

Yu Yang Soon<sup>1,2,3</sup>, Ian Marschner<sup>1</sup>, Manjula Schou<sup>1</sup>, Christopher J Sweeney<sup>4,5</sup>, Ian D. Davis<sup>6,7</sup>, Martin Stockler<sup>1,8,9</sup>, Andrew Martin<sup>1,10</sup>

<sup>1</sup>NHMRC Clinical Trials Centre, University of Sydney, Australia; <sup>2</sup>National University Cancer Institute, Singapore, Singapore; <sup>3</sup>National University of Singapore, Singapore; <sup>4</sup>South Australian Immunogenomics Cancer Institute, Australia; <sup>5</sup>University of Adelaide, Australia; <sup>6</sup> Monash University, Australia; <sup>7</sup>Eastern Health, Australia; <sup>8</sup>Concord Repatriation General Hospital, Australia; <sup>9</sup>Chris O'Brien Lifehouse, Australia; <sup>10</sup>University of Queensland, Australia

## 1. Background

A beneficial treatment effect reported at the first interim analysis (IA) may diminish at a subsequent analysis (SA)

We examined three challenges in interpreting treatment effects from randomized clinical trials (RCTs) after the first positive IA

- Overestimation bias
- Non-proportional hazards
- Heterogeneity over recruitment in participant profiles and treatment practice

## 2. Methods

We identified 71 oncology RCTs reporting positive results at the initial IA and an SA for event-free survival (EFS) and overall survival (OS).

