Simple parametric analysis for a multi-state model in hospital epidemiology

Maja von Cube, Martin Schumacher, Martin Wolkewitz

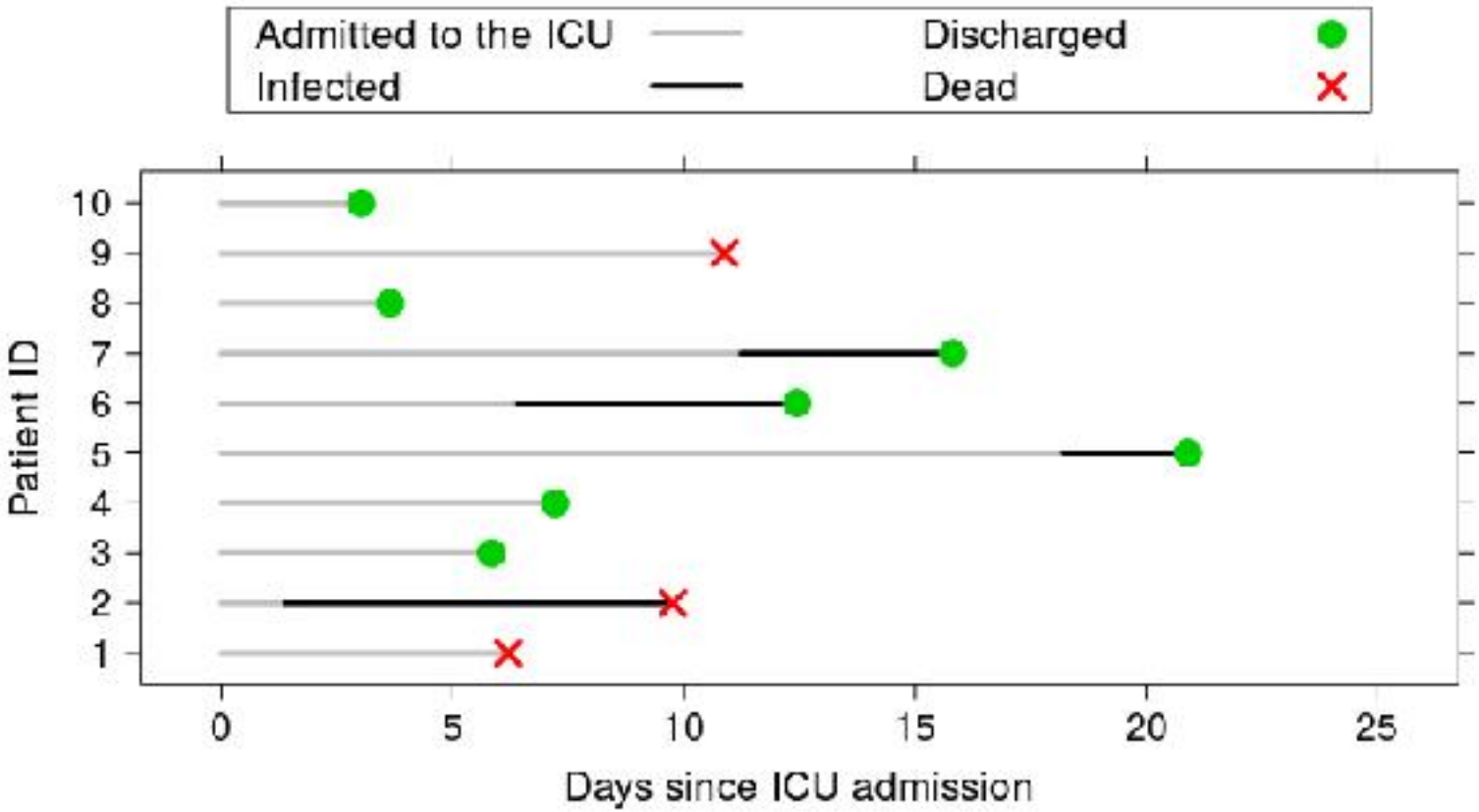
Institute for Medical Biometry and Statistics,
Faculty of Medicine and Medical Center - University of Freiburg, Germany

cube@imbi.uni-freiburg.de

July 13, 2017

- **Key message:**
  - When analysing a new dataset try to start out simple
- **Questions to discuss during the day:**
  - How do others bridge the gap between simplification and misspecification?
  - How do others bridge the gap between correct (but complex) modelling and interpretability?

