PSI 2025 Conference
9th June 2025

WHEN TO SCHEDULE THE INTERIM ANALYSIS IN THE PRESENCE OF MISSING DATA?

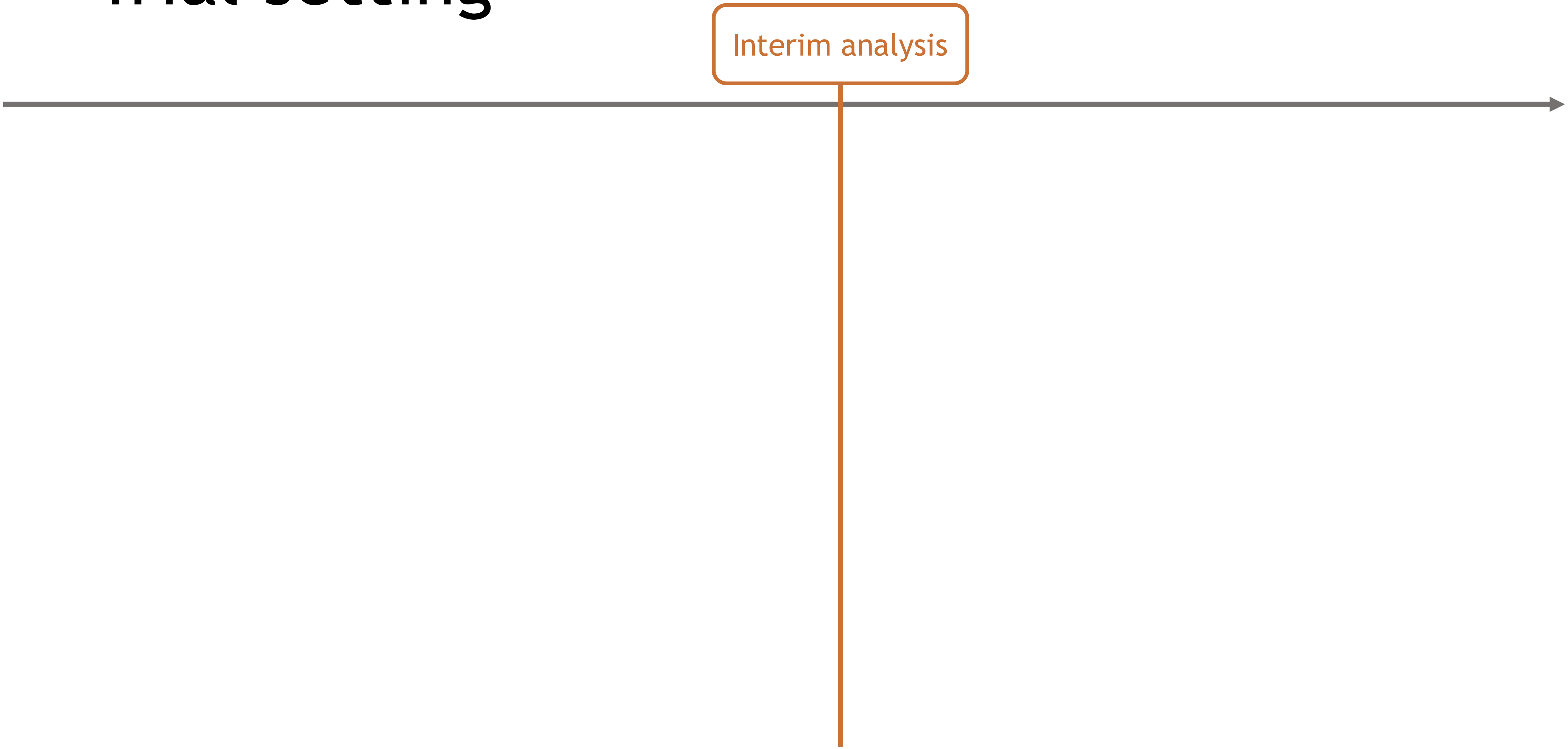
Neža Dvoršak [1], Jianmei Wang [2], Thomas Burnett [1], Christopher Jennison [1], Robin Mitra [3]

