
Modeling with Clinical Survival Data

© 2026 TreeAge Software, LLC.

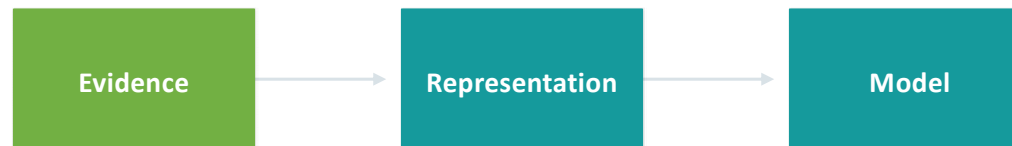
September 2026

Agenda

- Understanding our data - start with the evidence
- Common example: Kaplan-Meier survival data
- Convert survival to risk profile
- Fit distribution to the survival data (in TP or statistical software)
- Hazards and extrapolation
- What does the evidence actually represent?
- Using the clinical data for general survival decline
- Using the clinical data for a single event risk
- Calibrate event-driven model to general survival

Understand our data

- Building models requires that we first understand our data...
- What evidence do we have?
- What does that evidence represent?
- What model approach fits?
- How do we use the data within our model?

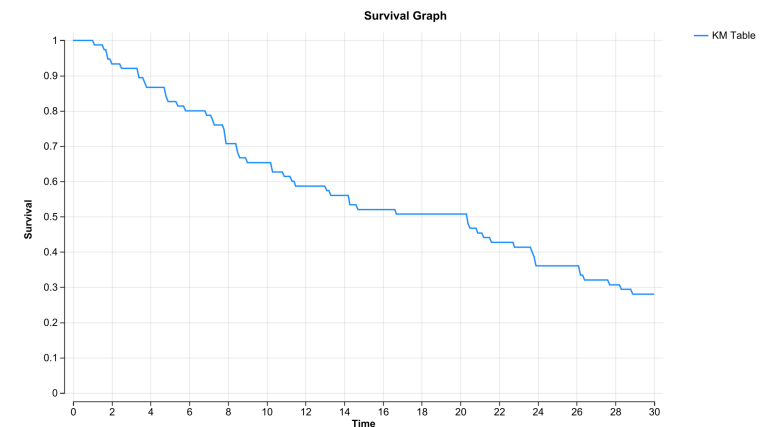


Start with the evidence

- Different evidence sources imply different modeling choices
- Kaplan-Meier curves or tables = Our focus today
- Patient-level data
- Fitted distributions
- Published probabilities and hazard ratios
- Treatment effects and clinical event data

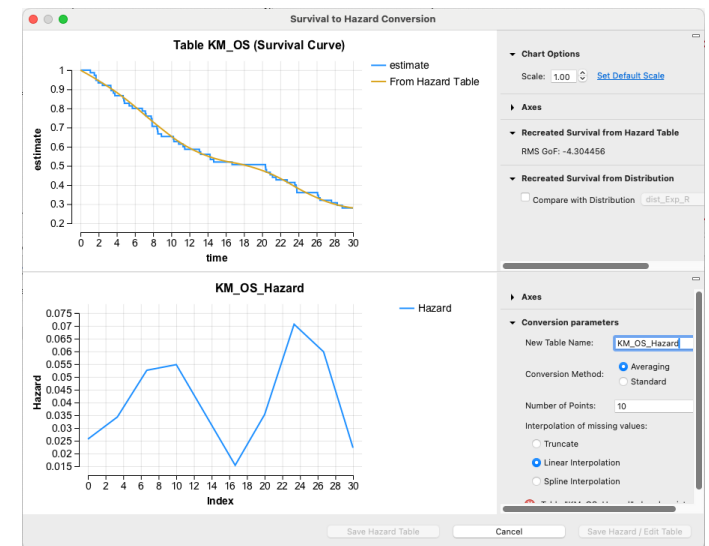
Common example: Kaplan-Meier survival data

- Observed survival often needs representation and extrapolation before it drives a model
- Observed clinical survival represents the impact of cumulative risk over time
 - Implies a changing risk over time
- Requires a proper representation of the health risk for use in your model
- Often requires extrapolation beyond observed period