Motivating example: Sample of the SIR-3 study

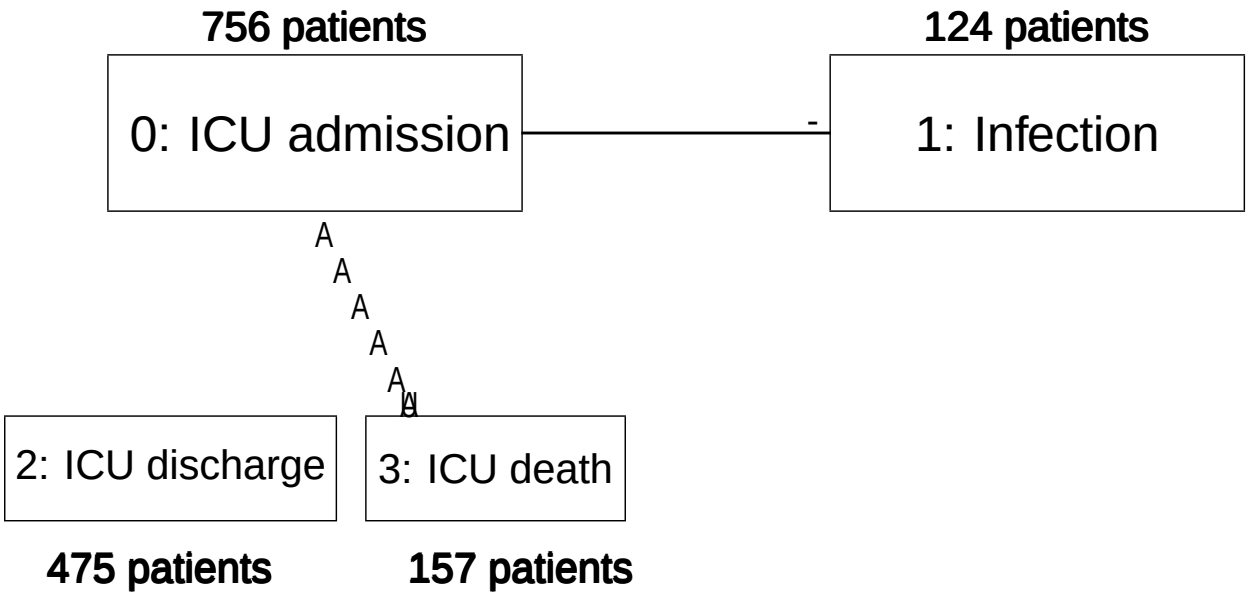


- Observational cohort study from the Charite university hospital in Berlin Germany on hospital-acquired pneumonia in intensive care
- Investigate burden of hospital-acquired infection

Motivating example: Sample of the SIR-3 study

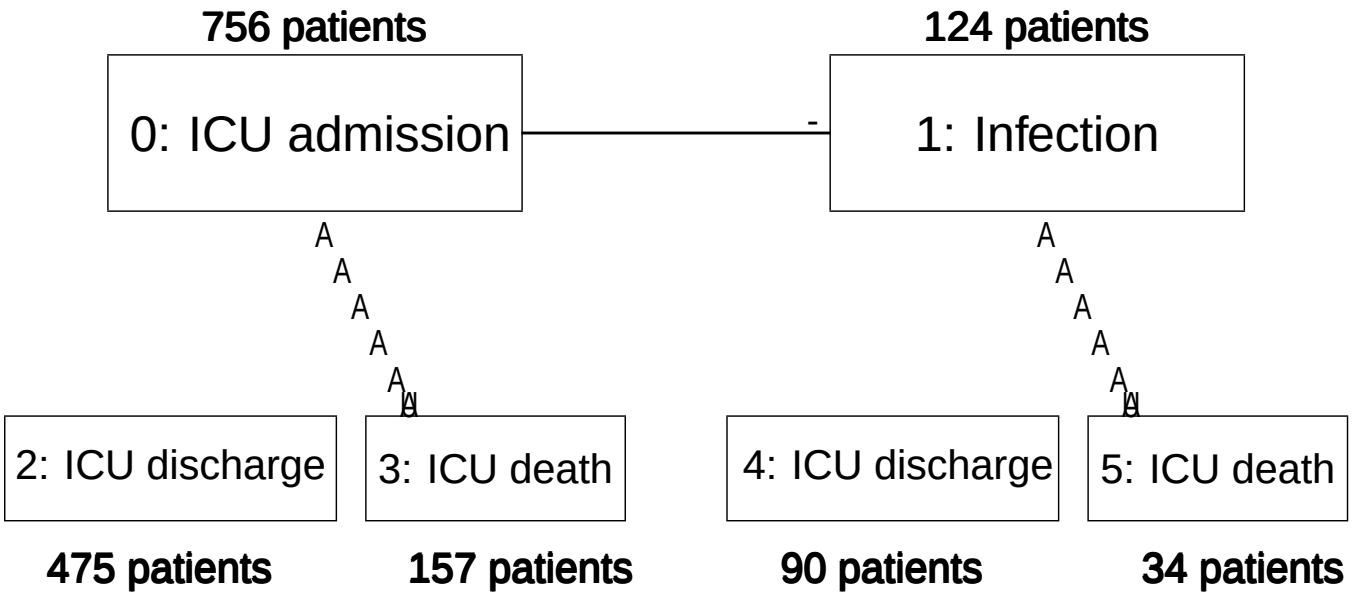
- Commonly used: Kaplan-Meier estimator for classical survival-model