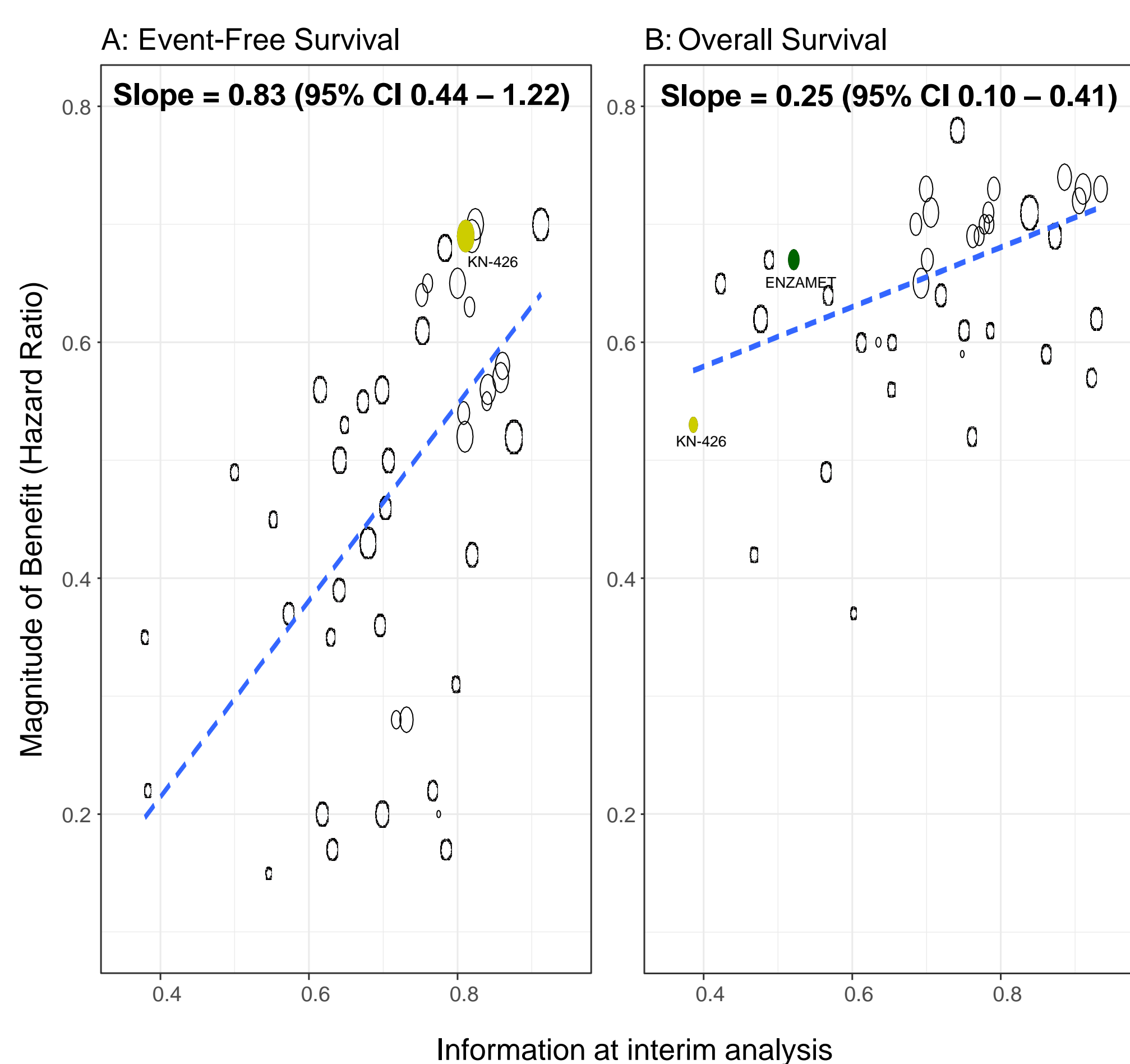
We modelled

- The hazard ratio at IA ( $HR_{IA}$ ) vs its timing as measured by the information fraction (IF; events at IA vs total event sought)
- The ratio of  $HR_{IA}$  to  $HR_{SA}$  vs IF
- Repeated for  $HR_{IA}$  adjusted for overestimation bias using a penalized estimation method (Marschner Stat Methods Med Res 2022)

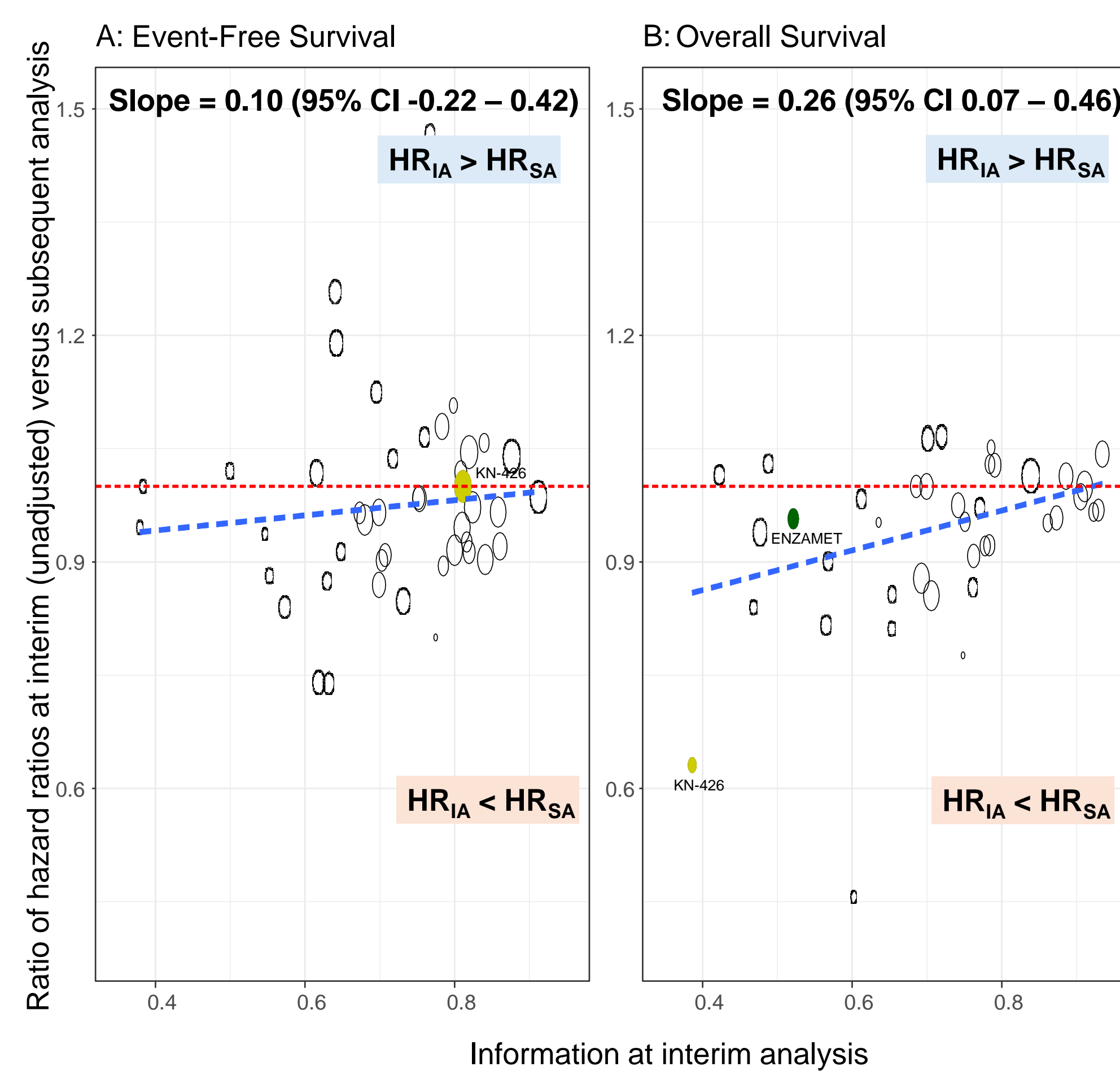
Examples (ENZAMET and KN-426 RCTs) of the other two challenges were sought

