specific hazard functions are useful for understanding the potential for etiological associations between the class-specific BMI trajectories and the occurrence of the competing events and thus more suited to the aims of our manuscript. On the other hand, the cumulative incidence functions are generally suited to understanding patient prognosis; for example,

“what is the probability that a patient in latent class X will experience death within 5 years”.

Importantly, the cumulative incidence function for one of the events, say death, depends on the hazard rates for both of the competing events. For example, whether a patient dies within 5 years depends partly on the rate of death, and partly on their rate of the competing event of transplant. Therefore, the association of a characteristic, say BMI, and the cumulative incidence function of an event, say death (without transplant), will depend on the associations between BMI and the hazard of both competing events, say death (without transplant) and transplant. Thus, in general, the association of a characteristic with the hazard of an event will be different to that with the cumulative incidence function of the same event, and in extreme cases could even be in opposite directions [4]. If the aims of our paper had been related to developing a model for patient prognosis, then we would have been interested in measures of predictive performance for the fitted models. Moreover, issues such as non-proportional hazards would have been more relevant since they can have a significant impact on the prognostic performance of the fitted model.

For a more thorough discussion of the differences between cause-specific hazard functions and cumulative incidence functions, and how they each align with the intended aims of a study, we refer the reader to Koller et al. [5].

8. Goodness of fit of the final model (observed vs predicted longitudinal trajectories)

eFigures 8 and 9 show observed and predicted, class-specific, longitudinal BMI trajectories. They are based on weighted means of the observed and predicted BMI values. The weighting refers to the estimated class membership probabilities for each individual, whilst the means are taken by splitting the distribution of observed measurement times into 15 quantiles (i.e. 15 “bins”).

The plots show that the predicted mean longitudinal trajectories generally provide a good fit to the observed data. There is some discrepancy between the observed and marginal predicted BMI values beyond 2.5 years for the “rapid BMI decline” class (i.e. the black curve in eFigure 8).

However, this is probably due to the fact that this class has a relatively small number of patients

overall, and has a high mortality rate early on in the follow up period. Therefore, the majority of the BMI measurements for this class are observed earlier in the follow up period. Because the bulk of the data is observed earlier in the follow up period the spline-based trajectory is seen to fit best in that region, whereas it has insufficient flexibility to capture the stabilising of the BMI curve after 2.5 years. Note however, that incorporating the subject-specific random intercept (i.e. the subject-specific predictions in eFigure 9) resolves the discrepancy between the observed and predicted BMI values.

9. GRoLTS checklist

eTable 2 shows a completed GRoLTS (Guidelines for Reporting on Latent Trajectory Studies) checklist for our study [6].

(To satisfy Item #15 of the GRoLTS checklist we have also included another file in our online Supplementary Materials entitled “table_full_model_estimates.txt”. This plain text file includes an unformatted table of the entire list of parameter estimates from the final model.)

References

- 1 Proust-Lima C, Philipps V, Diakite A, *et al.* lcmm: Extended mixed models using latent classes and latent processes. R package version: 1.7.7. Published Online First: 2017.<https://cran.r-project.org/package=lcmm>
- 2 Proust-Lima C, Philipps V, Lique B. Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcmm. *Journal of Statistical Software* 2017;**78**. doi:10.18637/jss.v078.i02
- 3 R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Published Online First: 2017.<https://www.R-project.org/>
- 4 Andersen PK, Geskus RB, de Witte T, *et al.* Competing risks in epidemiology: possibilities and pitfalls. *International Journal of Epidemiology* 2012;**41**:861–70. doi:10.1093/ije/dyr213
- 5 Koller MT, Raatz H, Steyerberg EW, *et al.* Competing risks and the clinical community: irrelevance or ignorance? *Statistics in Medicine* 2012;**31**:1089–97. doi:10.1002/sim.4384
- 6 van de Schoot R, Sijbrandij M, Winter SD, *et al.* The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling: A Multidisciplinary Journal* 2017;**24**:451–67. doi:10.1080/10705511.2016.1247646

eTable 1. Mean posterior probabilities of class membership, stratified by class membership (as determined by an individual's highest class-specific probability, and shown on the rows).