Motivating example: Sample of the SIR-3 study



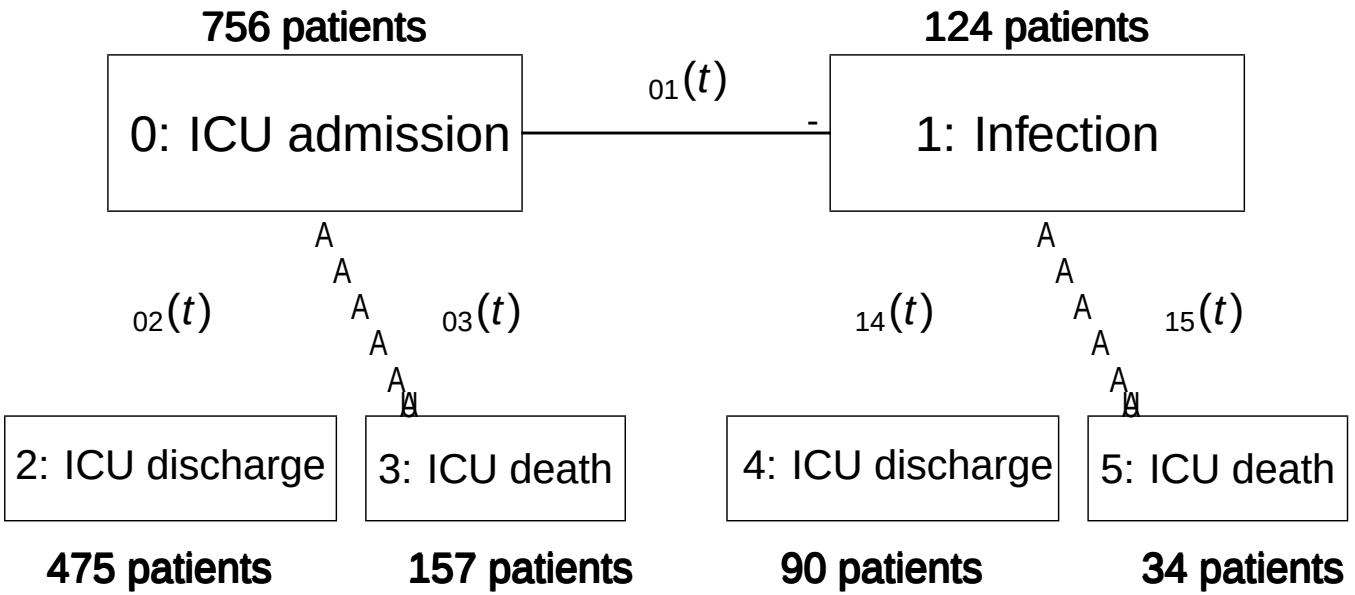
- Commonly used: Kaplan-Meier estimator for classical survival-model
- Source of bias! Kaplan-Meier estimator treats competing events as non-informative

Motivating example: Sample of the SIR-3 study



- Commonly used: Kaplan-Meier estimator for classical survival-model
- Source of bias! Kaplan-Meier estimator treats competing events as non-informative
- Watch out for: Competing risks bias, time-dependent bias

Motivating example: Sample of the SIR-3 study



- Transition specific hazard rates $_{ij}(t); i; j \; 2 \; f0; 1; 2; 3; 4; 5g$
- Transition probabilities $P_{ij}(s; t)$
- Hazard ratios of infection, relative risk of death

Parametric estimation: Assuming constant hazards

- Assume $_{ij}(t) = \ _{ij}$ for all possible transitions
- Simple summary analysis

Occurrence of infection	
Incidence rate per day	01

Parametric estimation: Assuming constant hazards

- Assume $_{ij}(t) = \ _{ij}$ for all possible transitions

- Simple summary analysis

Occurrence of infection	
Incidence rate per day	01
Risk of infection	$_{01} = (\ _{01} + \ _{02} + \ _{03})$

Parametric estimation: Assuming constant hazards

- Assume $_{ij}(t) = \ _{ij}$ for all possible transitions

- Simple summary analysis

Occurrence of infection	
Incidence rate per day	01
Risk of infection	$_{01} = (\ _{01} + \ _{02} + \ _{03})$
Mortality	
Death hazard ratio of infection	$_{15} = \ _{03}$
Discharge hazard ratio of infection	$_{14} = \ _{02}$

Parametric estimation: Assuming constant hazards

- Assume $_{ij}(t) = \ _{ij}$ for all possible transitions

- Simple summary analysis

Occurrence of infection	
Incidence rate per day	$_{01}$
Risk of infection	$_{01} = (\ _{01} + \ _{02} + \ _{03})$
Mortality	
Death hazard ratio of infection	$_{15} = \ _{03}$
Discharge hazard ratio of infection	$_{14} = \ _{02}$
ICU mortality risk of infected patients	$_{15} = (\ _{14} + \ _{15})$
ICU mortality risk of uninfected patients	$_{03} = (\ _{02} + \ _{03})$
ICU mortality risk ratio	$_{15} = \ _{03} \quad (\ _{02} + \ _{03}) = (\ _{14} + \ _{15})$

Parametric estimation: Assuming constant hazards

- Other quantities such as odds of infection/ICU death, extra length of hospital stay
- Even the transition probabilities of the extended illness-death model have closed mathematical forms!
- Example: $P_{ij}(s;t) = P_{ij}(0;\tau), \tau = t \quad s$
  - $P_{00}(s;t) = P_{00}(0;\tau) = \exp(\ (\ _{01} + \ _{02} + \ _{03})\tau)$
  - $P_{03}(s;t) = P_{03}(0;\tau) = \frac{_{03}}{_{01} + _{02} + _{03}}(1 \quad \exp(\ (\ _{01} + \ _{02} + \ _{03})\tau))$

Parametric estimation: Assuming constant hazards

- Estimate constant hazard rates with maximum likelihood estimator:

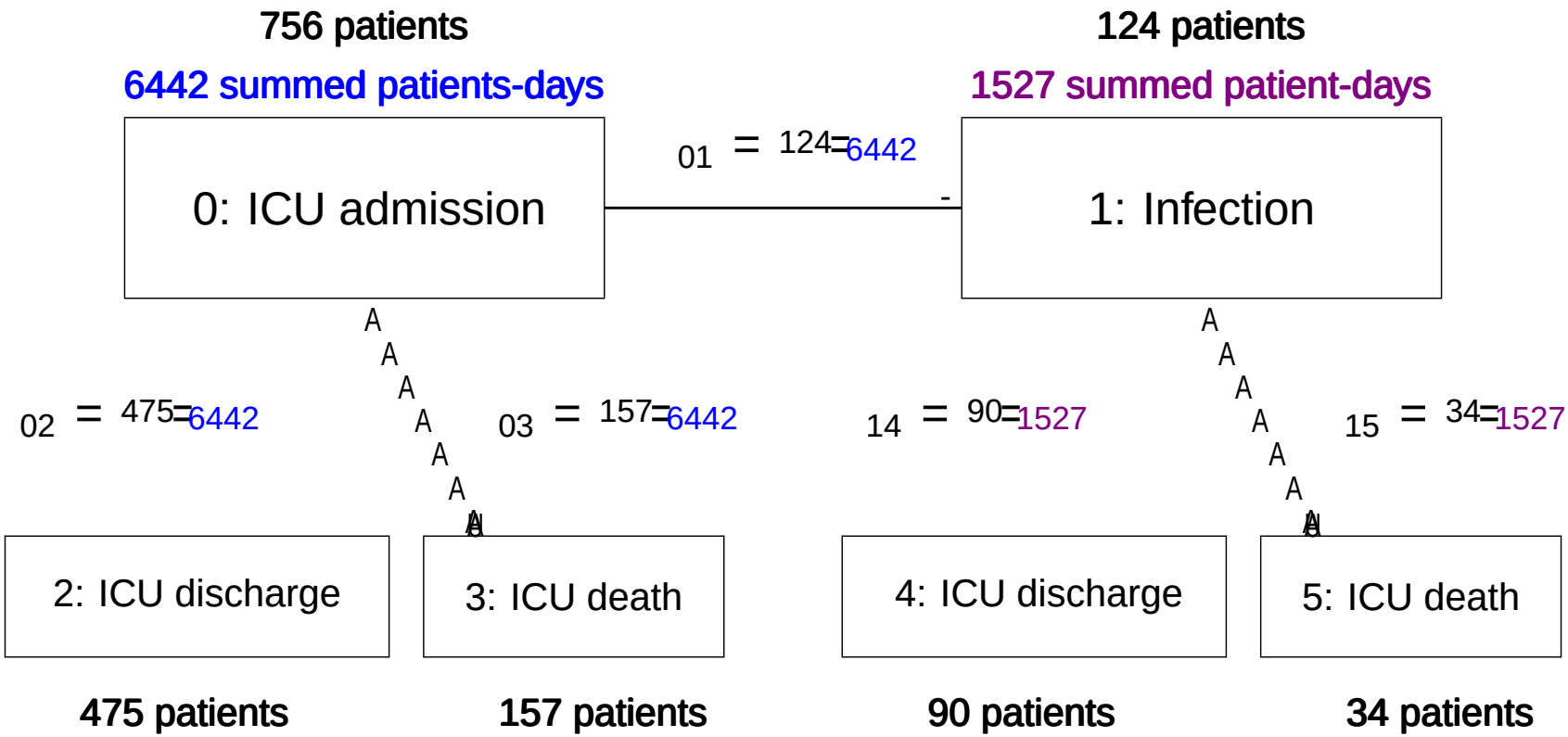
$$\hat{}_{ij} = \frac{\text{Number of } i \;!\; j \text{ transitions}}{\text{Summed patient-days in state } i}:$$

- Plug into the formulas

) **No sophisticated software is necessary for a basic multi-state analysis**

- All you need:
  - Summed patient days in state 0 and 1
  - Number of patients in each state at end of follow-up

