

# Marginal Structural Model for studying the causal effect between vaso-occlusive crises and occurrence of death or complications in sickle-cell disease patients

## Context

Sickle-cell disease (SCD) is one of the most prevalent inherited diseases. Previous research suggests a relationship between Vaso-Occlusive Crises (VOC) (also known as sickle cell pain crisis) death and complications. However, the time-varying effect of VOC on these outcomes prevents the establishment of a causal relationship using classically adjusted models. The number of VOC in the previous year could instead only imply the occurrence of acute complications and death.

## Objective

This study aimed to assess the effect of the number of VOCs experienced in the previous year on occurrence of death and common disease-related complications using an Inverse Probability Weighted Marginal Structural Model (MSM-IPW).

## Dataset

Prevalent patients with SCD between 2008 and 2018 were identified using ICD-10 diagnosis codes (D57.0, D57.1, D57.2) recorded in NHS England's Hospital Episode Statistics database linked to the Office for National Statistics Mortality Data. The database included information on patient demographics (sex, age, ethnicity), death and complications experienced through ICD-10 codes. The 20 most common complications were studied but only the 5 top complications were presented here.

Group ICD- 10 Codes Used HES analysis to find the complications:

ACS (Acute Chest Syndrome) J18 J22 J12 J13 J14 J15 J16 J17    Cardiomegaly I51.7  
Gall Stones K80 K81 K82 K83    Sepsis A41    Avascular Necrosis M81 M87

The VOC were identified using the D57.0 ICD-10 diagnosis code. The number of VOC experienced in the previous 12 months was defined into three groups, 0 VOCs, 1-2 VOCs and 3+ VOCs. These categories are consistent with a previous study (Platt 1991) which showed that patients experiencing an average of 0, 1-2 and 3+ VOCs each year had different survival trajectories. These time-varying VOC categories were aggregated at a one year time scale to allow maximum follow-up in the dataset.

## Discussion and limitations

The patients with more than 3 VOC during the previous 12 months had an estimated causal risk of death multiplied by 4.50 [2.91;6.97] and of ACS multiplied by 5.51 [3.67;8.27] compared to the patients without VOC during the previous 12 months. These estimations are valid if all hypotheses on the causal model are verified (positivity, consistency and exchangeability), the DAG is consistently defined, and the statistical model are correctly specified. The exchangeability may be a problem in studies with health record administrative dataset as some important factors could be missing or badly measured. For chronic or long-term complications, the estimations of a causal hazard ratios are also more challenging as it is difficult to measure the impact of the number of VOC in the previous 12 months only.