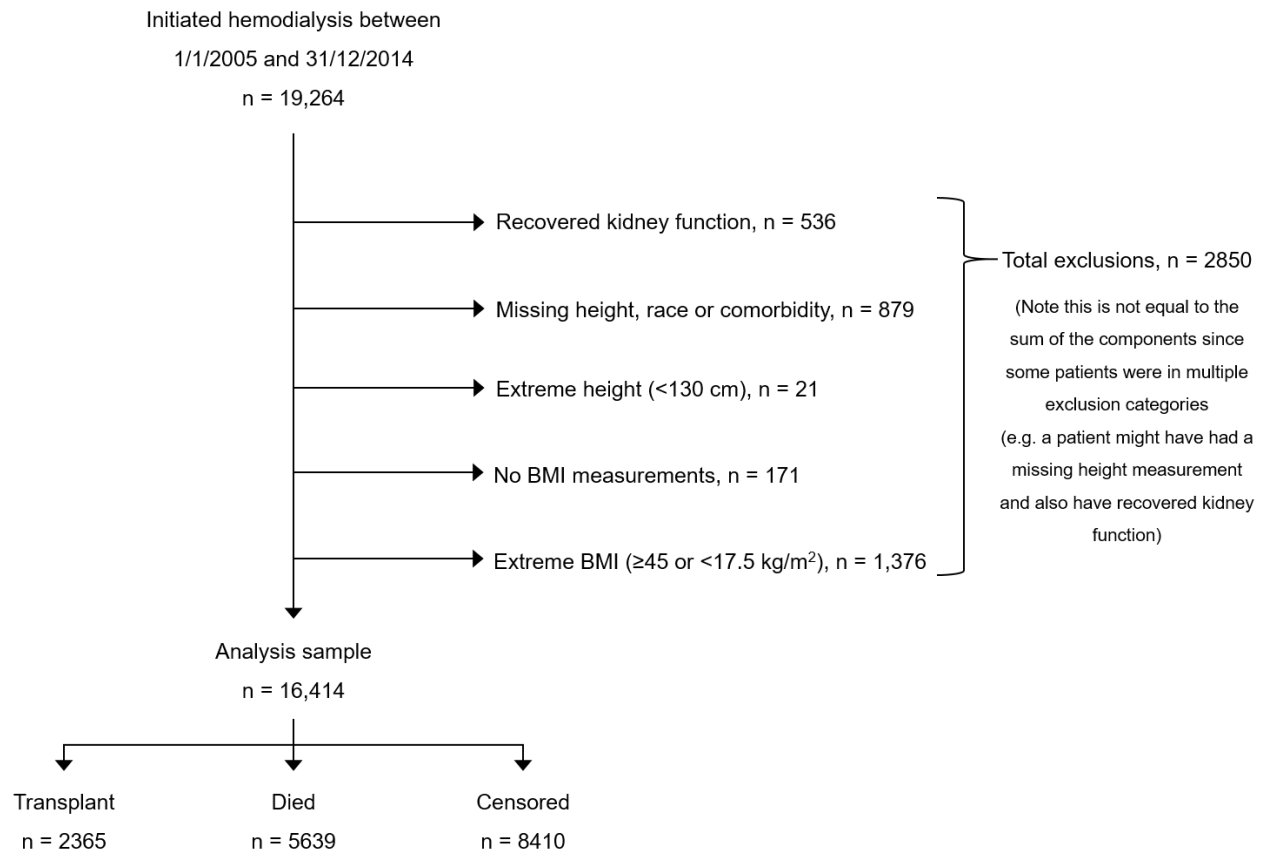
Class membership (% of patients)	Mean probabilities of class membership				
	Prob (Class A)	Prob (Class B)	Prob (Class C)	Prob (Class D)	Prob (Class E)
Class A (74.7%)	0.89	0.04	0.04	0.01	0.03
Class B (13.8%)	0.11	0.73	0.09	0.03	0.04
Class C (6.2%)	0.13	0.08	0.77	0.03	0.00
Class D (1.5%)	0.07	0.02	0.10	0.81	0.00
Class E (3.9%)	0.13	0.08	0.00	0.00	0.78

eTable 2. GRoLTS checklist.