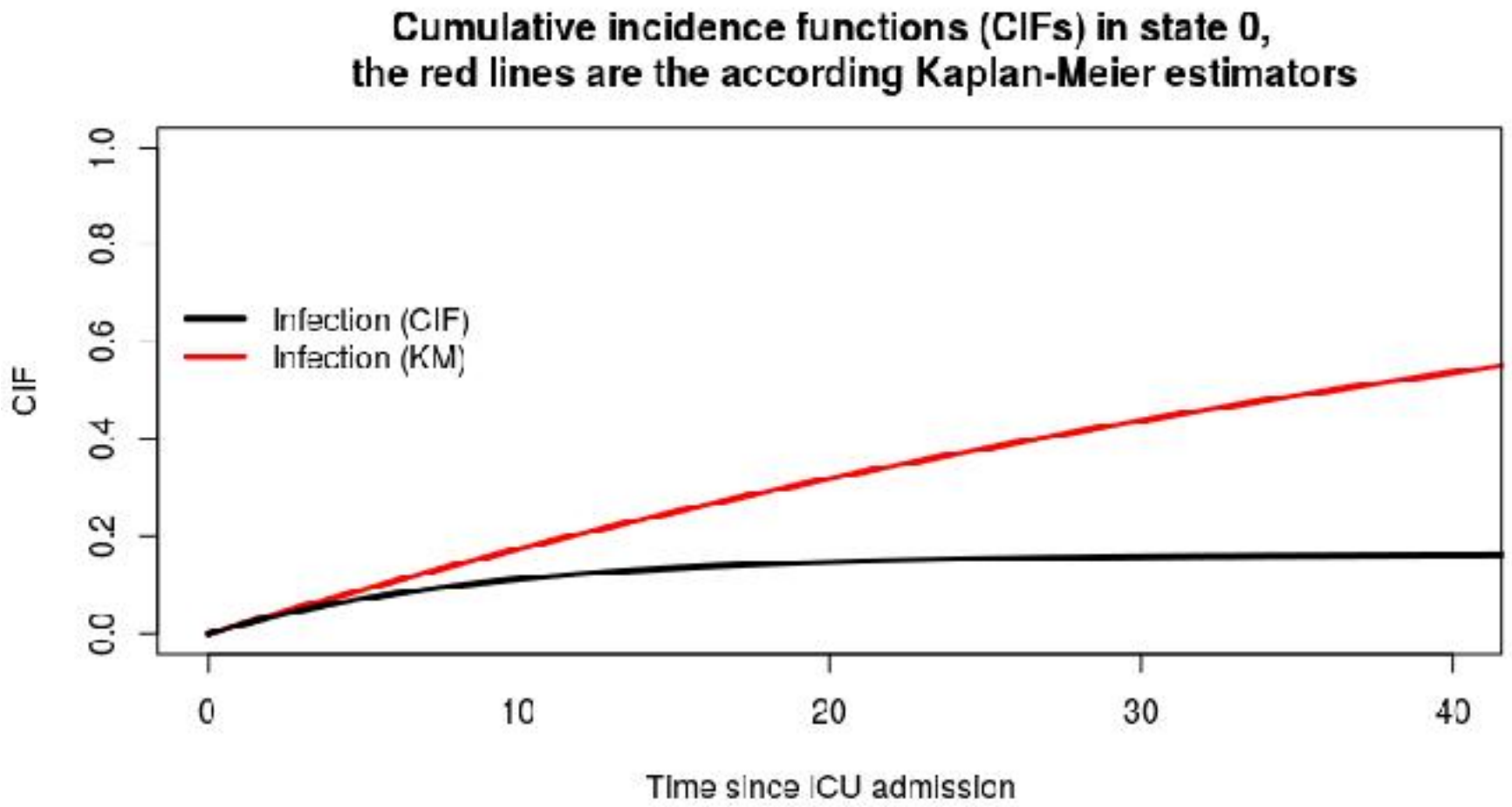
Applied to the sample of the SIR-3 study



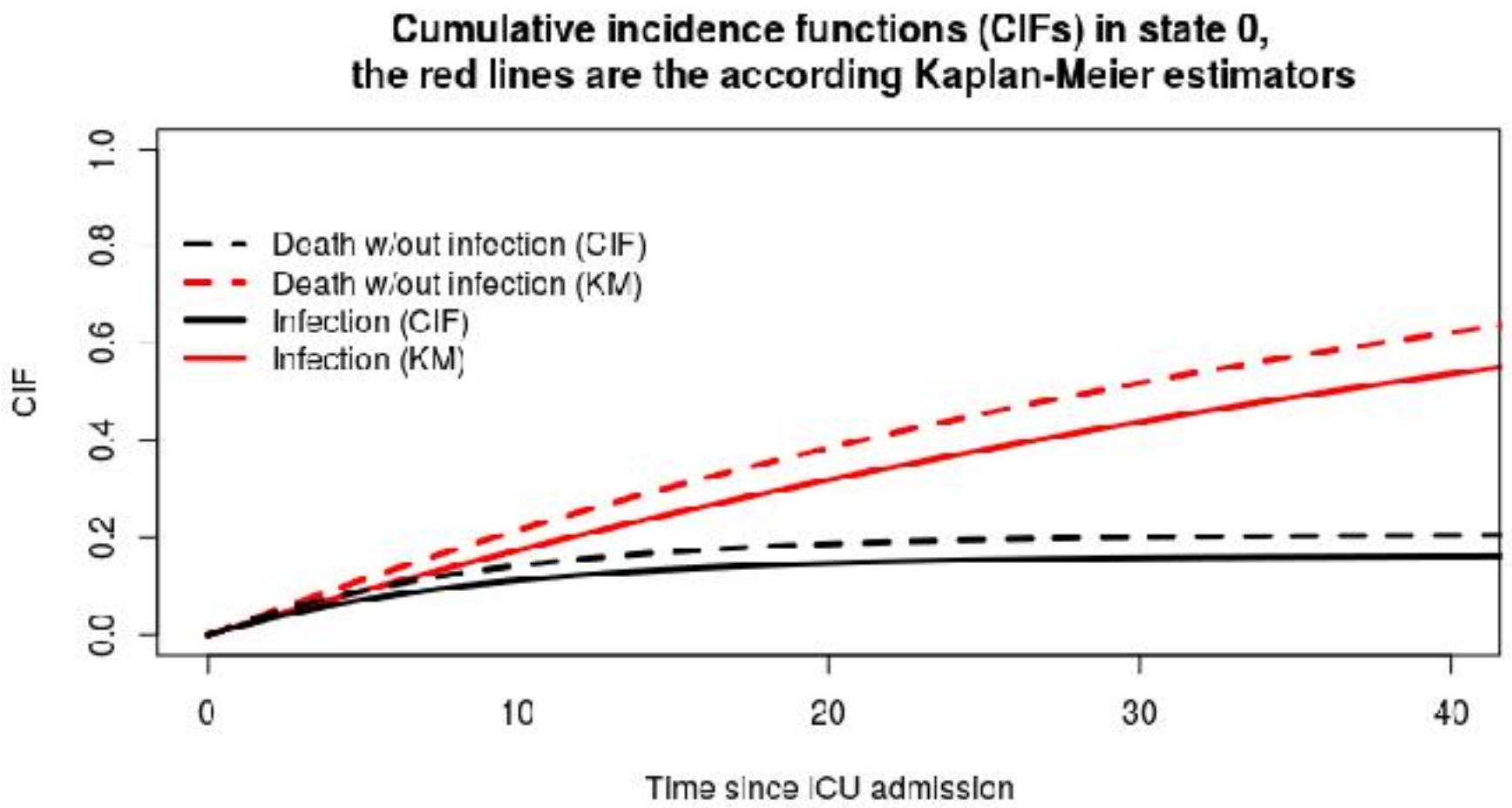
Data example: Summary analysis of sample of SIR-3 study