## 3. Results

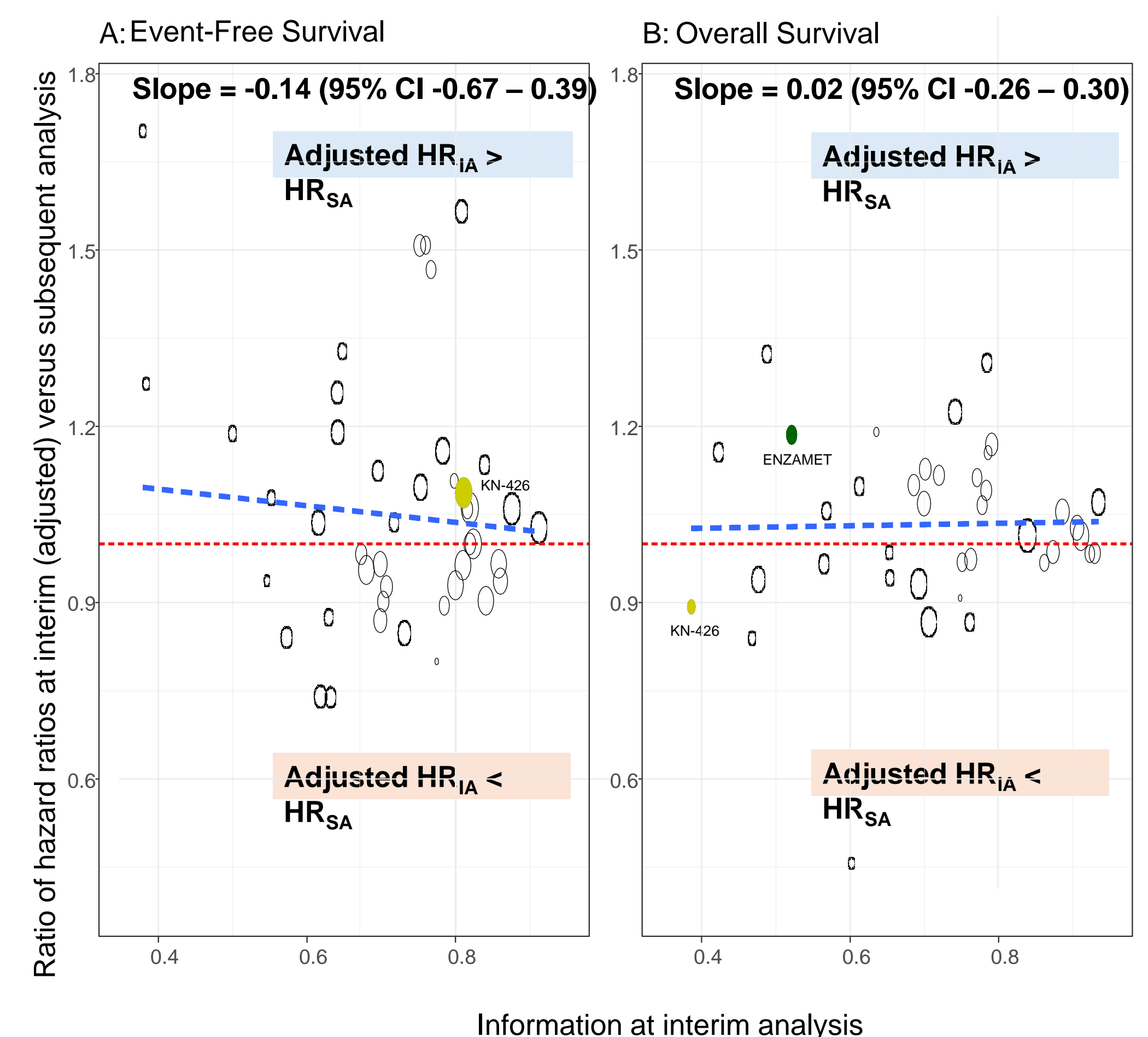
### $HR_{IA}$ were positively associated with the IF