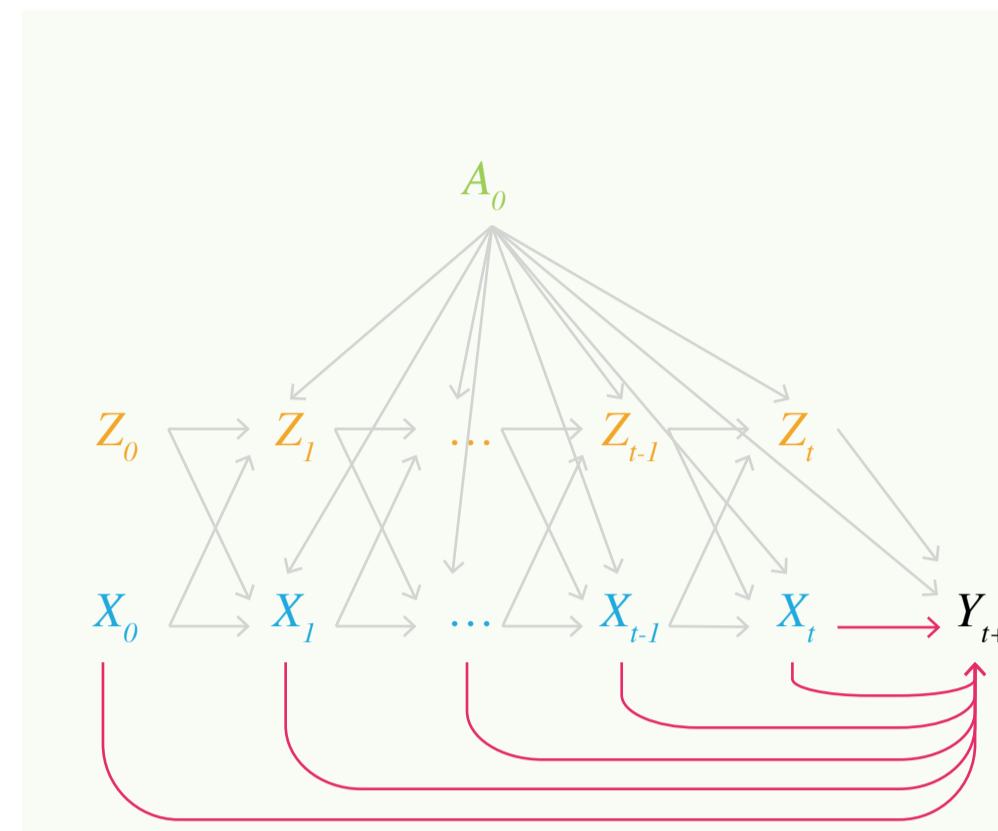
## Conclusion

The IPW analysis is an interesting tool in case of dynamic time-varying exposure. In this illustration having more than one VOC in the previous 12 months was associated with a higher risk of death and complications. Preventing or reducing the number of VOCs experienced each year may significantly reduce the occurrence of death and some common complications.

## Method

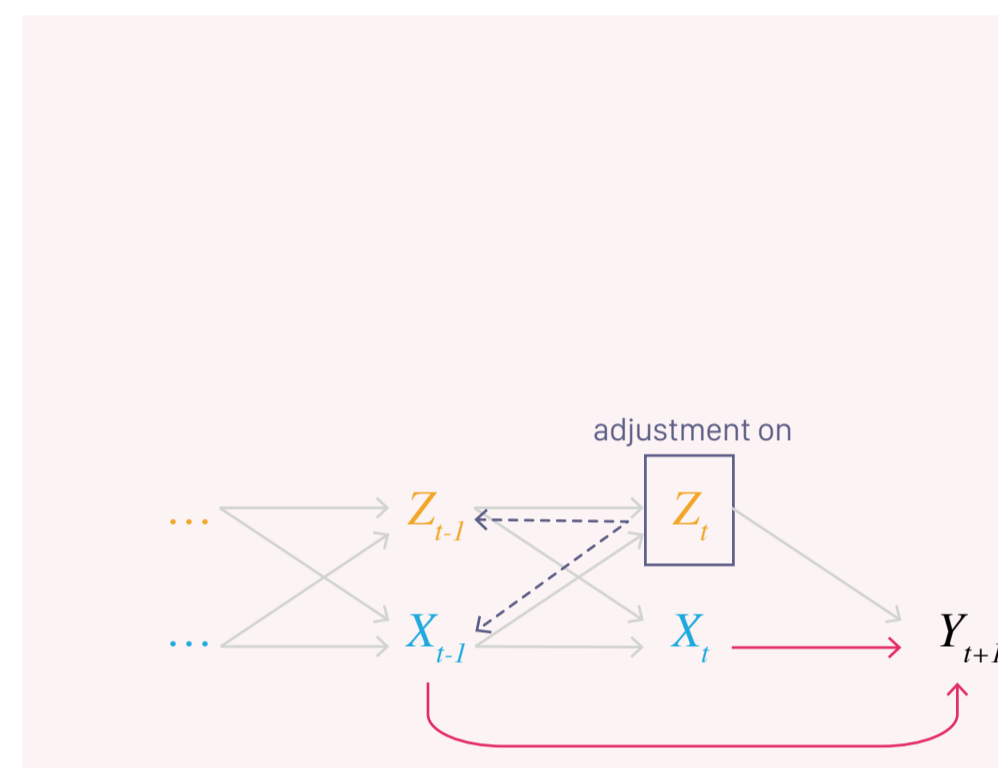
### IPW OR TRADITIONAL ADJUSTMENT METHODS?

$Z_i$  = Complication at  $i$      $X_i$  = VOC group at  $i$      $A_i$  = Age, sex ethnicity     $Y_i$  = Death at  $i$      $U$  = Unmeasured factor     $\leftarrow \dots \rightarrow$  Spurious correlation     $\rightarrow$  Causal effect



The association between the number of VOC in the previous year and the acute/chronic complication or death is a typical example of a study of a dynamic treatment regimen on a survival outcome (Hernan 2006). The traditional survival model was created to compare non-dynamic regimens i.e. treated versus non treated without any change during the follow up. In the case of a dynamic regimen, the model must include the time-varying treatment (the number of VOC each year) but also all time-varying covariates that could influence the treatment assignment (the presence of each complication in the previous year).