Occurrence of infection	
Incidence rate per day	0.019
Risk of infection	16.4%
Mortality	
Death hazard ratio of infection	0.91
Discharge hazard ratio of infection	**0.80**
ICU mortality risk of infected patients	27.4%
ICU mortality risk of uninfected patients	24.8%
ICU mortality risk ratio	1.11

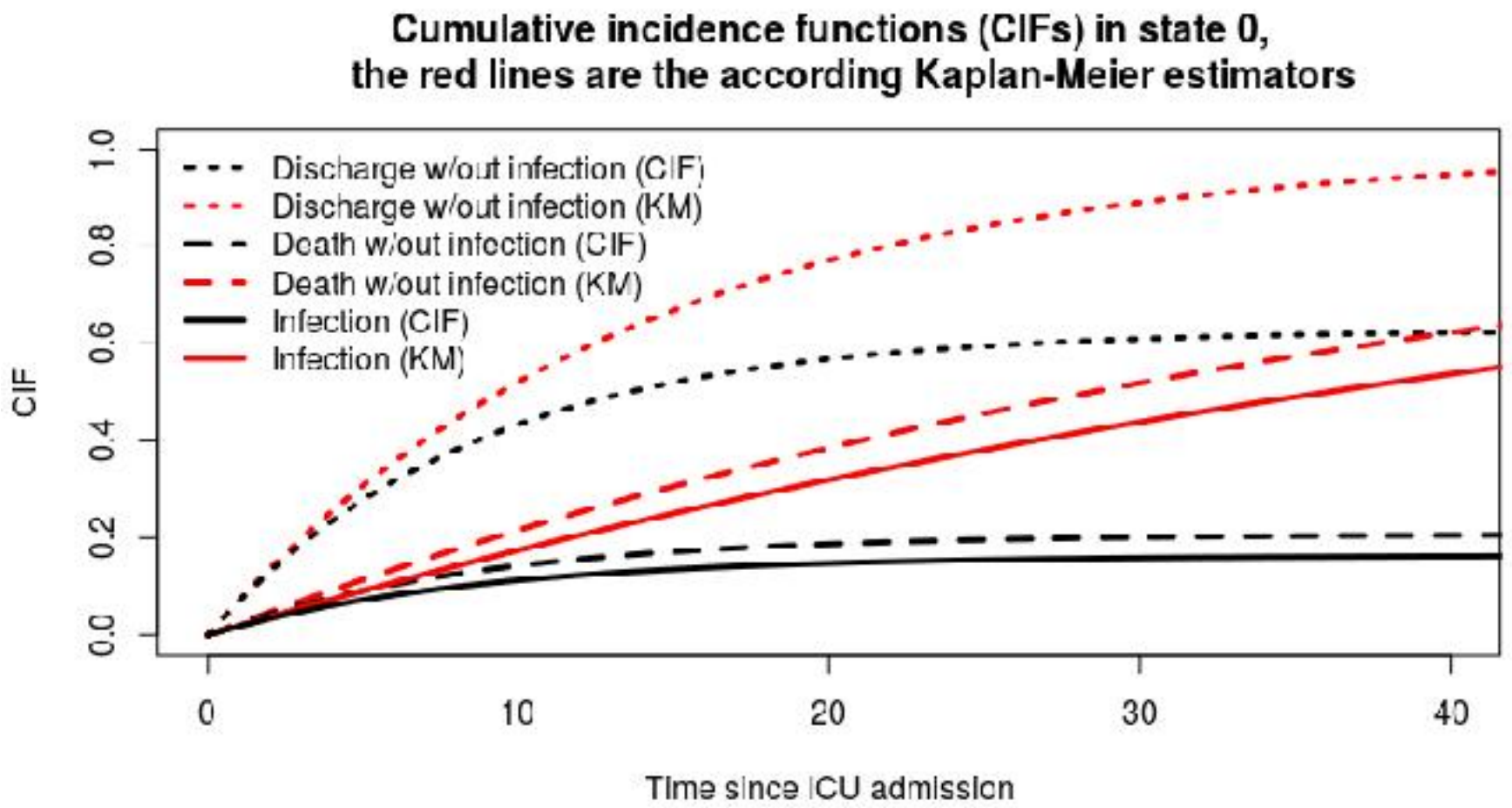
Parametric cumulative incidence functions to demonstrate competing-risks bias



Parametric cumulative incidence functions to demonstrate competing-risks bias



Parametric cumulative incidence functions to demonstrate competing-risks bias



Summary

- Basic multi-state analysis can be performed by assuming constant hazards
  -) Provides first insights into data structure and time-dynamics

Summary

- Basic multi-state analysis can be performed by assuming constant hazards
  -) Provides first insights into data structure and time-dynamics
- Competing risks bias, time-dependent bias are avoided

Summary

- Basic multi-state analysis can be performed by assuming constant hazards
) Provides first insights into data structure and time-dynamics
- Competing risks bias, time-dependent bias are avoided
- Method can be used for reanalysis of published data and meta analysis

) In publications: always provide mean length of stay spent in the transient states and the number of observed events!