[1] University of Bath, [2] Roche UK, [3] University College London

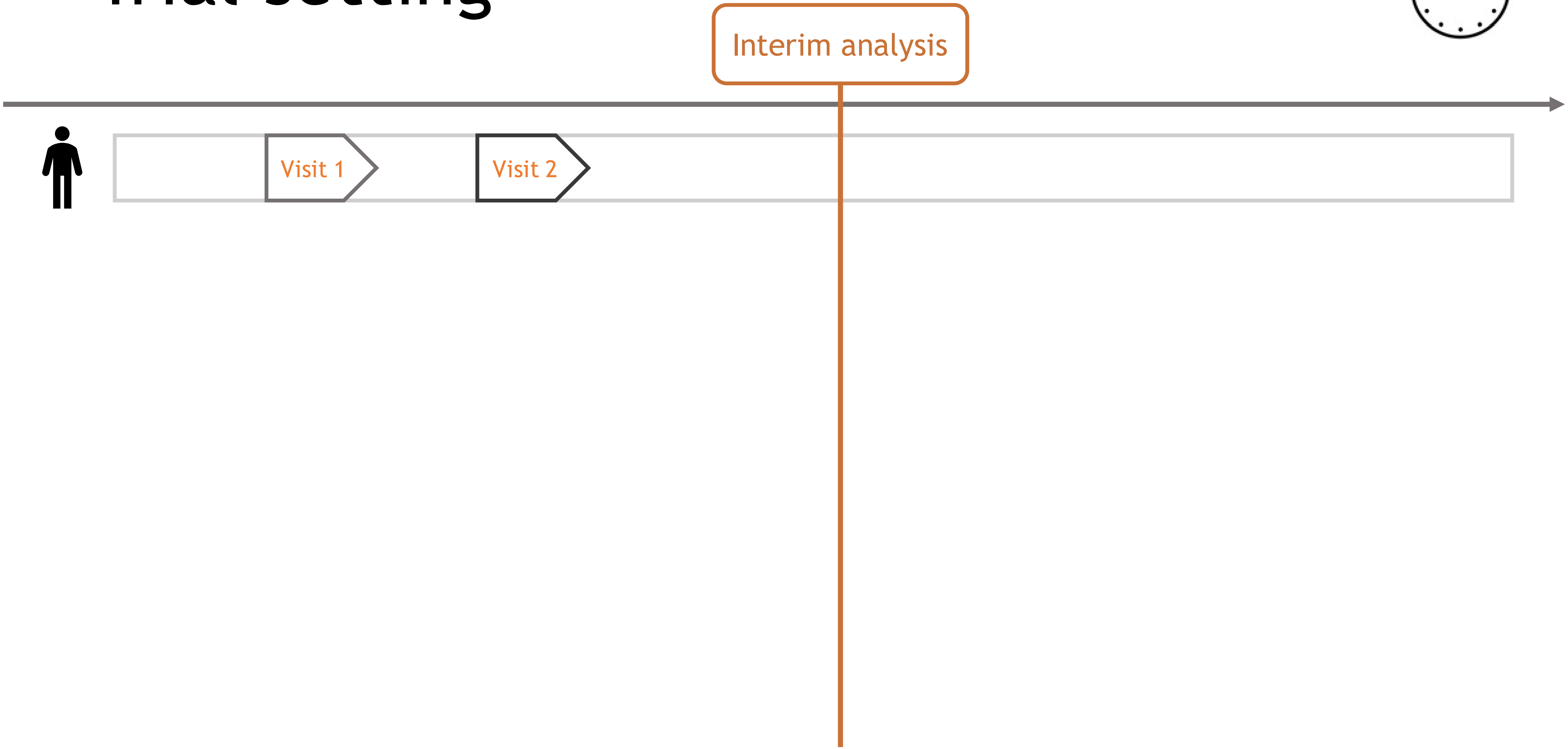
UNIVERSITY OF BATH

Trial setting

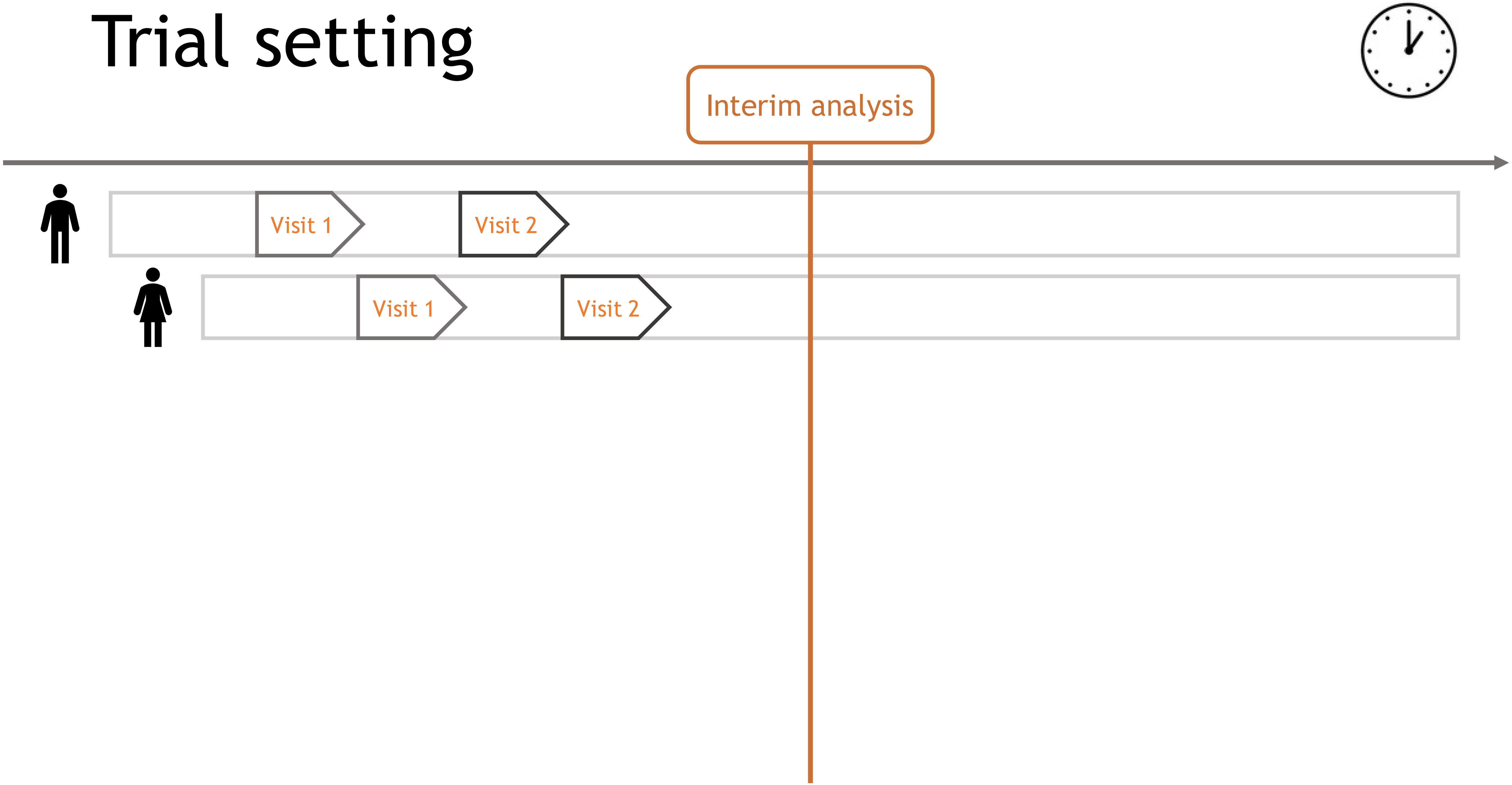
Trial setting



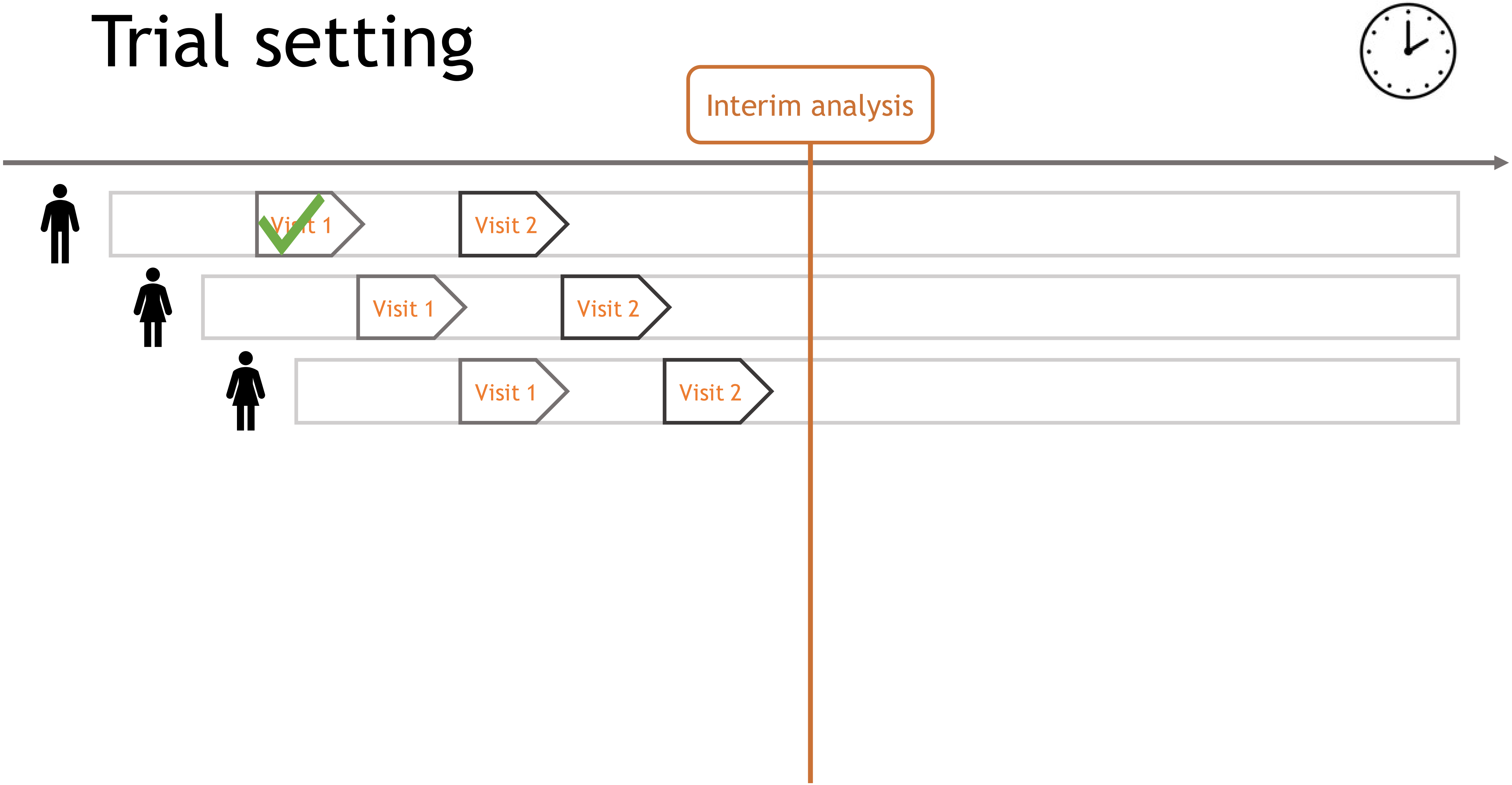
Trial setting



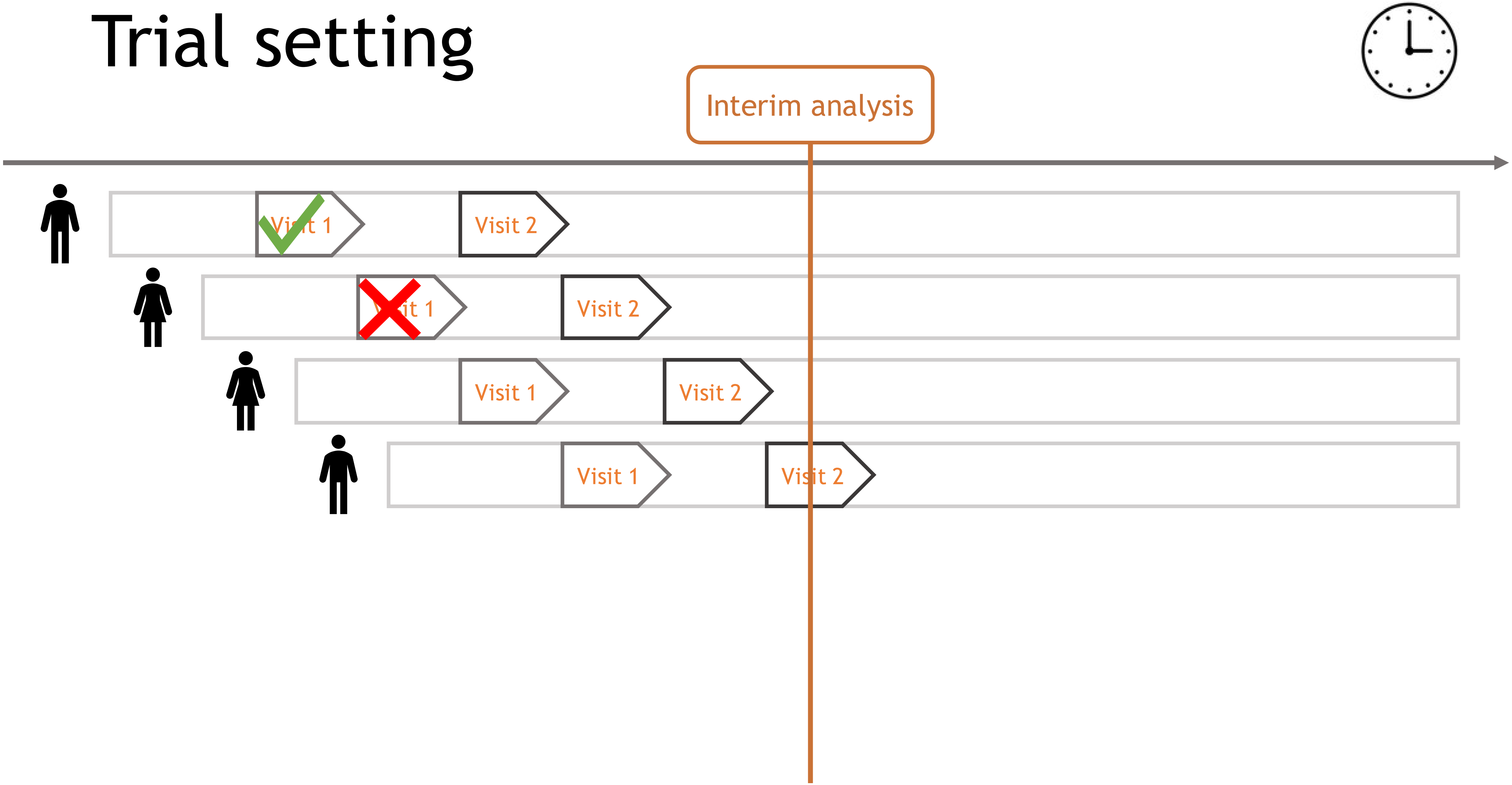
Trial setting



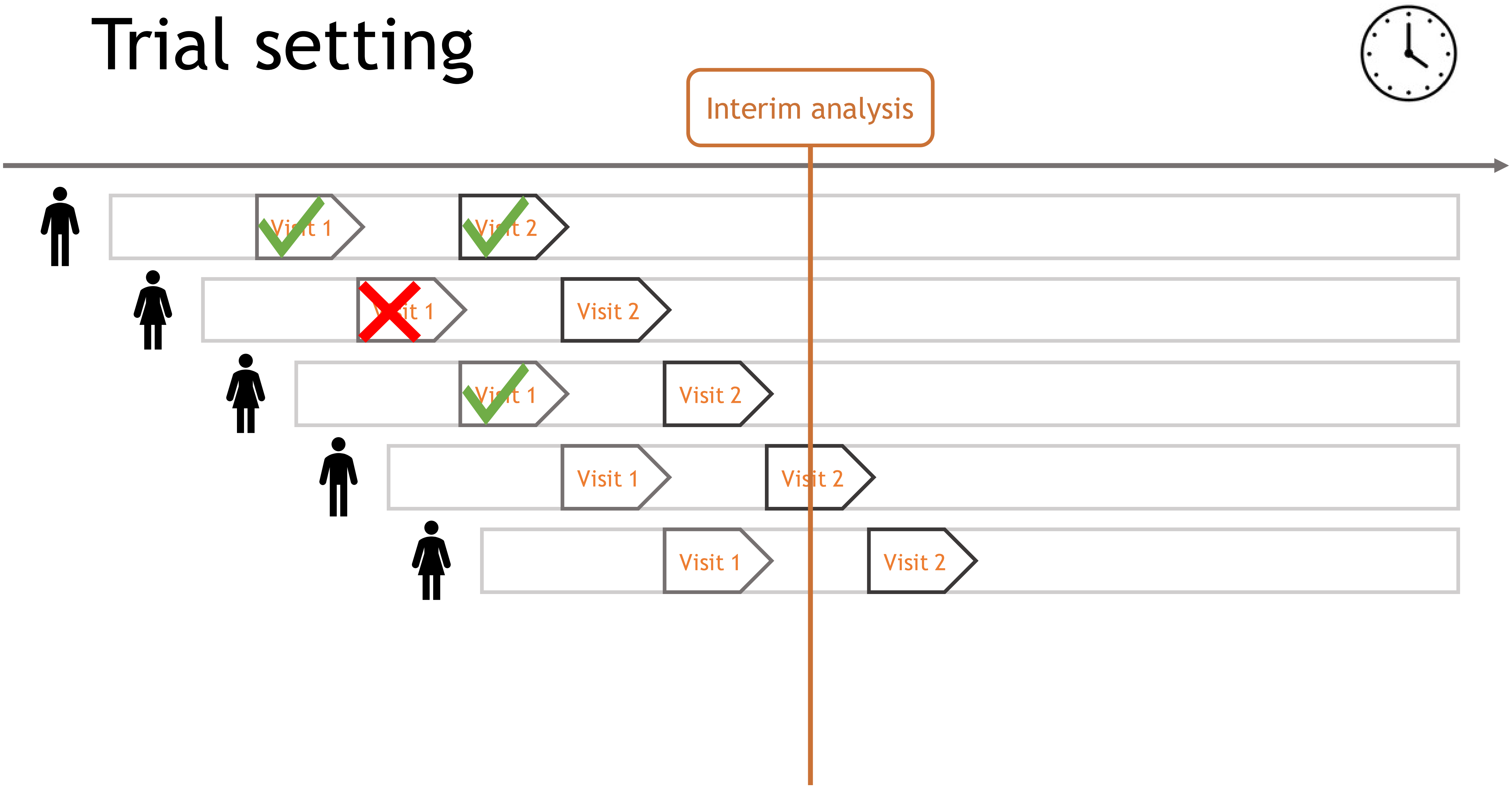
Trial setting



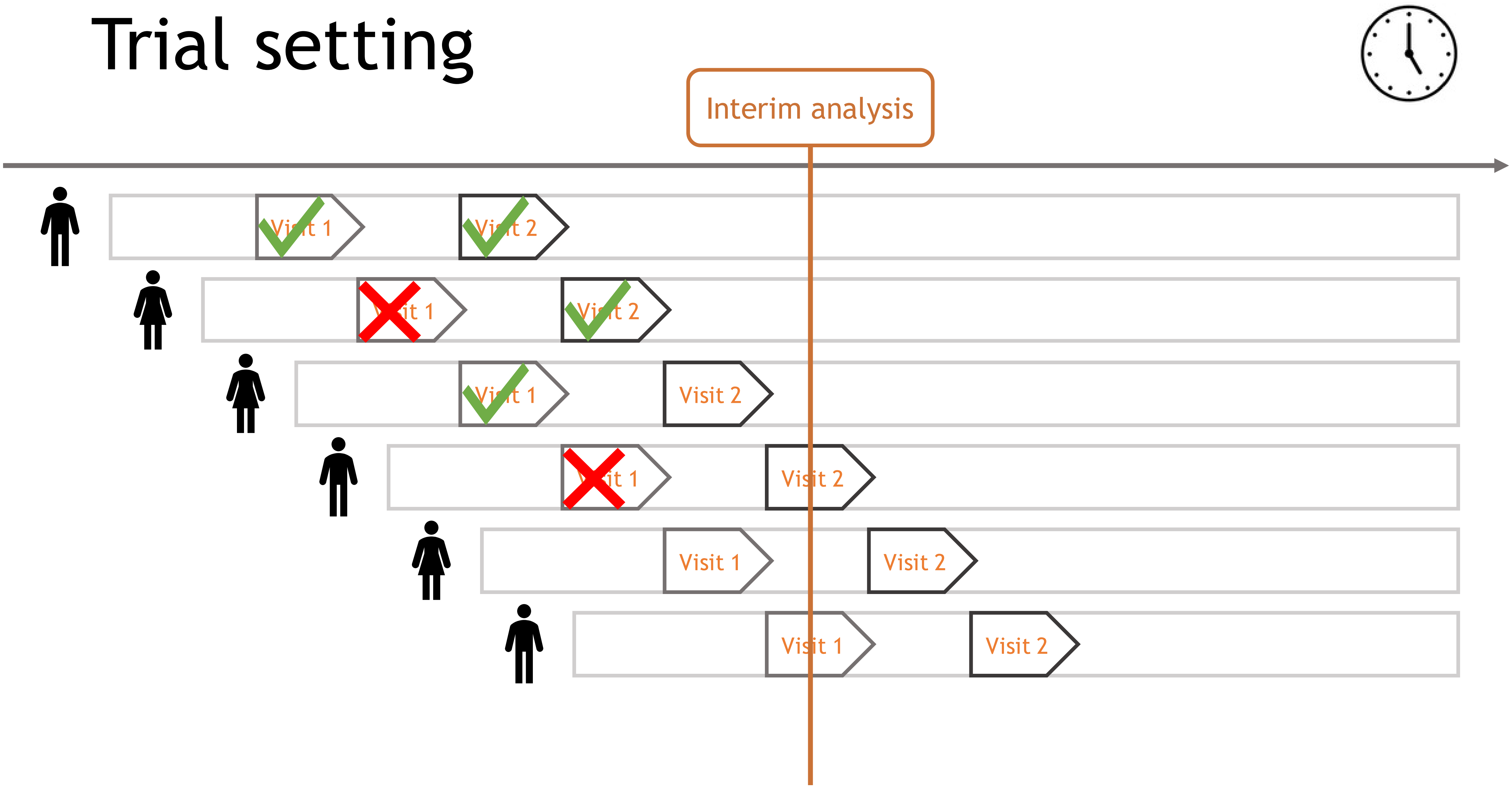
Trial setting



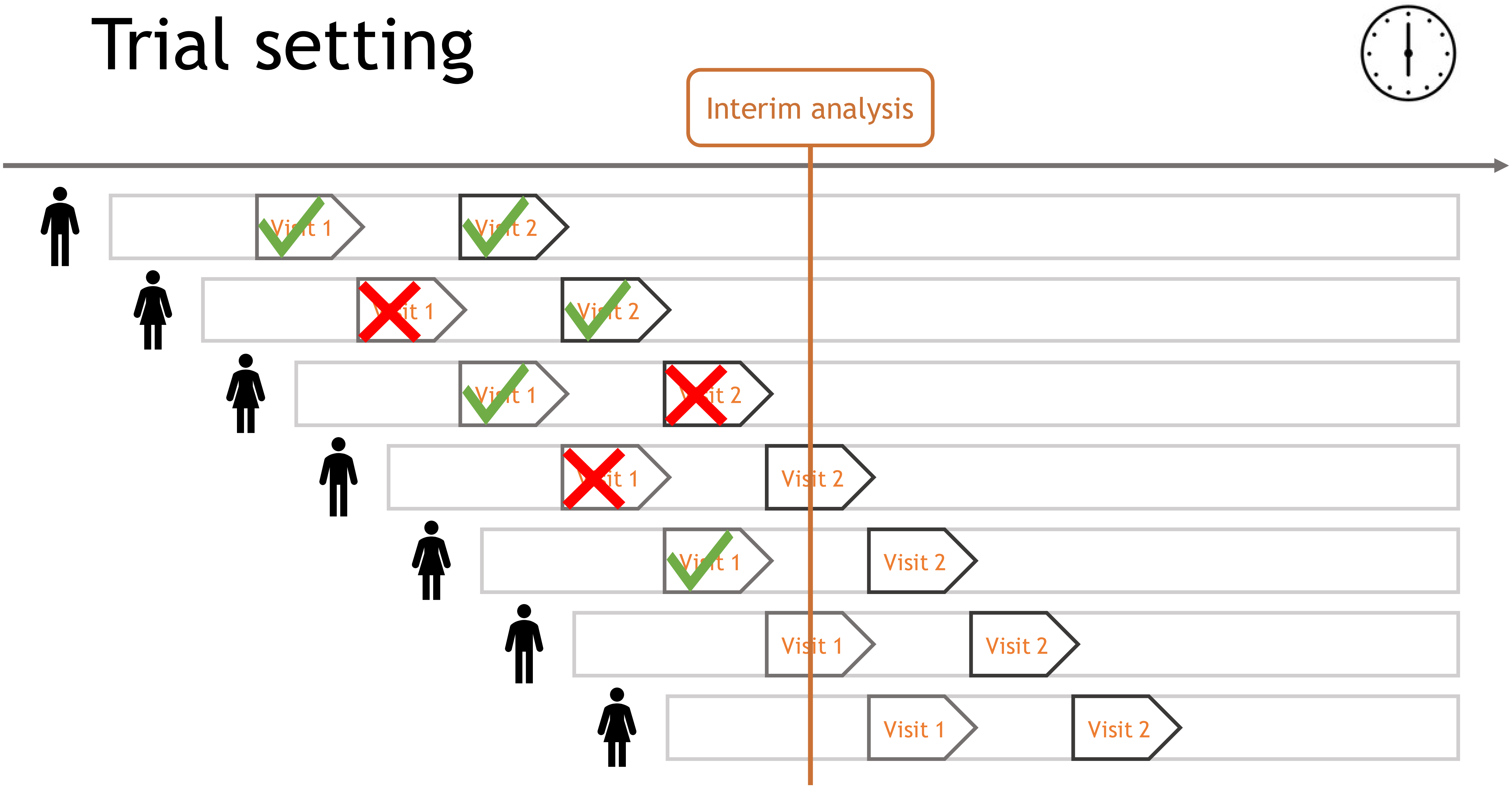
Trial setting



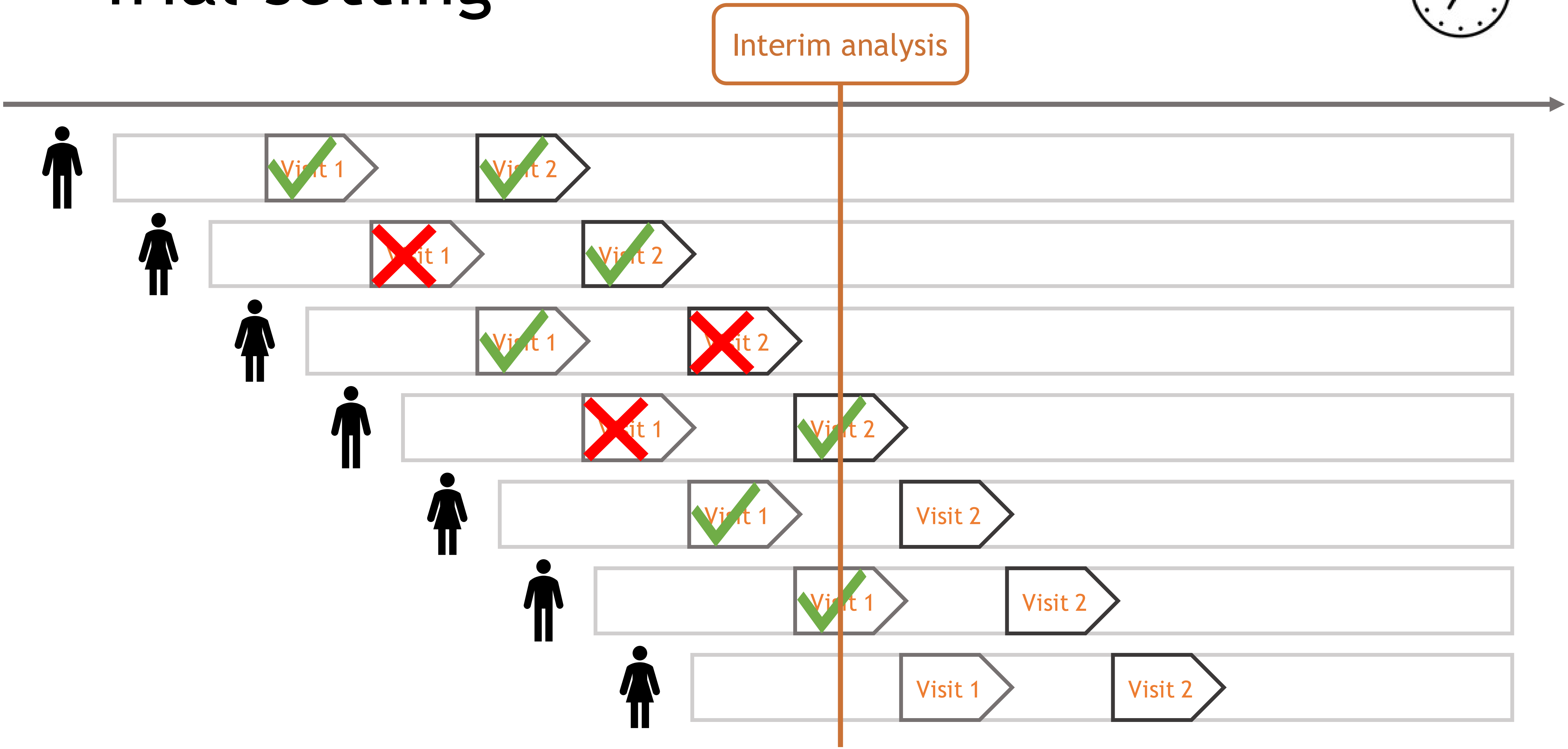
Trial setting



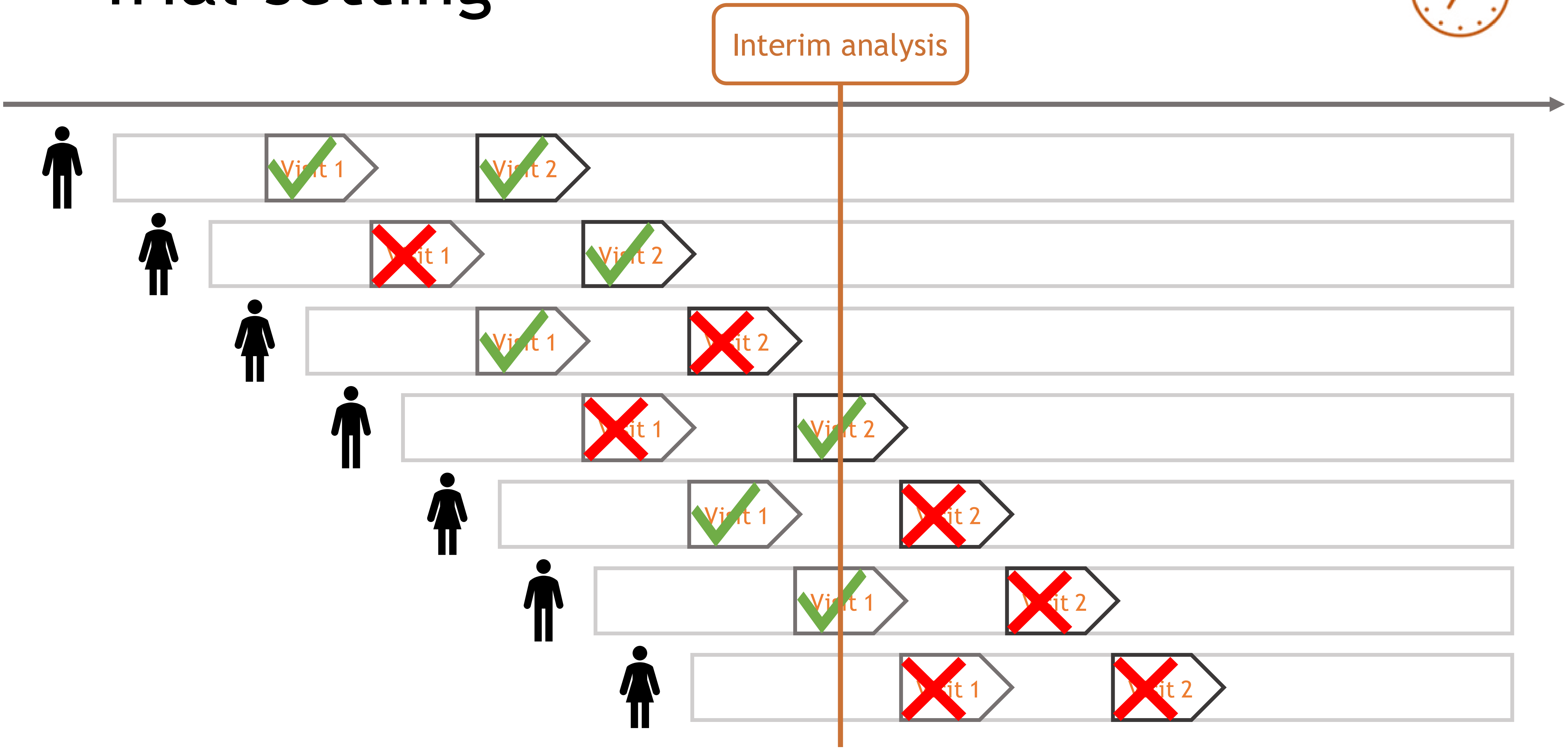
Trial setting



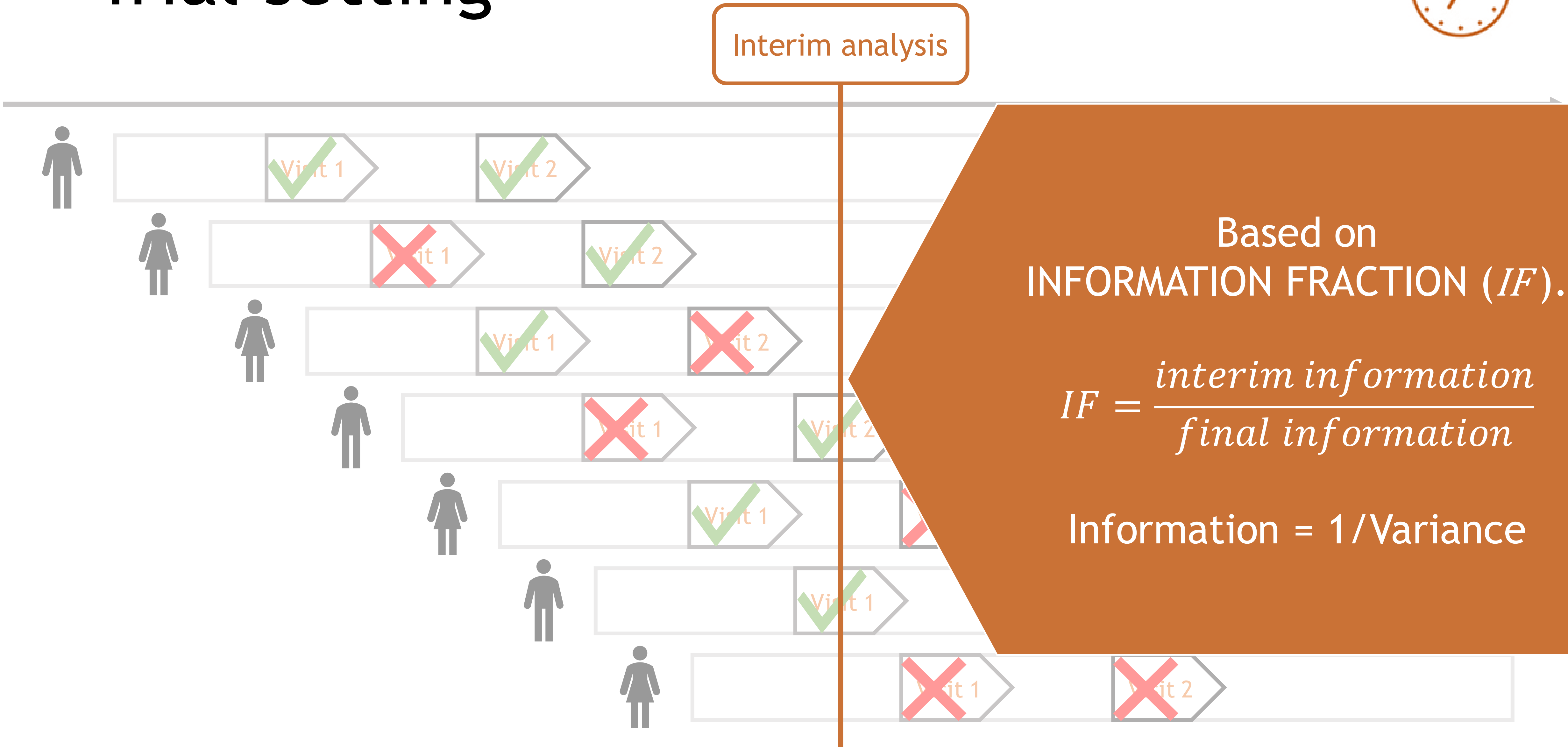
Trial setting



Trial setting



Trial setting



Based on
INFORMATION FRACTION (IF).

$$IF = \frac{interim\ information}{final\ information}$$

Information = 1/Variance