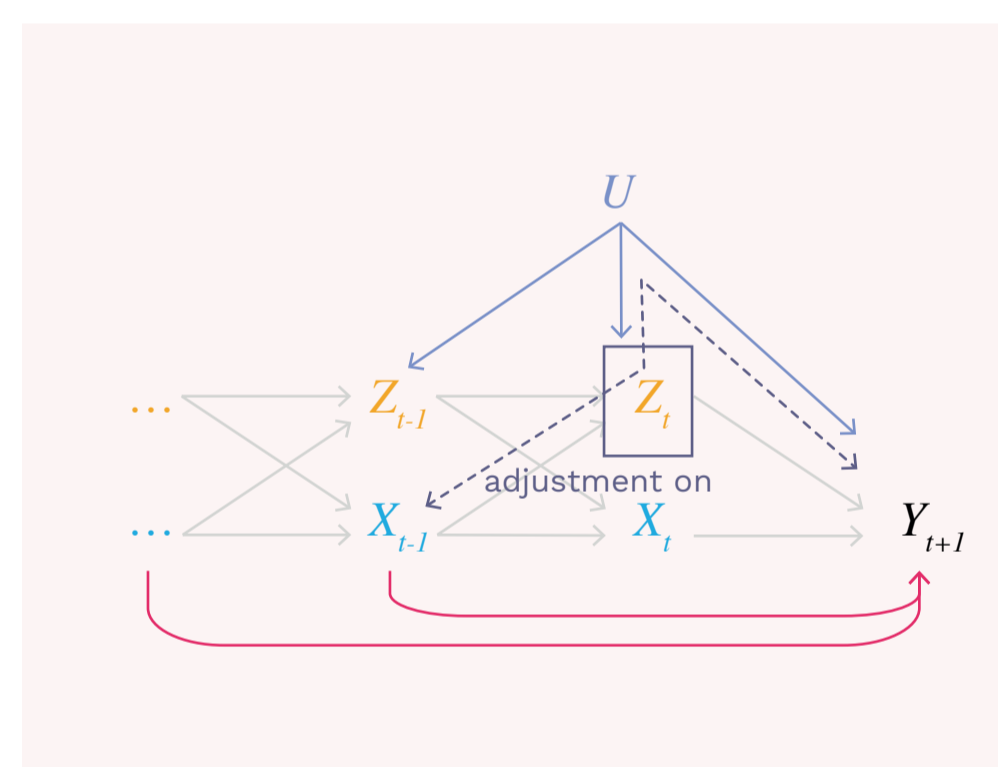
As we intended to estimate the causal effect of VOC on death, the following directed acyclic graph (DAG) was assumed (similar DAG could be drawn for each complications).



### Spurious association in traditional model

In case the objective of the study is to quantify the effect of VOC on death, the traditional survival model consists of a Cox model adjusted on VOC and complications. If we adjusted on the effect of a complication in the last year, we would have adjusted on a collider. Indeed, the complication in the last year is the common consequence of VOC two years before and the same complication two years before. This would have created a spurious correlation between the VOC and the complication two years before and imply a modification on the estimation of effect of VOC on death.

Thus the method that adjusts on time-varying VOC and complication is not appropriate for the purpose of this study.



### Case of unmeasured confounder

Moreover, in case of an unmeasured variable that could cause some complications and death, there is a risk that a backdoor path is opened.

As the purpose of electronic health records was not to record clinical diagnosis but to record the reimbursement of health consumption, it is plausible to have some unmeasured confounders. Thus, another method was sought that aimed to reduce these risks.

## THE MSM-IPW METHOD

$$w_i = \prod_{j \leq i} \frac{P(C_j = 0 | A_0)P(X_j = voc | A_0, C_j = 0)}{P(C_j = 0 | A_0, X_{j-1}, Z_{j-1})P(X_j = voc | A_0, X_{j-1}, Z_{j-1})}$$

The probability of being in each VOC category was estimated by multinomial logistic models.

The probability of being censored each year was estimated by logistic models.