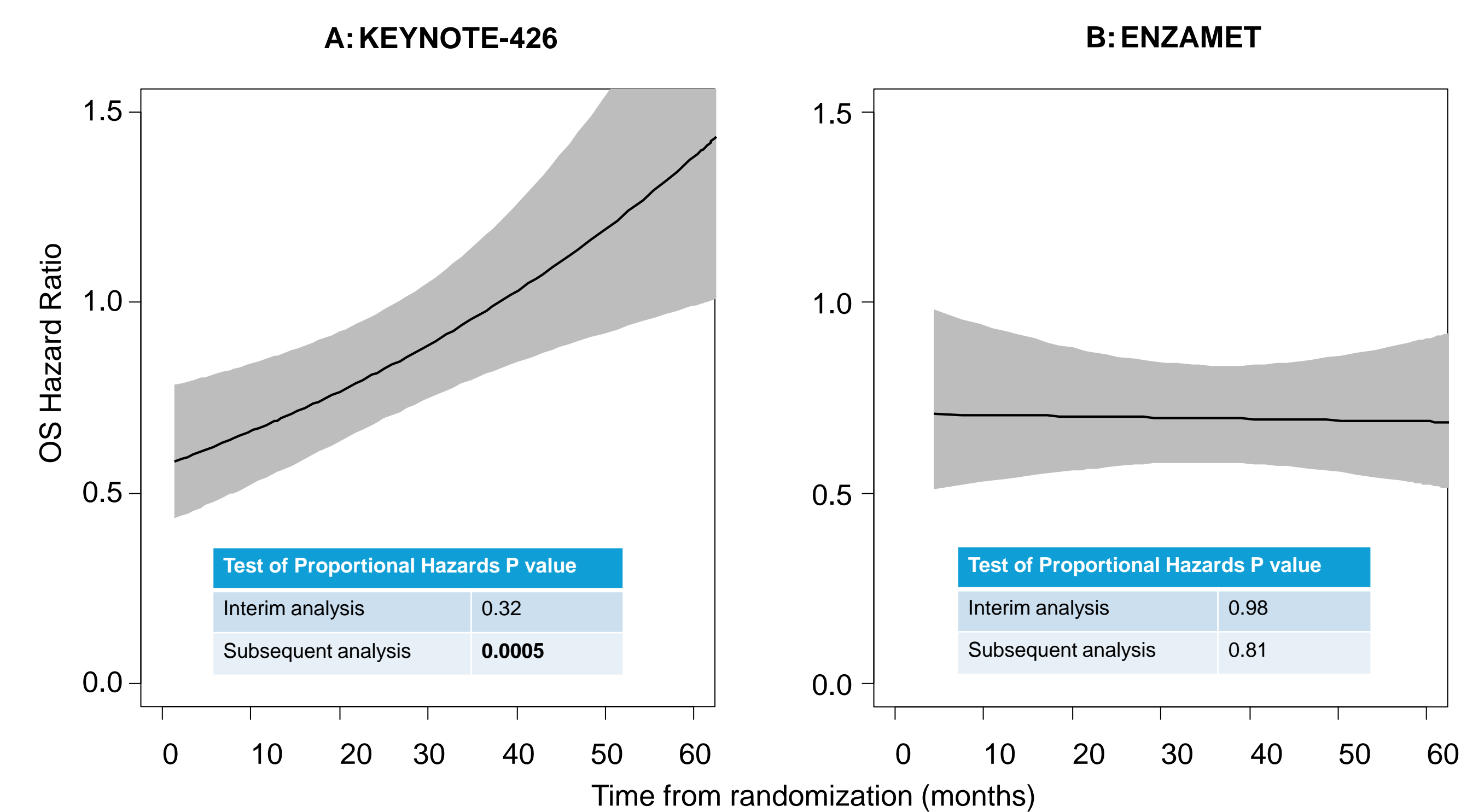
### $HR_{IA}$ tended to exaggerate $HR_{SA}$



### Adjusted $HR_{IA}$ did not exaggerate $HR_{SA}$



### Non-proportional hazards can occur at subsequent analysis



### Heterogeneity over recruitment in participant profiles and treatment practice

Characteristic	Recruitment period 1: 31 Mar 2014 – 24 Nov 2015			Recruitment period 2: 25 Nov 2015 – 4 Aug 2016			Recruitment period 3: 5 Aug 2016 – 24 Nov 2017		
	Control (n = 187)	Enzalutamide (n = 189)	Overall	Control (n = 188)	Enzalutamide (n = 187)	Overall	Control (n = 188)	Enzalutamide (n = 187)	Overall
Australia	141 (75)	145 (77)	286 (76)	86 (46)	96 (51)	182 (49)	94 (50)	83 (44)	177 (47)
Planned early docetaxel	42 (23)	44 (23)	86 (23)	115 (61)	110 (59)	225 (60)	93 (50)	99 (53)	192 (51)
$HR_{IA}$ (95% CI)	0.65 (0.45 – 0.95)			0.90 (0.57 – 1.41)			0.47 (0.27 – 0.82)		
$HR_{SA}$ (95% CI)	0.65 (0.48 – 0.88)			0.91 (0.66 – 1.24)			0.57 (0.41 – 0.79)		

Estimates from positive first interim may overestimate treatment effects

A penalized maximum likelihood estimator may correct this overestimation bias

Follow-up post-interim analysis allows accurate assessment of the treatment effect