Data patterns

Pattern	Visit 1	Visit 2
1	✓	✓
2	✗	✓
3	✓	✗
0	✗	✗

UNIVERSITY OF BATH

Data structure

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Data structure

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
	μ_1	μ_2	

$$(\bar{X}_{11}, \bar{X}_{12}) \sim N\left(\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}, \frac{\sigma^2}{n_1}\begin{bmatrix}1 & \rho^2 \\ \rho^2 & 1\end{bmatrix}\right)$$

$$\bar{X}_{22} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$$

$$\bar{X}_{31} \sim N\left(\mu_1, \frac{\sigma^2}{n_3}\right)$$

Data modelling

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3

$$(\bar{X}_{11}, \bar{X}_{12}) \sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \frac{\sigma^2}{n_1}\begin{bmatrix}1 & \rho^2\\ \rho^2 & 1\end{bmatrix}\right)$$

$$\bar{X}_{22} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$$

$$\bar{X}_{31} \sim N\left(\mu_1, \frac{\sigma^2}{n_3}\right)$$

$$\boldsymbol{Y} = \begin{pmatrix}\bar{X}_{11}\\ \bar{X}_{12}\\ \bar{X}_{22}\\ \bar{X}_{31}\end{pmatrix} \quad E(\boldsymbol{Y}) = \begin{pmatrix}\mu_1\\ \mu_2\\ \mu_2\\ \mu_1\end{pmatrix} = D\boldsymbol{\mu} = \begin{pmatrix}1 & 0\\ 0 & 1\\ 0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$$

$$Var(\boldsymbol{Y}) = \sigma^2\begin{pmatrix}\frac{1}{n_1} & \frac{\rho^2}{n_1} & 0 & 0\\ \frac{\rho^2}{n_1} & \frac{1}{n_1} & & \\ 0 & & \frac{1}{n_2} & 0\\ 0 & & 0 & \frac{1}{n_3}\end{pmatrix}$$

Data modelling

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3

$$(\bar{X}_{11}, \bar{X}_{12}) \sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \frac{\sigma^2}{n_1}\begin{bmatrix}1 & \rho^2\\ \rho^2 & 1\end{bmatrix}\right)$$