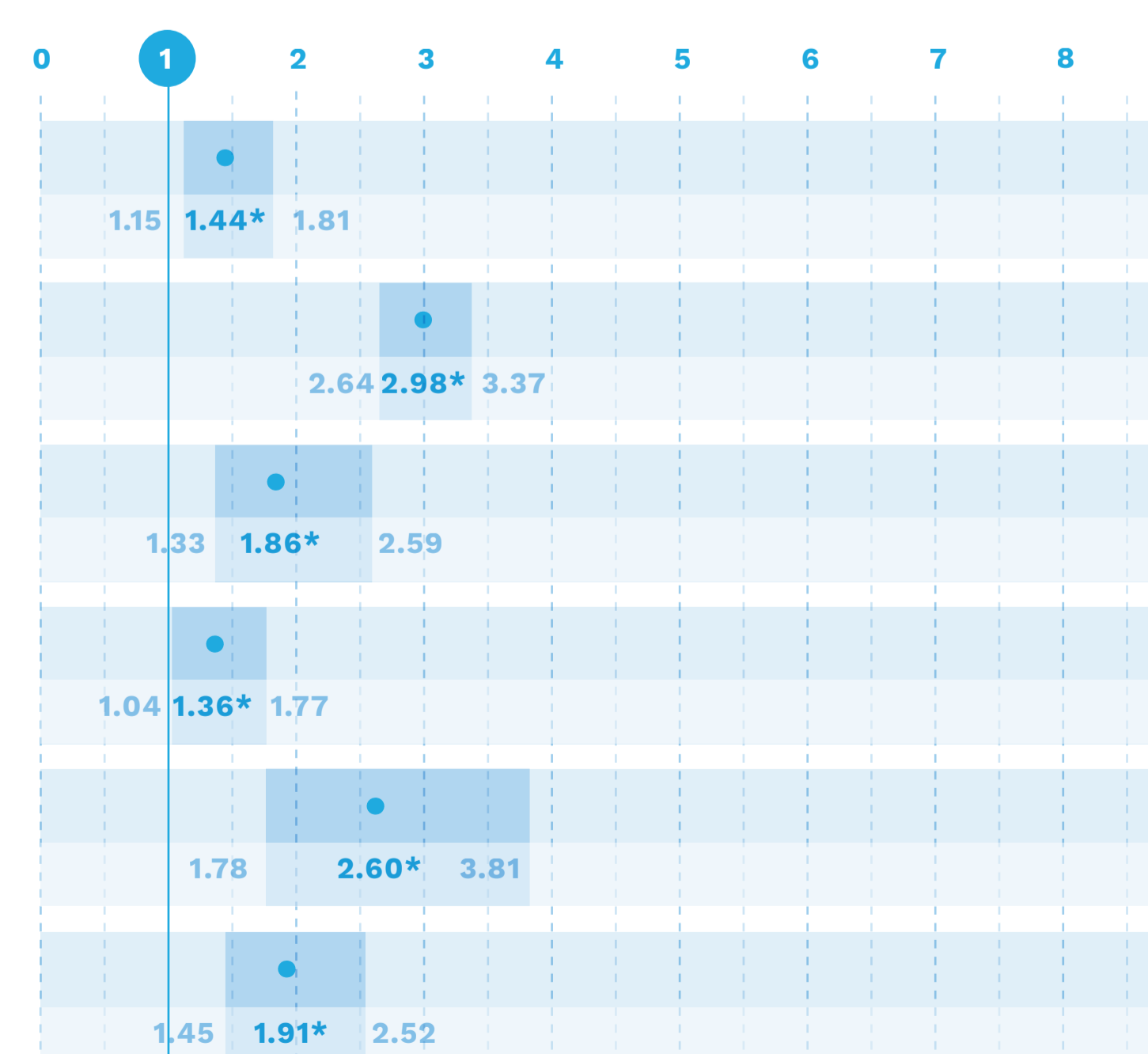
Supported by the limitation described above, the estimated causal effect of number of VOC on each complication and death was instead estimated using MSM-IPW Cox models approximated by weighted pooled logistic regressions.

The stabilized weights were estimated using a combination of the probability of being in each VOC category and the probability of being censored each year. In each year, these models were adjusted for age, gender, ethnicity, complications and comorbidities as reported the year before.