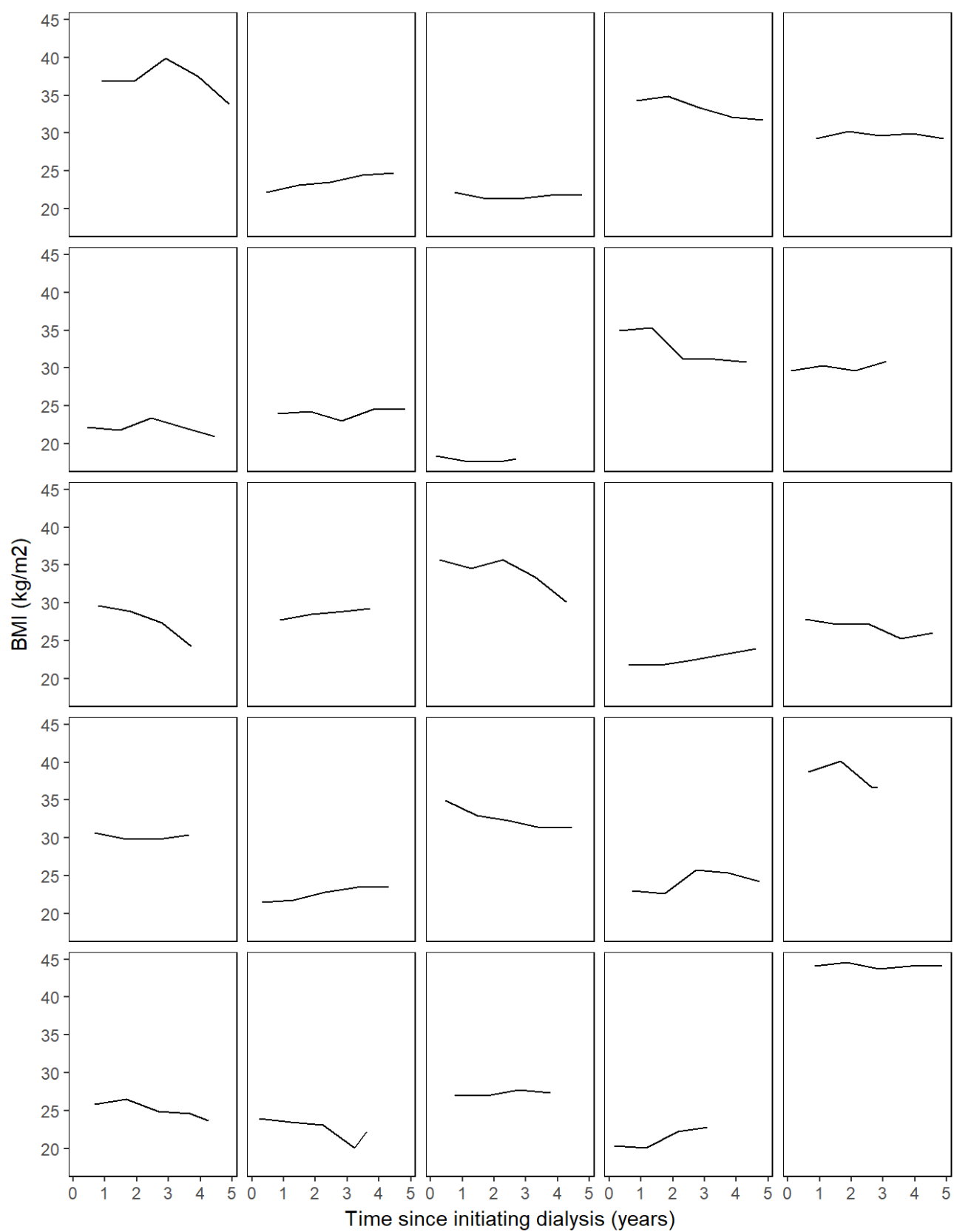
GRoLTS checklist item	Yes/No	Additional comments
1. Is the metric of time used in the statistical model reported?	Yes	Time metric is years.
2. Is information presented about the mean and variance of time within a wave?	NA	Exact time of the measurement (ANZDATA survey date - baseline date) is used in the analysis, so time-structured data is not relevant in this study.
3a. Is the missing data mechanism reported?	Yes	
3b. Is a description provided of what variables are related to attrition/missing data?	Yes	
3c. Is a description provided of how missing data in the analyses were dealt with?	Yes	Described in the Methods section in the manuscript, and supported by eFigure 1 in the Supplementary Materials.
4. Is information about the distribution of the observed variables included?	Yes	Described in the model specification in the manuscript.
5. Is the software mentioned?	Yes	
6a. Are alternative specifications of within-class heterogeneity considered (e.g., LGCA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	We wish to allow for some between-individual variation within classes. Accordingly, we allow for this within-class heterogeneity through individual-level random effects, as described in the model specification.
6b. Are alternative specifications of the between-class differences in variance-covariance matrix structure considered and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	Our supplementary materials describe our choice of structure for the individual-level random effects, as well as the predictions under a model with a more extensive individual-level random effects structure (including random effects for the spline terms).
7. Are alternative shape/functional forms of the trajectories described?	Yes	We believe cubic splines with 3 df provide sufficient flexibility to capture the underlying functional form of the longitudinal trajectories. In addition, the goodness of fit plots (observed vs predicted) suggest that this is the case.
8. If covariates have been used, can analyses still be replicated?	NA	The data for this study is not publically available.
9. Is information reported about the number of random start values and final iterations included?	Yes	In the Supplementary Materials.
10. Are the model comparison (and selection) tools described from a statistical perspective?	Yes	A subsection is contained in the Methods section of the manuscript.