$$\bar{X}_{22} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$$

$$\bar{X}_{31} \sim N\left(\mu_1, \frac{\sigma^2}{n_3}\right)$$

Parameter of interest: μ_2

$$\boldsymbol{Y} = \begin{pmatrix}\bar{X}_{11}\\ \bar{X}_{12}\\ \bar{X}_{22}\\ \bar{X}_{31}\end{pmatrix} \quad E(\boldsymbol{Y}) = \begin{pmatrix}\mu_1\\ \mu_2\\ \mu_2\\ \mu_1\end{pmatrix} = D\boldsymbol{\mu} = \begin{pmatrix}1 & 0\\ 0 & 1\\ 0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$$

$$Var(\boldsymbol{Y}) = \sigma^2\begin{pmatrix}\begin{matrix}\frac{1}{n_1} & \frac{\rho^2}{n_1}\\ \frac{\rho^2}{n_1} & \frac{1}{n_1}\end{matrix} & 0 & 0\\ 0 & \frac{1}{n_2} & 0\\ 0 & 0 & \frac{1}{n_3}\end{pmatrix}$$

UNIVERSITY OF BATH

Variance of $\hat{\mu}_2$

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3

$$\boldsymbol{Y} = \begin{pmatrix} \bar{X}_{11} \\ \bar{X}_{12} \\ \bar{X}_{22} \\ \bar{X}_{31} \end{pmatrix} \quad E(\boldsymbol{Y}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_1 \end{pmatrix} = D\boldsymbol{\mu} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$Var(\boldsymbol{Y}) = \sigma^2 \begin{pmatrix} \begin{matrix} \frac{1}{n_1} & \frac{\rho^2}{n_1} \\ \frac{\rho^2}{n_1} & \frac{1}{n_1} \end{matrix} & 0 & 0 \\ 0 & \frac{1}{n_2} & 0 \\ 0 & 0 & \frac{1}{n_3} \end{pmatrix}$$

$$\hat{\mu}_2 = (0 \quad 1)(D^T Var(\boldsymbol{Y})^{-1} D)^{-1} D^T Var(\boldsymbol{Y})^{-1} \boldsymbol{Y}$$

$$Var(\hat{\mu}_2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} (D^T Var(\boldsymbol{Y})^{-1} D)^{-1}$$

$$\boxed{Var(\hat{\mu}_2) = \frac{\sigma^2}{\dfrac{n_1(n_1+n_2+n_3) + n_2 n_3 (1-\rho^2)}{n_1 + n_3(1-\rho^2)}}}$$

