

How the data is set up: green (first time **years\_2y\_X**) is not missing and  $\geq 1$  year

Example individual A

10.5y	12y	14y	16y	18y	mc1
-------	-----	-----	-----	-----	-----

Example individual B

10.5y	12y	14y	16y	18y	mc1
-------	-----	-----	-----	-----	-----

Possible solution: line up by period number t

Example individual A

10.5y	t=1, 12y	t=2, 16y	t=3, 18y	t=4, mc1	t=5, CENSORED
-------	----------	----------	----------	----------	---------------

NOTE: small issue with this approach occurs if individual A does not experience the outcome at t = 4 such that  $Y_4 = 0$  and they are considered censored at period t = 5 (meaning they are missing exposures at t=5 since this goes past the study end date). In the cleaned data, we will carry-forward the exposure (and time-varying covariates) from period t = 4 and make  $Y_5$  unknown due to the censoring. However, alternatively, we could consider the individual censored at time t = 4.

Example individual B

t=1, 10.5y	t=2, 12y	t=3, 16y	t=4, 18y	t=5, mc1
------------	----------	----------	----------	----------