11. Are the total number of fitted models reported, including a one-class solution?	Yes	Table 1 in the manuscript presents the different models that were considered. Note that the one-class solution is not appropriate for answering the research question in this study, since the one-class joint model corresponds to the assumption of no association between the longitudinal BMI trajectories and death/transplant event rates.
12. Are the number of cases per class reported for each model (absolute sample size, or proportion)?	Yes	Table 1 in the manuscript.
13. If classification of cases in a trajectory is the goal, is entropy reported?	Yes	Relative entropy is shown in Table 1. In addition, the mean posterior probabilities of class membership, stratified by class membership are presented in eTable 1.
14a. Is a plot included with the estimated mean trajectories of the final solution?	Yes	Figure 2 in the manuscript.
14b. Are plots included with the estimated mean trajectories for each model?	Yes	We provide plots of the six-class model, and an alternative model specification that includes additional individual-level random effects. It is infeasible to include plots of every model in our manuscript or supplementary.
14c. Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?	Yes	It is infeasible for us to plot all observed trajectories for the given sample size. But, we have provided plots of the mean predicted and mean observed BMI values across follow up; this answers a slightly different but nonetheless related question about goodness of fit.
15. Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?	Yes	Included in a separate .txt document in the supplementary materials.
16. Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?	Yes	Example code is provided in the Supplementary Materials. The data is not publically available.

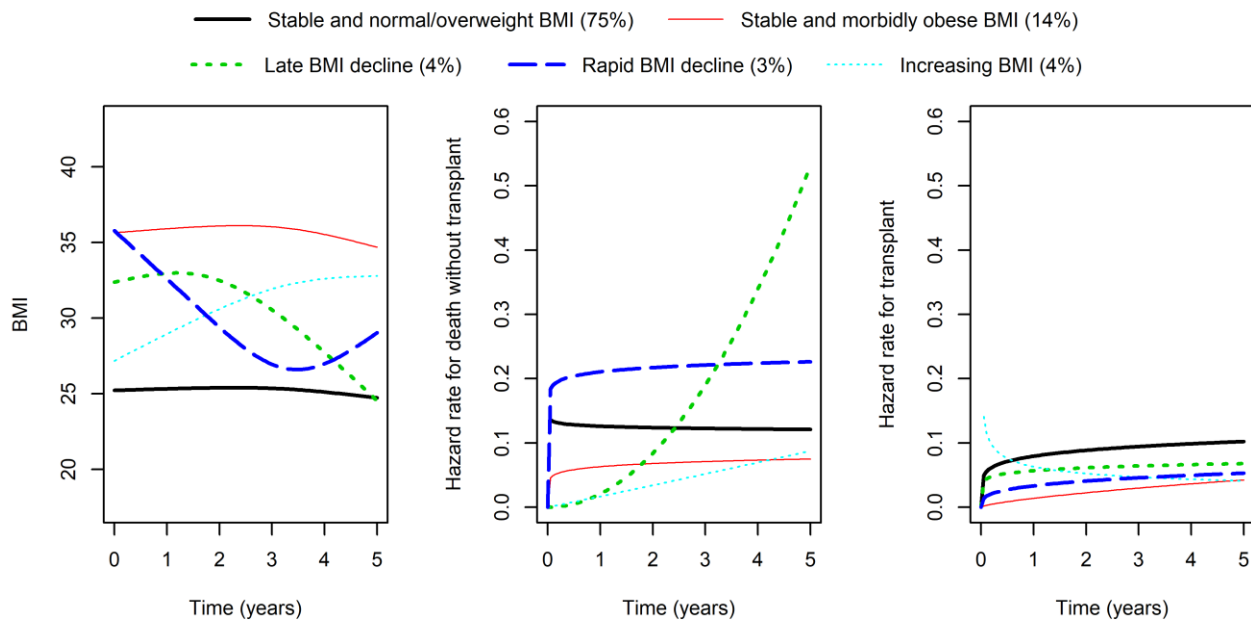
eFigure 1. Flowchart showing numbers of patients excluded from the analysis.