UNIVERSITY OF BATH

Information of $\hat{\mu}_2$

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3

$$\boldsymbol{Y} = \begin{pmatrix} \bar{X}_{11} \\ \bar{X}_{12} \\ \bar{X}_{22} \\ \bar{X}_{31} \end{pmatrix} \quad E(\boldsymbol{Y}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_1 \end{pmatrix} = D\boldsymbol{\mu} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$Var(\boldsymbol{Y}) = \sigma^2 \begin{pmatrix} \begin{matrix} \frac{1}{n_1} & \frac{\rho^2}{n_1} \\ \frac{\rho^2}{n_1} & \frac{1}{n_1} \end{matrix} & 0 & 0 \\ 0 & \frac{1}{n_2} & 0 \\ 0 & 0 & \frac{1}{n_3} \end{pmatrix}$$

$$\hat{\mu}_2 = (0 \quad 1)(D^T Var(\boldsymbol{Y})^{-1} D)^{-1} D^T Var(\boldsymbol{Y})^{-1} \boldsymbol{Y}$$

$$Var(\hat{\mu}_2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} (D^T Var(\boldsymbol{Y})^{-1} D)^{-1}$$

$$Var(\hat{\mu}_2) = \frac{\sigma^2}{\frac{n_1(n_1 + n_2 + n_3) + n_2 n_3 (1 - \rho^2)}{n_1 + n_3(1 - \rho^2)}}$$

$$\boxed{I(\hat{\mu}_2) = \frac{\frac{n_1(n_1 + n_2 + n_3) + n_2 n_3 (1 - \rho^2)}{n_1 + n_3(1 - \rho^2)}}{\sigma^2}}$$

Equivalent sample size

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3

<u>Complete patients only (Pattern 1):</u>

$$I(\hat{\mu}_2) = \frac{\boxed{n}}{\sigma^2}$$

<u>Partial patients:</u>

$$I(\hat{\mu}_2) = \frac{\boxed{\dfrac{n_1(n_1+n_2+n_3)+n_2n_3(1-\rho^2)}{n_1+n_3(1-\rho^2)}}}{\sigma^2}$$

How many complete patients are the partial patients equivalent to?

$$n \equiv \frac{n_1(n_1+n_2+n_3)+n_2n_3(1-\rho^2)}{n_1+n_3(1-\rho^2)}$$

UNIVERSITY OF BATH

Timing of the interim analysis

→Based on information fraction (IF), determined by equivalent sample size

$$\boldsymbol{IF} = \frac{\text{interim information}}{\text{final information}} = \frac{\frac{\text{interim equivalent sample size}}{\sigma^2}}{\frac{\text{final equivalent sample size}}{\sigma^2}} = \boldsymbol{\frac{\textbf{interim equivalent sample size}}{\textbf{final equivalent sample size}}}$$

UNIVERSITY OF BATH

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION?

PRACTICAL EXAMPLE

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- No missing data (Pattern 1 only)

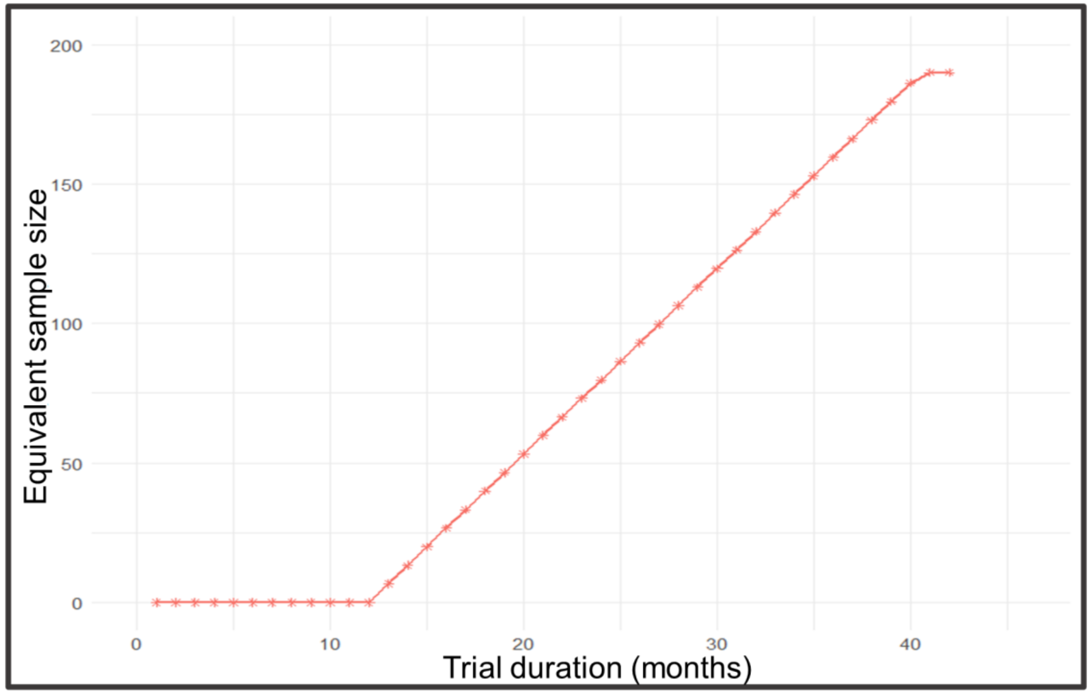
UNIVERSITY OF BATH

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- No missing data (Pattern 1 only)



Pattern	Visit 1	Visit 2
1	$\bar{X}_{11}$	$\bar{X}_{12}$

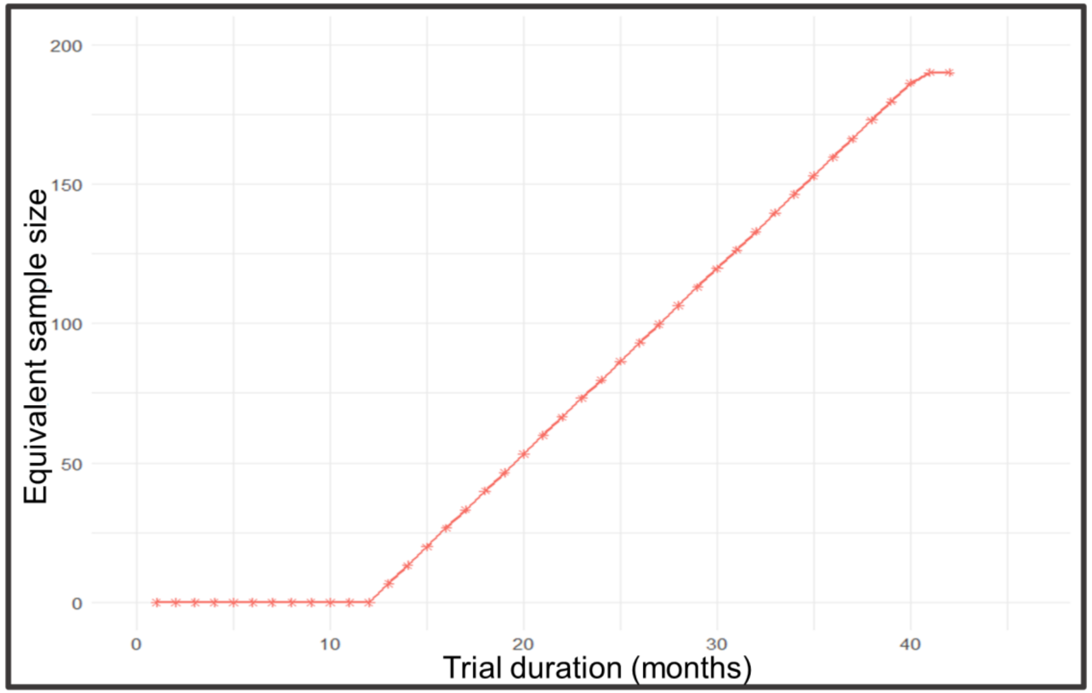
n_1

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- No missing data (Pattern 1 only)

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION?