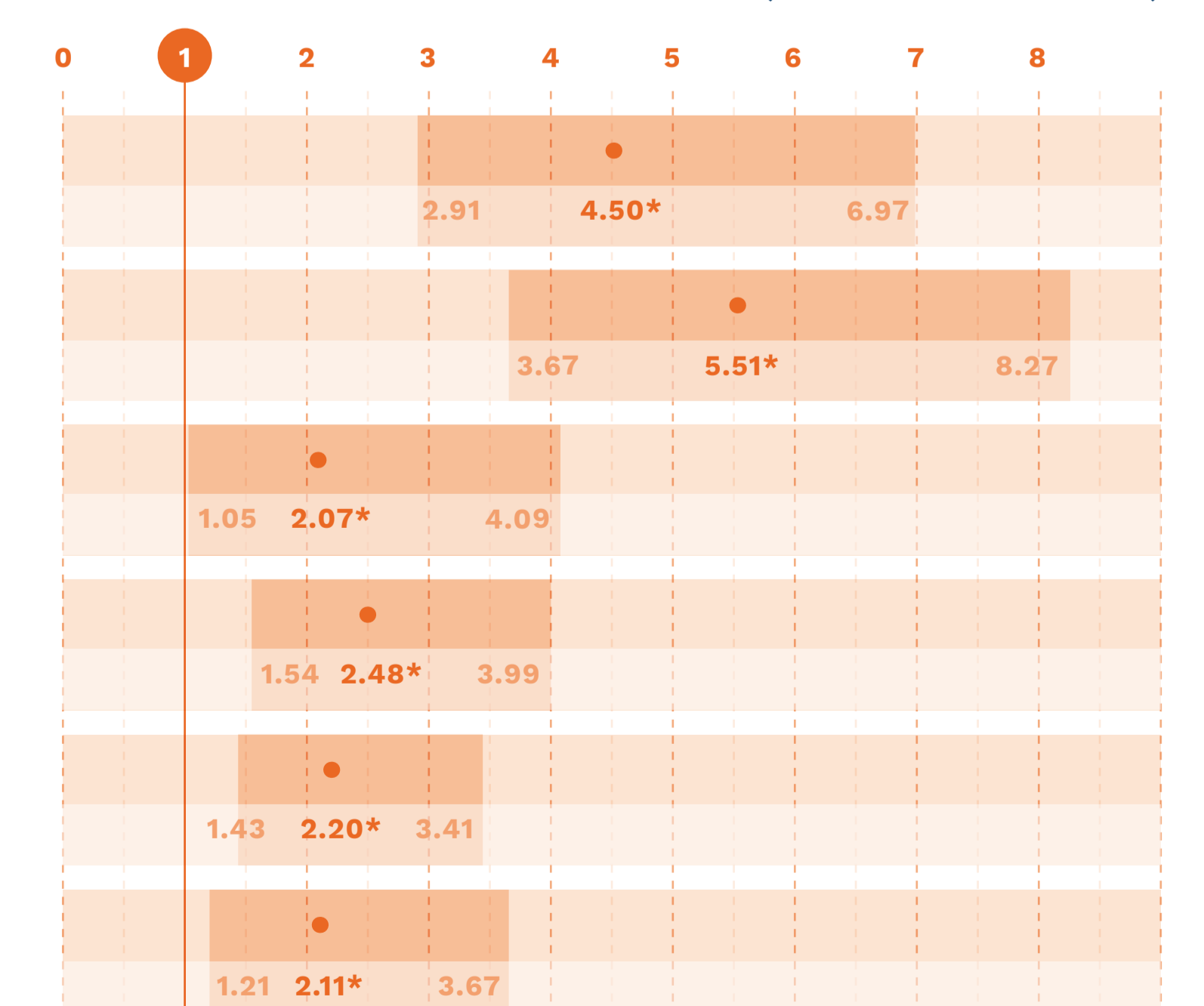
## Results

Complication	Patients
Death	1,141 patients (8%)
ACS	4,097 patients (27%)
Avascular necrosis	1,295 patients (9%)
Cardiomegaly	1,037 patients (7%)
Gall Stones	1,571 patients (10%)
Sepsis	1,098 patients (7%)

### 1-2 VOC vs 0 VOC<sup>1</sup>



### 3+ VOC vs 0 VOC<sup>1</sup>



\* Significant results at 5% level

## REFERENCES

Platt OS, Thorington BD, Brambilla DJ, Milner PF, Rosse WF, Vichinsky E, Kinney TR. Pain in sickle cell disease. Rates and risk factors. N Engl J Med. 1991 Jul 4;325(11):11-6.  
Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. Basic Clin Pharmacol Toxicol. 2006 Mar;98(3):237-42.