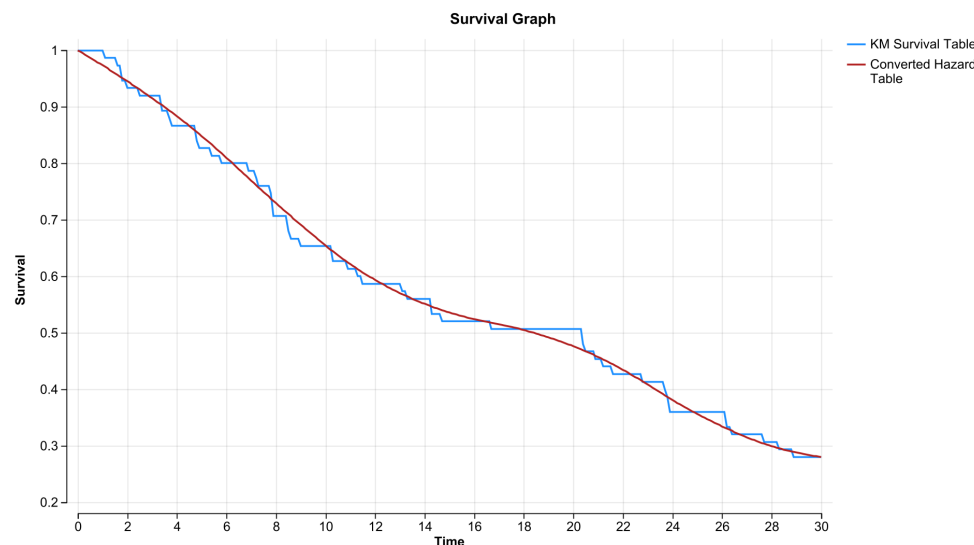
Convert survival to changing risk (hazard)

- TreeAge Pro can convert the survival to hazard for use in all model types
- When the converted hazard is applied over time, it results in the survival which matches well with the observed survival
 - Overlaid in the top graph
 - Survival → Hazard → Resulting Survival



Convert survival to risk profile (hazard)

- Compare survival from the converted hazard to the original observed survival data in the Survival Plot



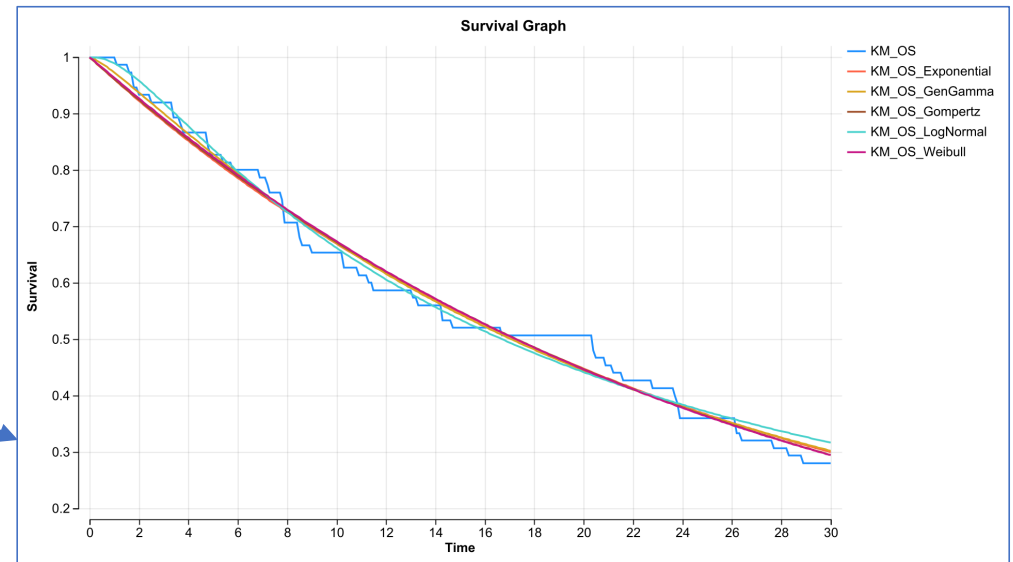
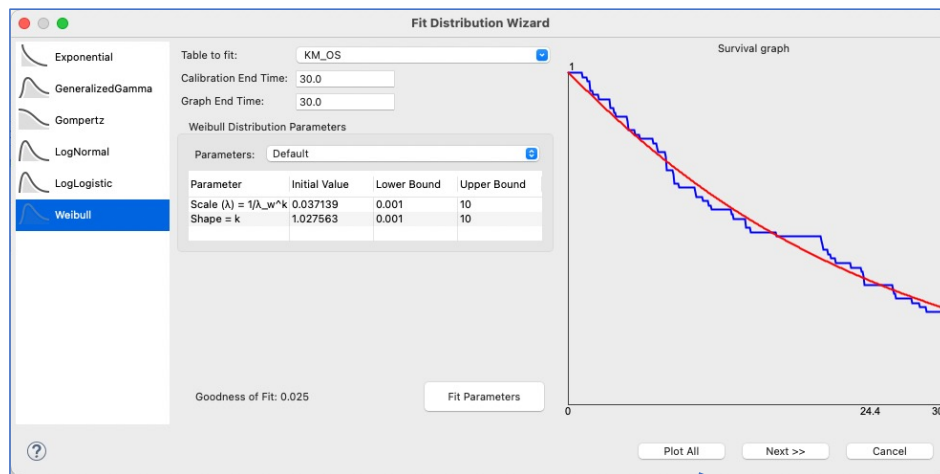
- The converted hazard table provides flexibility to use it in multiple ways to drive disease progression in models