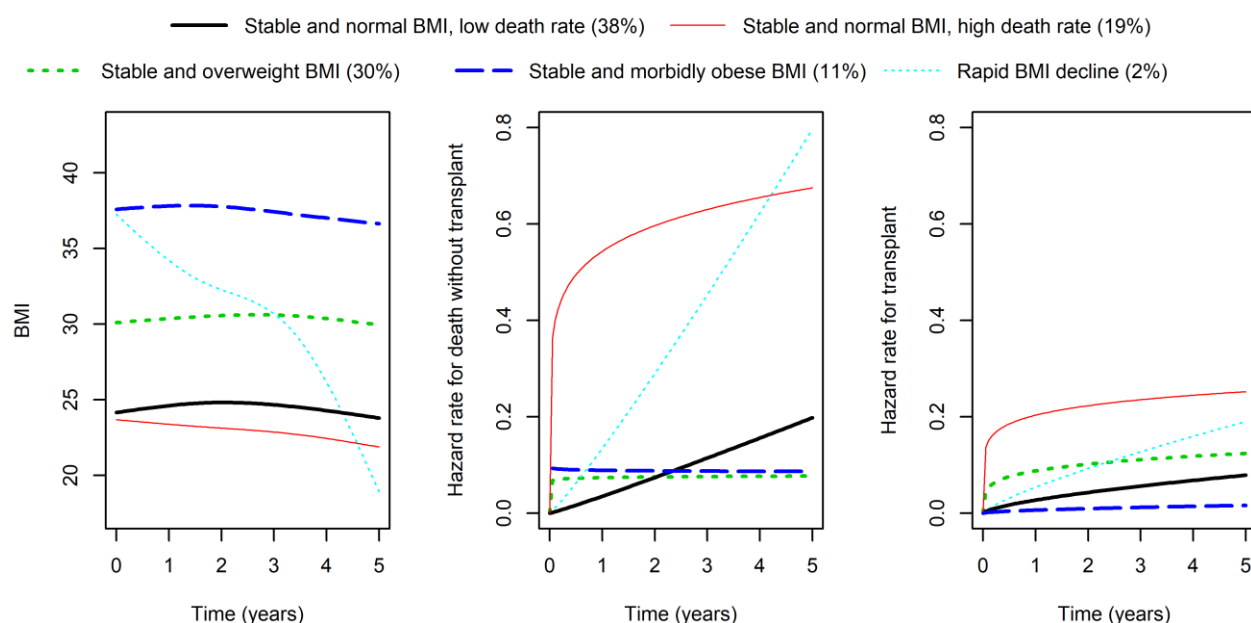
eFigure 2. Observed BMI trajectories for a random sample of 25 patients.



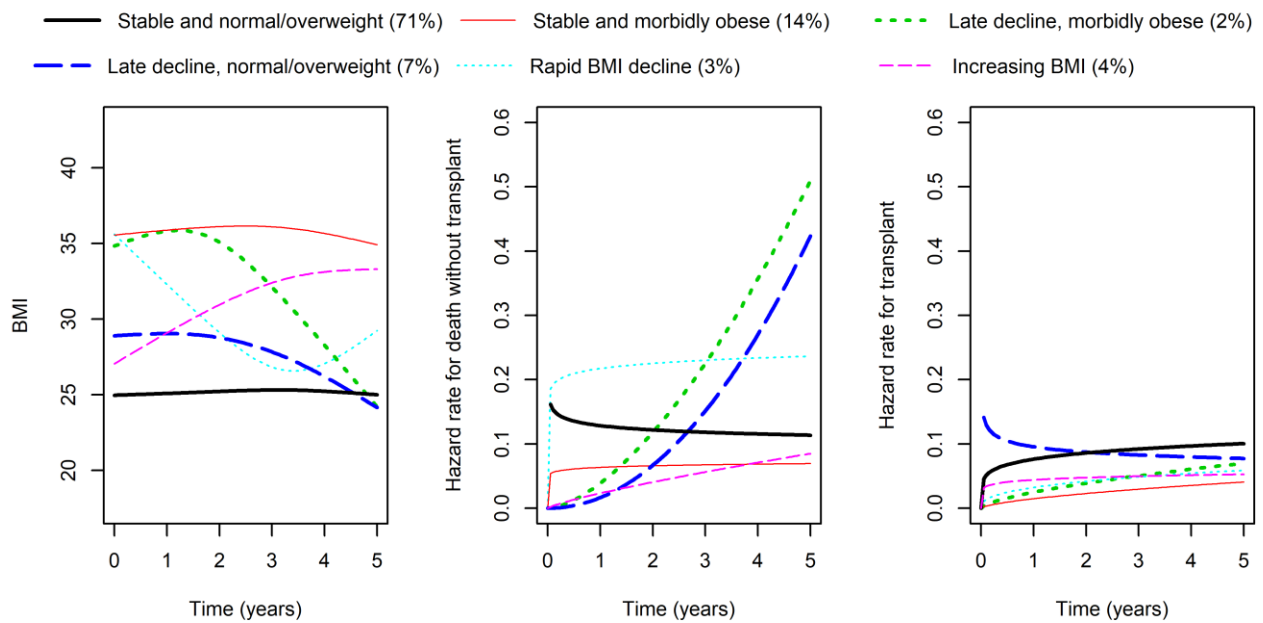
eFigure 3. Predicted longitudinal BMI trajectories (left panel) and cause-specific hazard functions for death without transplant (middle panel) and transplant (right panel) for the alternative solution for the five-class model found using random initial values. The BMI predictions are on average (since no covariates were included in the BMI submodel), whilst the event outcome predictions are for a Caucasian male, aged ≤ 50 years, initiating RRT between 2005-09 with diabetic nephropathy, cerebrovascular disease and coronary artery disease.