Summary

- Basic multi-state analysis can be performed by assuming constant hazards
) Provides first insights into data structure and time-dynamics
- Competing risks bias, time-dependent bias are avoided
- Method can be used for reanalysis of published data and meta analysis

) In publications: always provide mean length of stay spent in the transient states and the number of observed events!
- Method applicable to any other data setting with competing risks and binary exposure

Conclusion

Start simple,

Conclusion

Start simple,
then increase complexity

References

Allignol, A., Schumacher, M., and Beyersmann, J. (2011). Empirical transition matrix of multi-state models: The etm package. *Journal of Statistical Software*, 38(4):1–15.

Andersen, P. K. and Keiding, N. (2002). Multi-state models for event history analysis. *Statistical methods in medical research*, 11(2):91–115.

Beyersmann, J., Allignol, A., and Schumacher, M. (2011). *Competing risks and multistate models with R*. Springer.

de Wreede, L. C., Fiocco, M., and Putter, H. (2011). mstate: An R package for the analysis of competing risks and multi-state models. *Journal of Statistical Software*, 38(7):1–30.

Hougaard, P. (1999). Multi-state models: a review. *Lifetime data analysis*, 5(3):239–264.

Paola Rebora, A. S. and Reilly, M. (2014). *bshazard: Nonparametric Smoothing of the Hazard Function*. R package version 1.0.

Therneau, T. (2014). A package for survival analysis in S. R package version 2.37-4. *See http://CRAN. R-project. org/package= survival*.

von Cube, M., Schumacher, M., and Wolkewitz, M. (2017). Simple parametric analysis for a multi-state model in hospital epidemiology. *Submitted*.

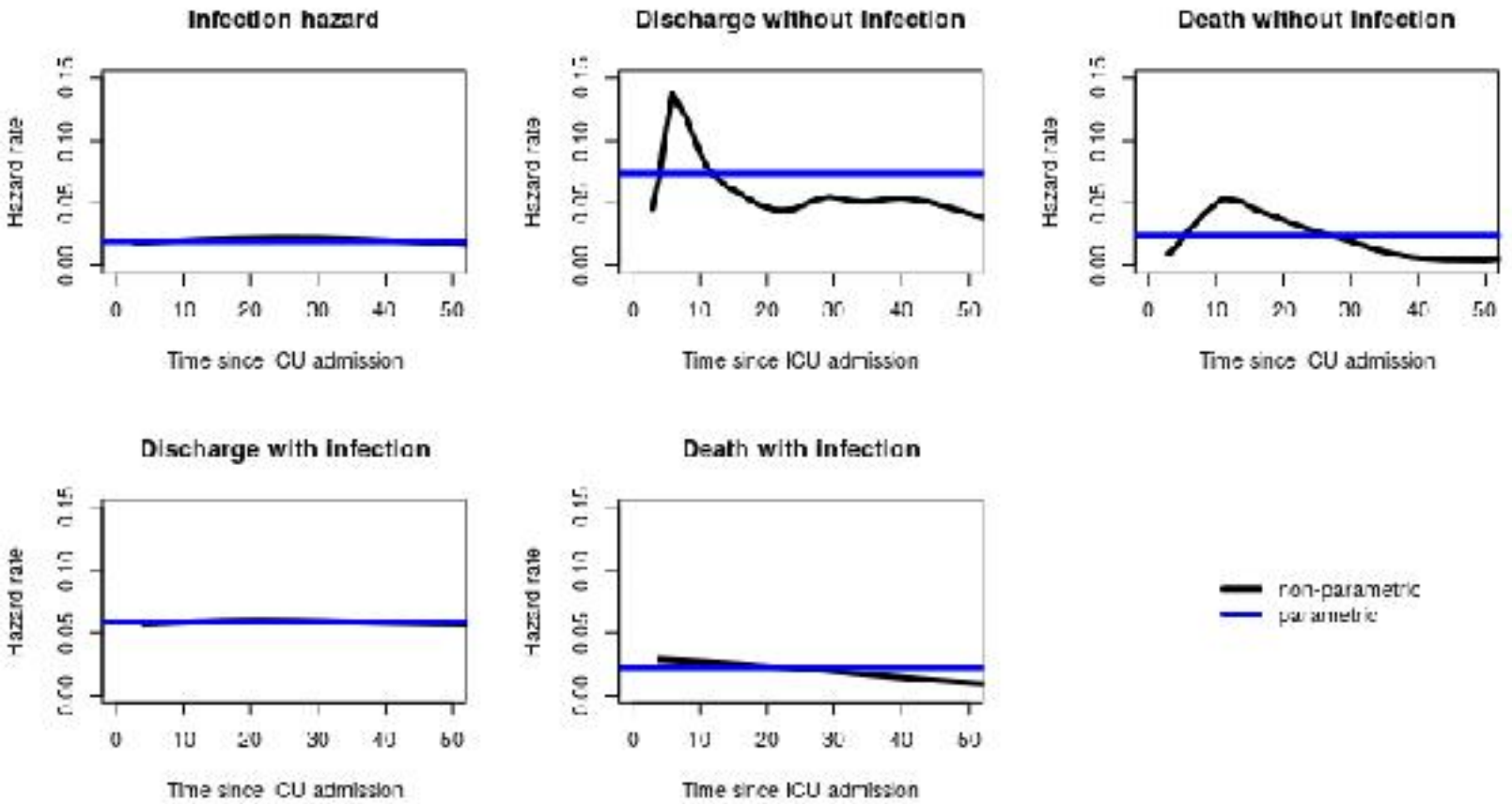
Wolkewitz, M., von Cube, M., and Schumacher, M. (2017). Multi-state modelling to analyze nosocomial infection data: an introduction with real-data demonstration. *Infection Control & Hospital Epidemiology*.