Fit distribution to the survival data

- Modelers frequently fit distributions to observed data, then use the distributions to extrapolate risk into the future
- This can be done using...
 - TreeAge Pro's Fit Distribution function (curve fitting)
 - Statistical software regression analysis – SAS, Stata, R
- Either way, the fitted distribution can be used to drive disease progression in your model

Fit distribution to the survival data in TreeAge Pro

- TreeAge Pro uses curve fitting to generate distributions that match the observed data



In TP: fit distribution

Fit distribution to the survival data externally

- If your distribution is determined by regression analysis, create the TreeAge Pro distribution with the appropriate type and parameters

The screenshot shows the 'Add/Change Distribution' dialog box in TreeAge Pro. The 'General' tab is active, and the 'Weibull' distribution is selected from the left-hand list. The 'Weibull Distribution Parameters' section is highlighted, showing the following values:

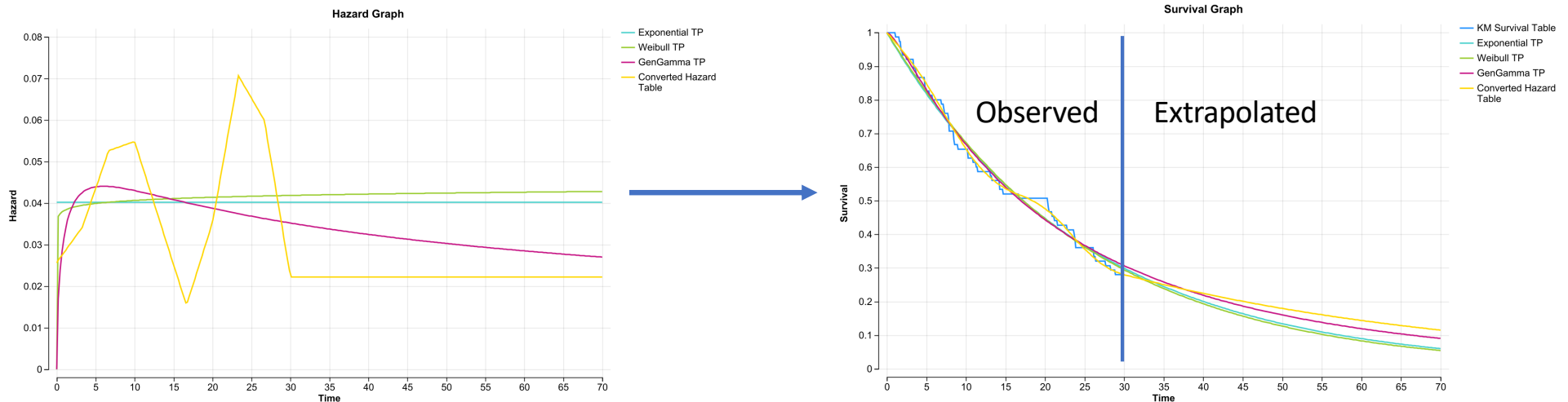
- Parameters: R flexsurvreg
- Scale: 24.027
- Shape: 1.141

Below the parameters, the formula for the scale is given as $\text{Scale } (\lambda) = 1/\lambda_w^k$, where $\lambda_w = 1/((24.027)^{1.141})$. The shape is given as $k = 1.141$. There are also fields for 'Override mean' and 'Minimum time', both currently empty. The 'Sampling Rate' section has three options: 'Resample per EV/group of trials' (selected), 'Resample per Markov stage', and 'Resample per individual trial'. On the right, there are two plots: 'Probability density' (a red curve peaking at 0.032) and 'Cumulative probability' (a blue curve). The x-axis for both plots ranges from 0 to 75.4, with a major tick at 22.9. The y-axis for the probability density plot ranges from 0 to 0.032. The y-axis for the cumulative probability plot ranges from 0 to 1. At the bottom right, there are 'Cancel' and 'OK' buttons, and a 'Range' section with two radio buttons: '[0; 97.5%]' (selected) and '[0; 99%]'.