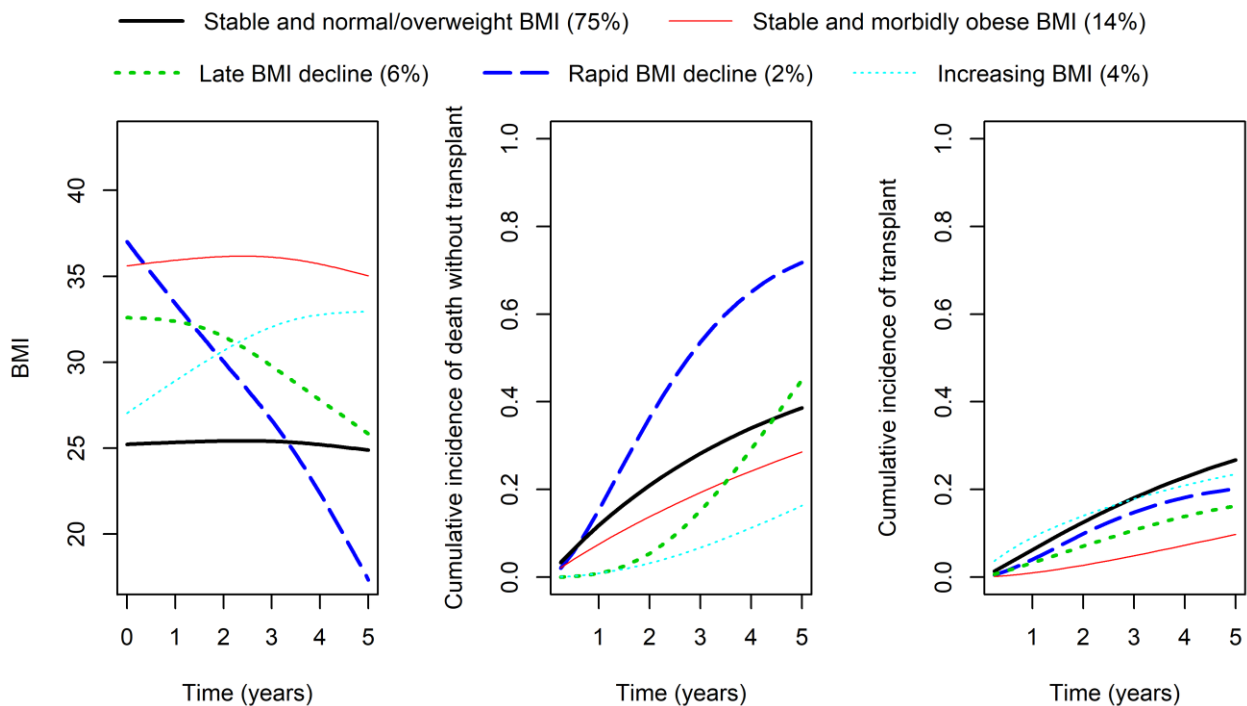
eFigure 4. Predicted longitudinal BMI trajectories (left panel) and cause-specific hazard functions for death without transplant (middle panel) and transplant (right panel) for the five-class joint model after **including individual-level random effects for the cubic splines** (for the longitudinal BMI trajectories). The BMI predictions are on average (since no covariates were included in the BMI submodel), whilst the event outcome predictions are for a Caucasian male, aged ≤ 50 years, initiating RRT between 2005-09 with diabetic nephropathy, cerebrovascular disease and coronary artery disease.