Non-parametric estimation in multi-state models

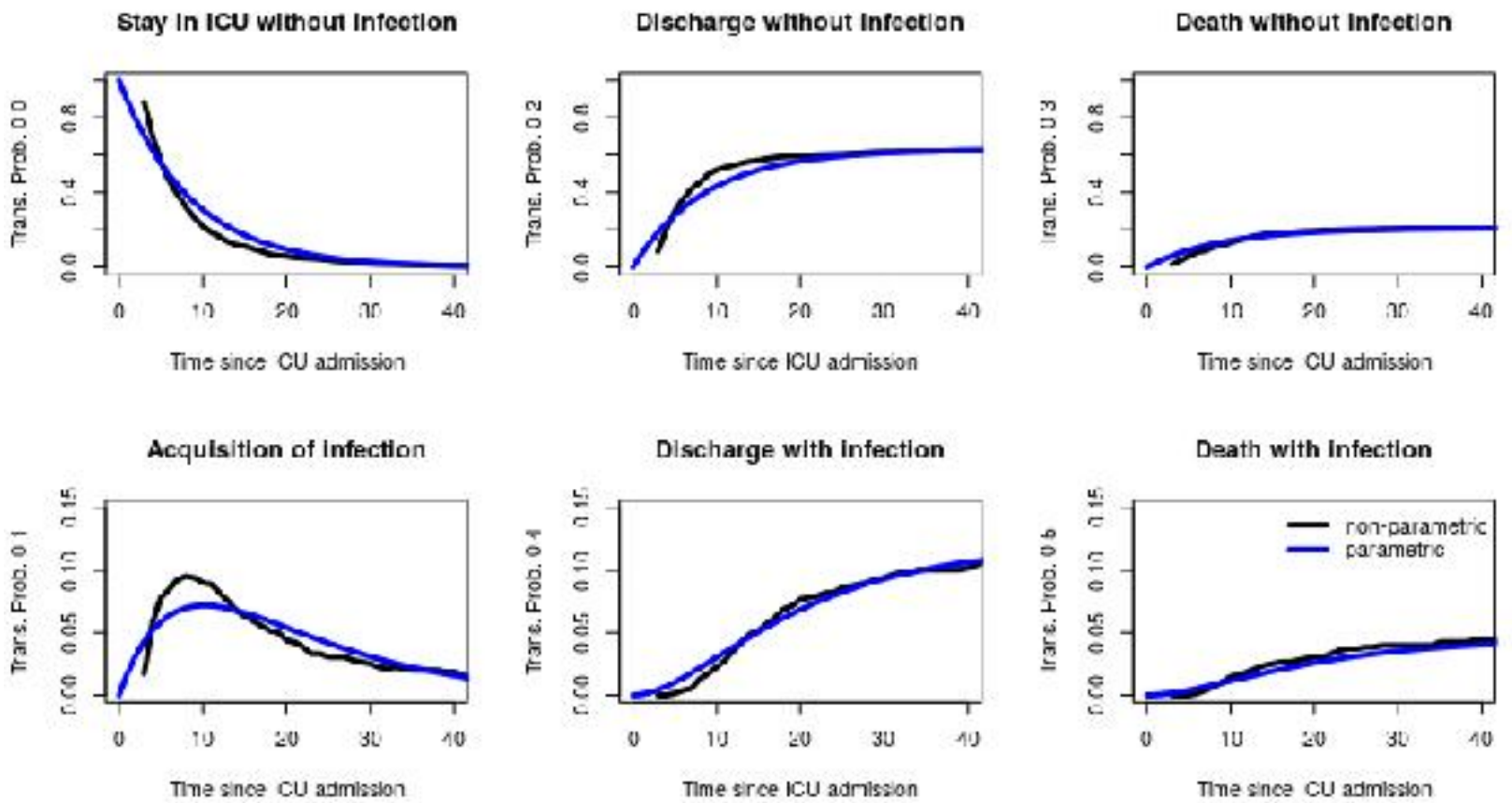
- Nelson-Aalen estimator of cumulative transition-specific hazards
  R-package: mvna
- Aalen-Johansen estimator of transition-probabilities
  R-package: etm, mstate
- Smoothing methods can be used to obtain the transition-specific hazard rates
  R-package: bshazard
- Cox proportional hazards model to estimate hazard ratios
  R-package: survival

) For more details and application see ”Competing risks and Multistate models with R” by Beyersmann et al. (2011)

Misspecification due to false assumption of constant hazards?



Parametric and non-parametric transition probabilities from state 0



Data example: Summary analysis of sample of SIR-3 study

Occurrence of infection / length of stay due to infection	
Incidence rate per day	0.019
Risk of infection	16.4%
Change in length of stay due to infection (cLOS)	1.77 days

- cLOS in naive approach (time-dependent bias):

(?LOS of infected patients) (?LOS of uninfected patients) =

21:0 8:5 = 12:5