In TP: open R distribution

Hazards and Extrapolation

- While all the representations of hazard may match the observed period well, the implied hazards impact long-term survival when extrapolated into the future



Choices affect long-term survival estimates

What does the evidence actually represent?

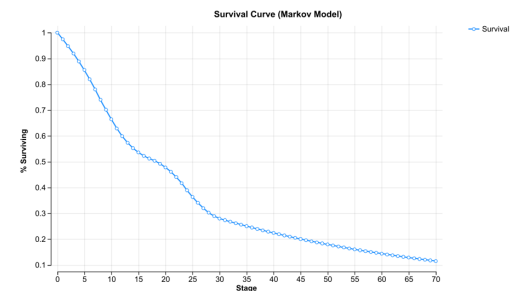
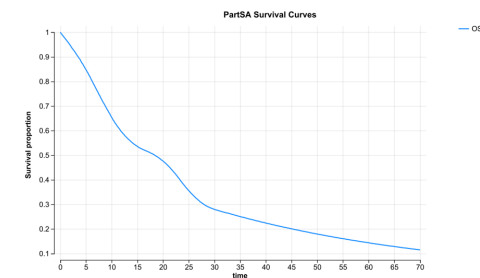
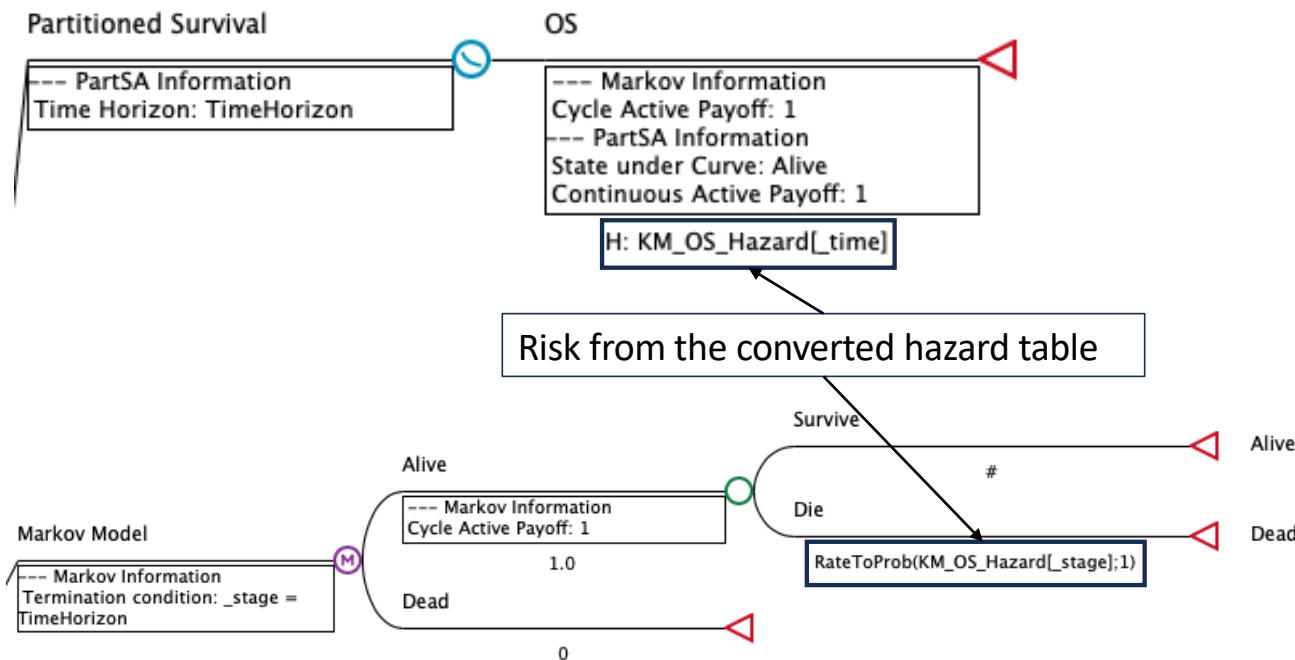
- We need to understand the survival data affects patients before we can include it in a model
- Does the survival data represent a single event?
 - Progression event
 - Death from a single cause
- Does the survival data represent overall cumulative risk from multiple causes?
 - Progression-free survival
 - Overall survival

What does the evidence actually represent?