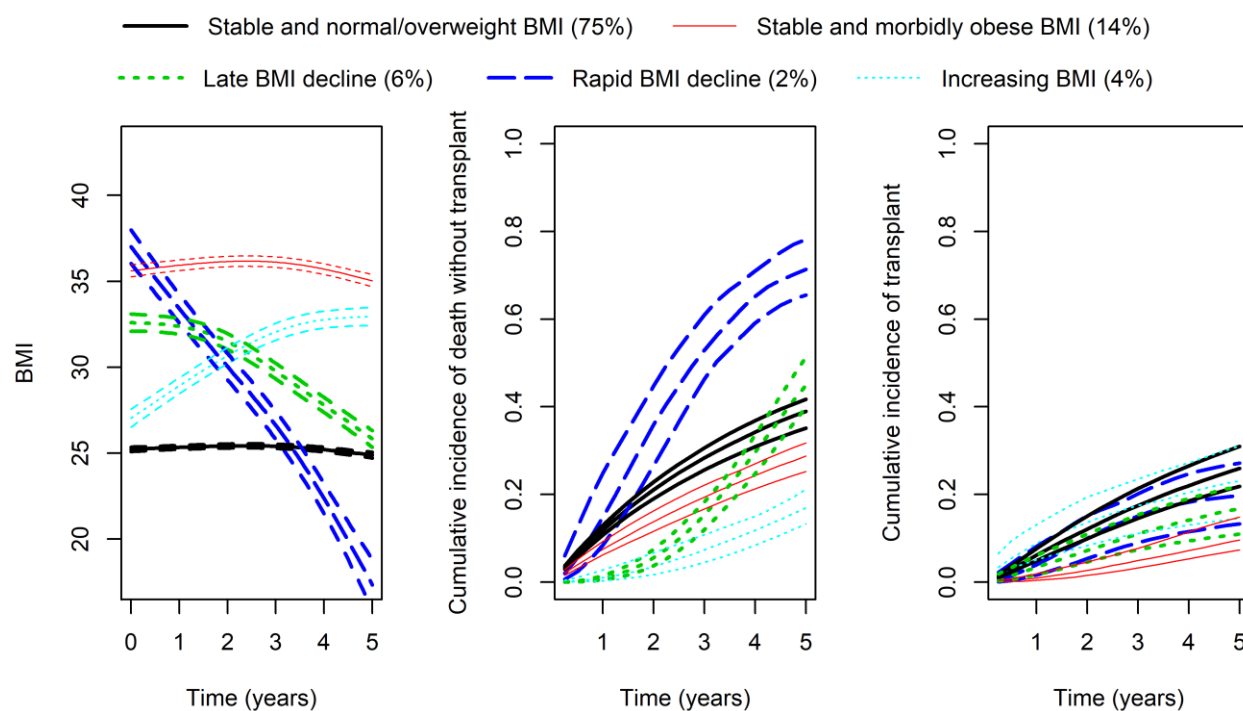
eFigure 5. Predicted longitudinal BMI trajectories (left panel) and cause-specific hazard functions for death without transplant (middle panel) and transplant (right panel) from the **six-class** model. The BMI predictions are on average (since no covariates were included in the BMI submodel), whilst the event outcome predictions are for a Caucasian male, aged ≤ 50 years, initiating RRT between 2005-09 with diabetic nephropathy, cerebrovascular disease and coronary artery disease.