- If clinical survival curves represent...
 - The application of cumulative risk
 - Fits most easily into a **Partitioned Survival Model**
 - The application of a specific event risk
 - Fits most easily into event-driven model types...
 - **Markov**
 - **Discrete Event**
 - **Hazard Model** (coming Jan 2027)

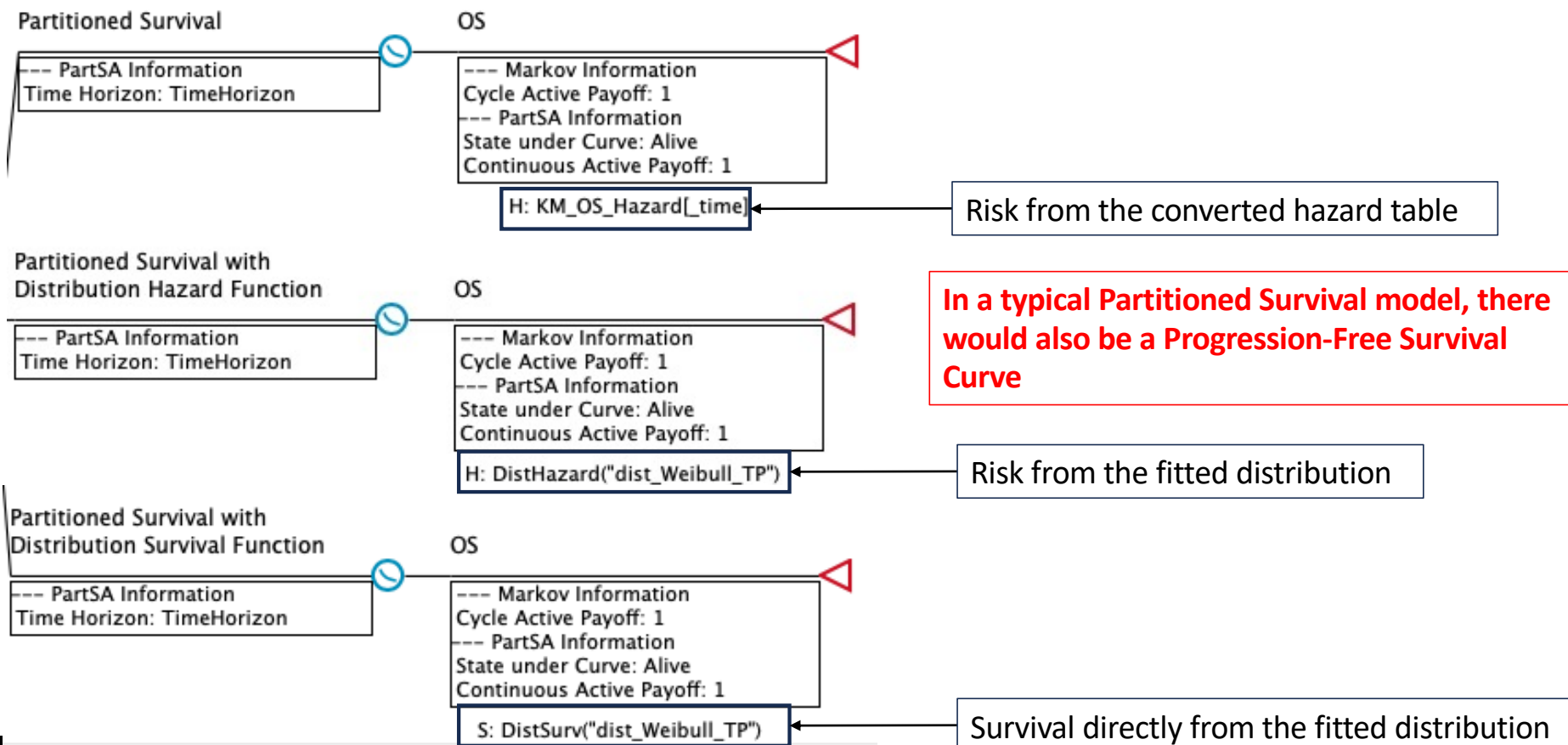
What does the evidence actually represent?

- In a super simple model with only one event, these are equivalent

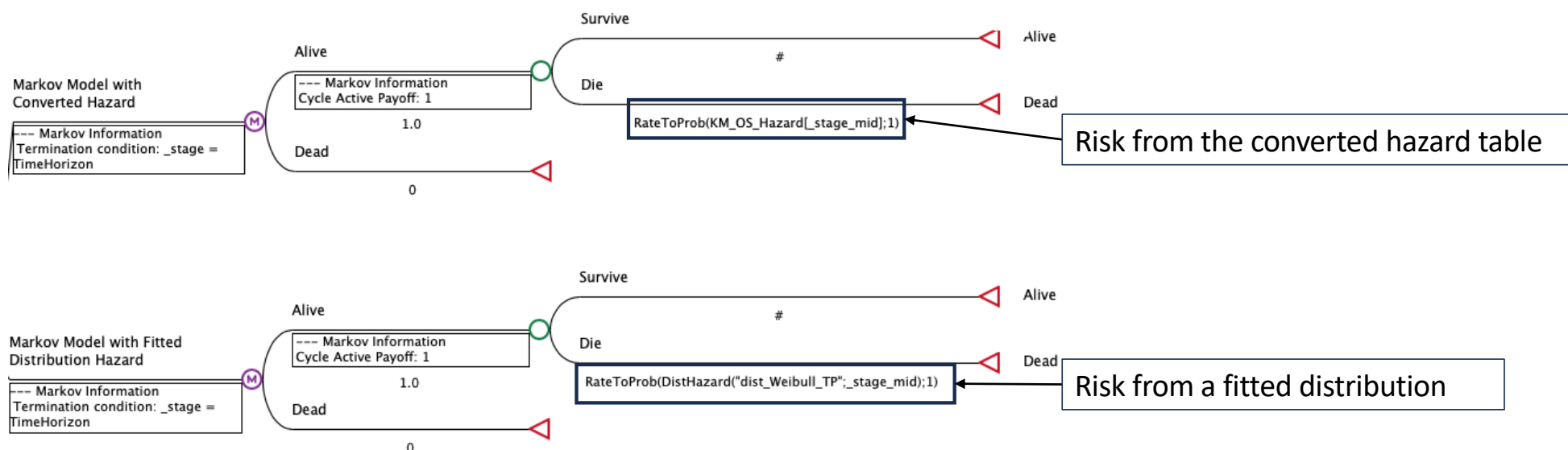


- However, the impact is not equivalent with multiple risks (see later)