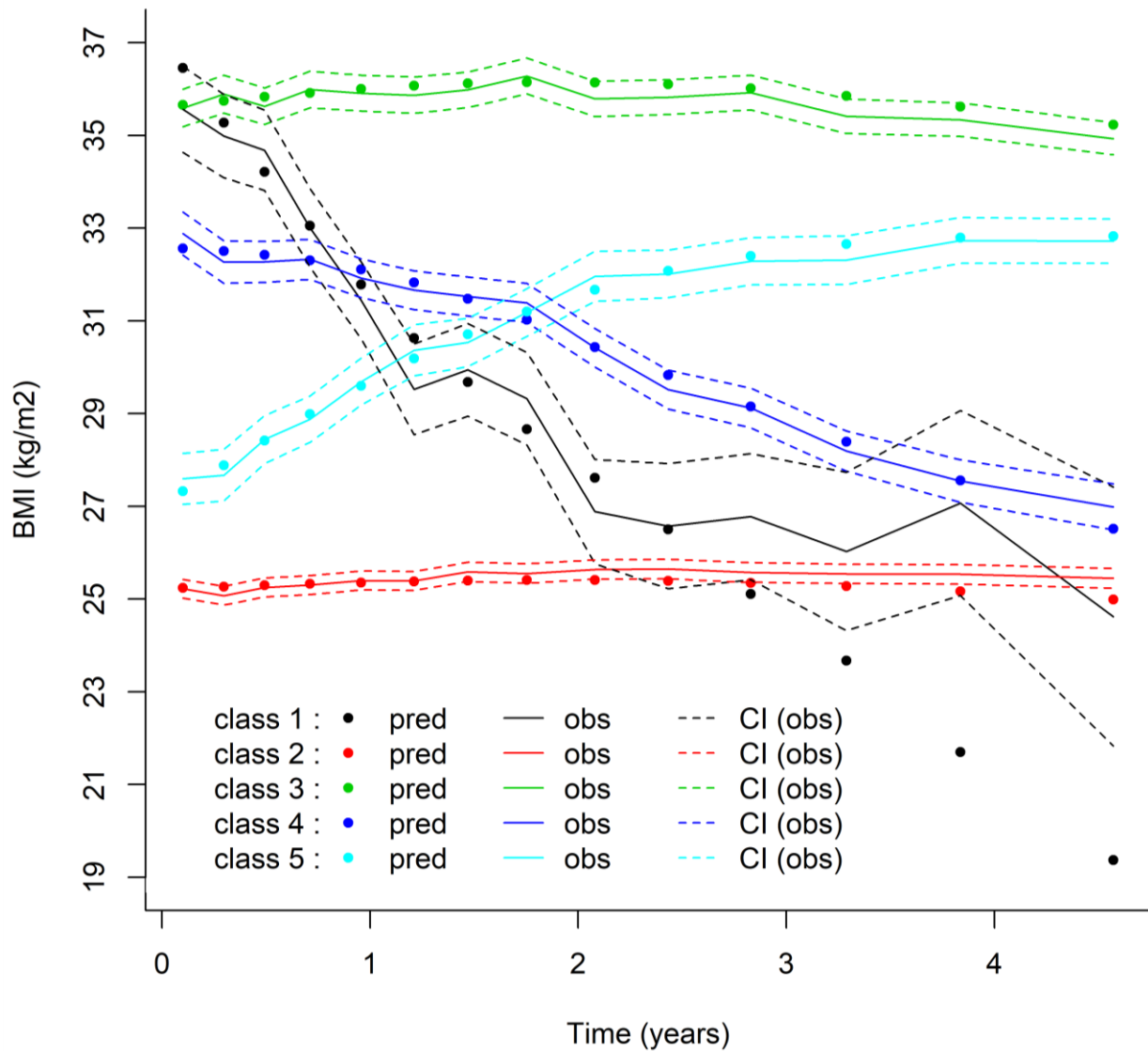
eFigure 6. Predicted longitudinal BMI trajectories (left panel) and **cumulative incidence functions** for death without transplant (middle panel) and transplant (right panel) from the five-class model. The predictions are shown for each of the five possible latent classes. The BMI predictions are on average (since no covariates were included in the BMI submodel), whilst the event outcome predictions are for a Caucasian male, aged ≤ 50 years, initiating RRT between 2005-09 with diabetic nephropathy, cerebrovascular disease and coronary artery disease. These are the cumulative incidence functions for the same covariate profile as for the hazard functions shown in Figure 2 in the main manuscript.



eFigure 7. These are the same figures as described in eFigure 6, but with 95% confidence limits included in the plots.



eFigure 8. Observed and predicted BMI trajectories (**marginal** predictions). The plot shows class-specific BMI trajectories based on weighted means of the observed BMI data (with 95% confidence limits) and weighted means of the **marginal** predictions. The weighting is based on class membership probabilities, whilst the means are taken by splitting the distribution of observed measurement times into 15 quantiles (i.e. 15 “bins”).



eFigure 9. Observed and predicted BMI trajectories (**subject-specific** predictions). The plot shows class-specific BMI trajectories based on weighted means of the observed BMI data (with 95% confidence limits) and weighted means of the **subject-specific** predictions. The weighting is based on class membership probabilities, whilst the means are taken by splitting the distribution of observed measurement times into 15 quantiles (i.e. 15 “bins”).