Using the clinical data for general survival decline

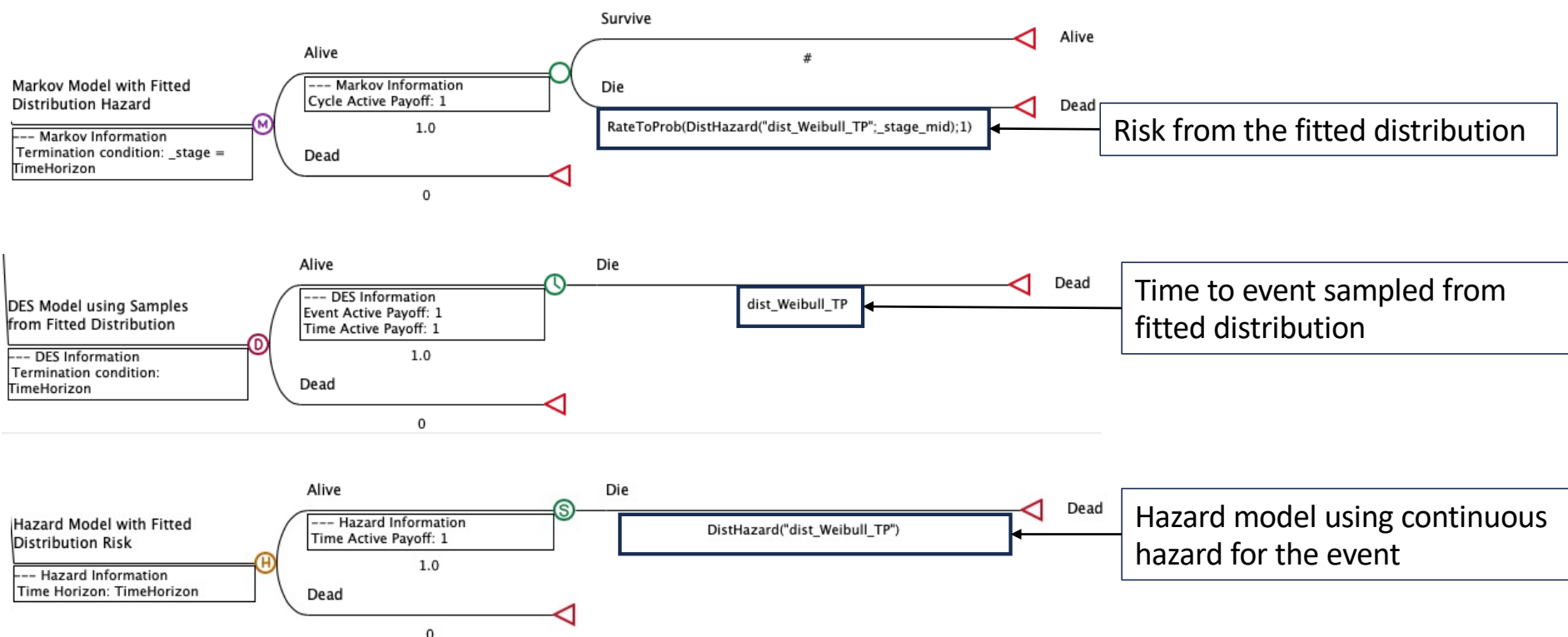


Using the clinical data for a single event risk



In a typical Markov model, there would be more states and competing events

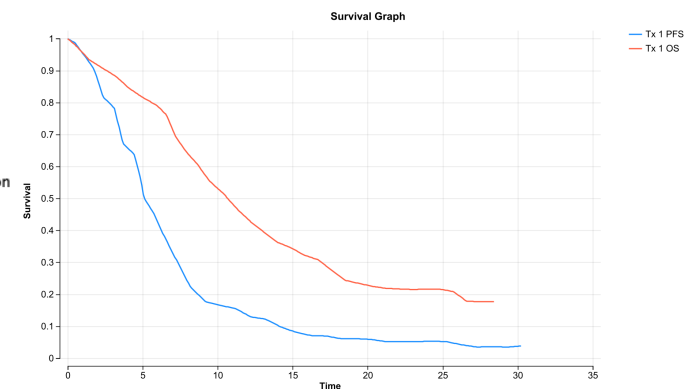
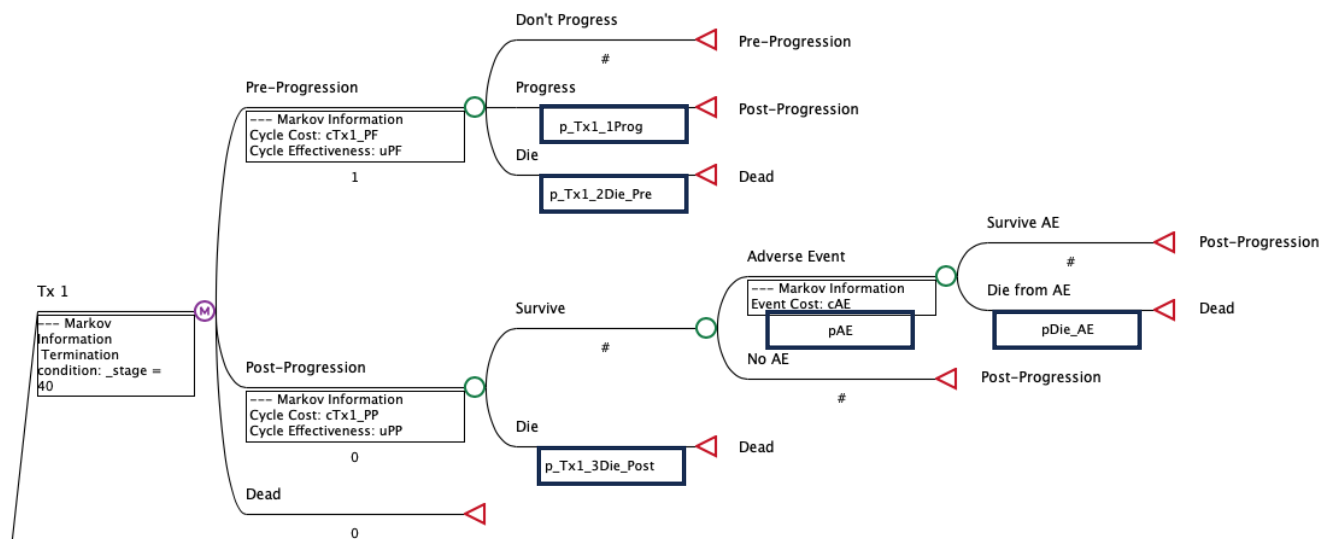
Using the clinical data for a single event risk



Hazard models coming in January 2027

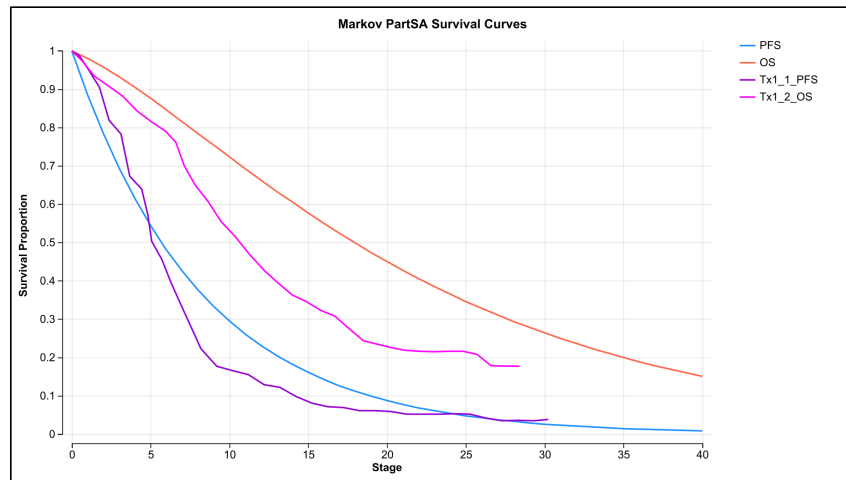
Calibrate event-driven model to general survival

- What if...?
 - You may need an event-driven model because the clinical pathways require key events, but...
 - You may have general survival as source data

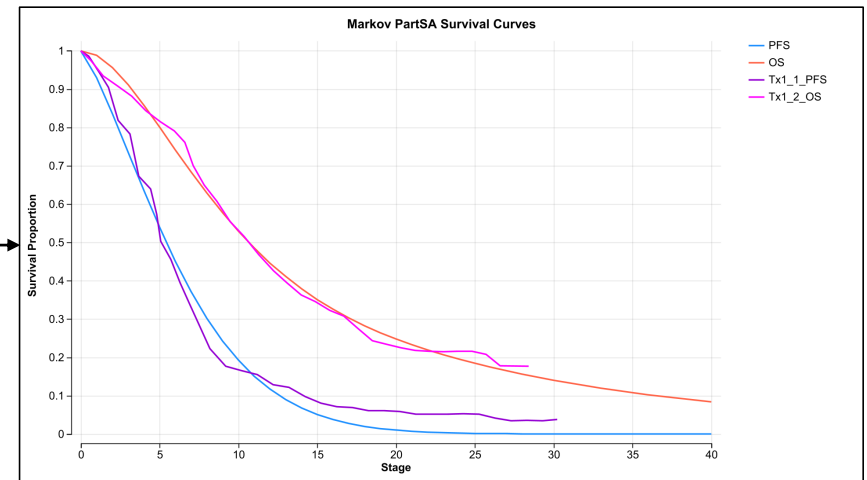


Calibrate event-driven model to general survival

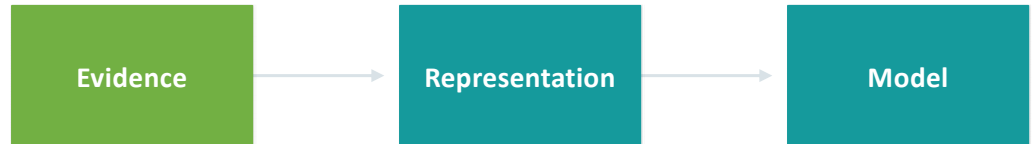
- Calibration can help you match up your model-driven progression to observed clinical data



Calibration



Conclusion



- We want accurate models, so...
- Start with the evidence
- Calculate and visualize how risk behaves over time and its impact on survival into the future
- Define what the risk represents within the model
 - Single cause?
 - Multiple cause?
- Integrate the risk into the model
- TreeAge Pro provides tools to help every step of the way