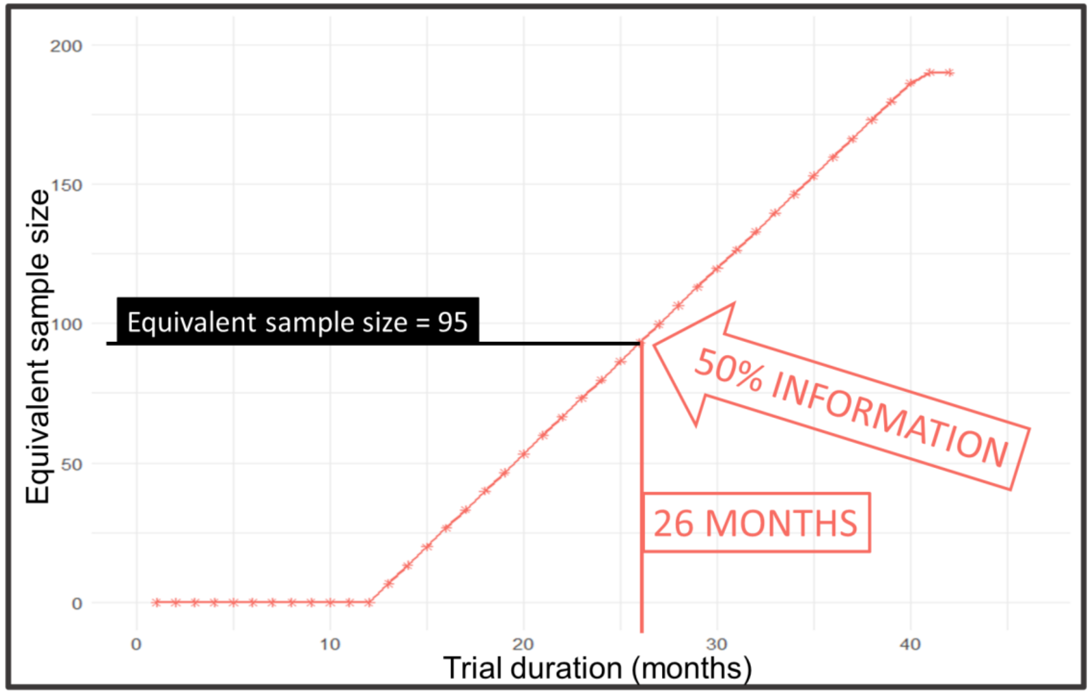
Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- No missing data (Pattern 1 only)

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION? AT 26 MONTHS



Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- Missing data from staggered entry and dropout ($P(dropout) = 0.05$)

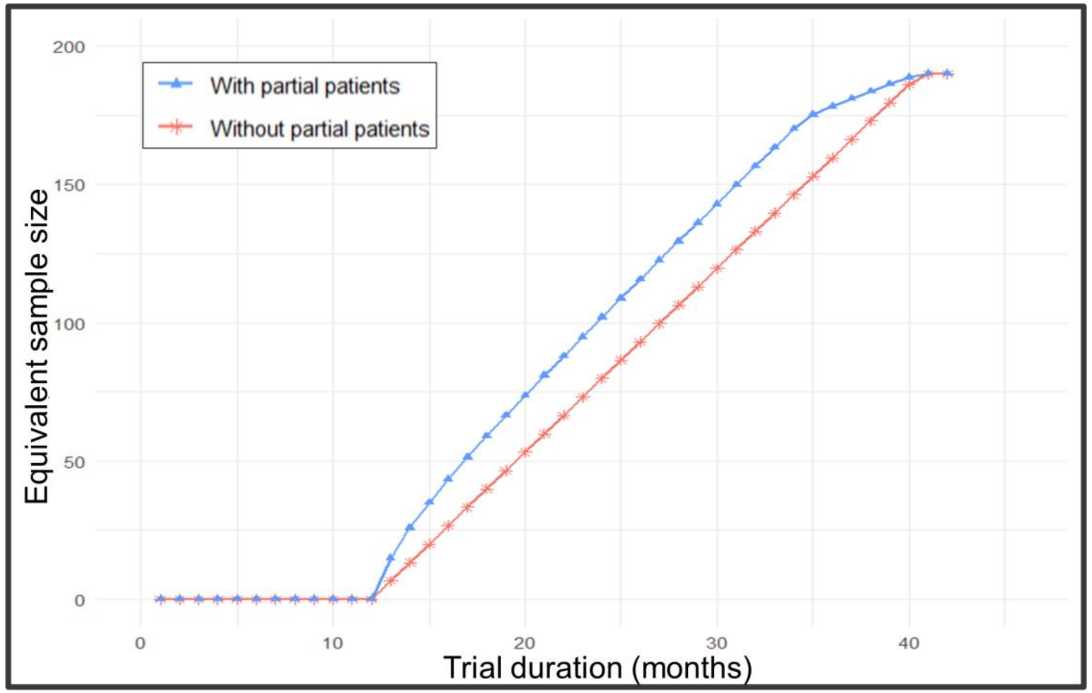
UNIVERSITY OF BATH

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- Missing data from staggered entry and dropout ($P(dropout) = 0.05$)



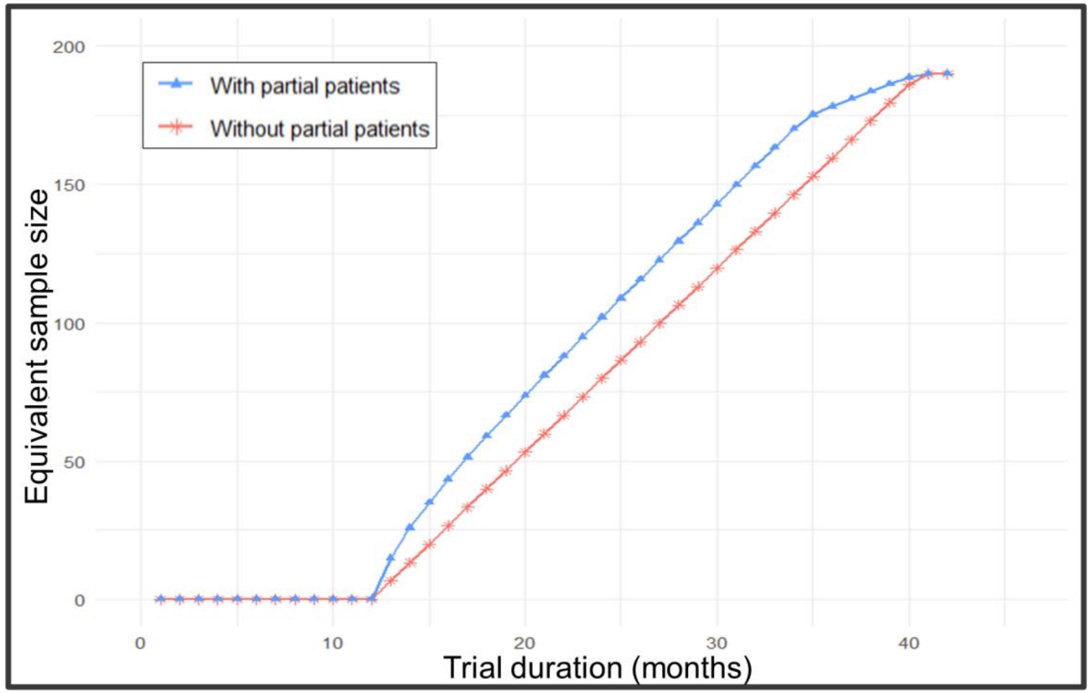
Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

1. Design stage

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- Missing data from staggered entry and dropout ($P(dropout) = 0.05$)

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION?



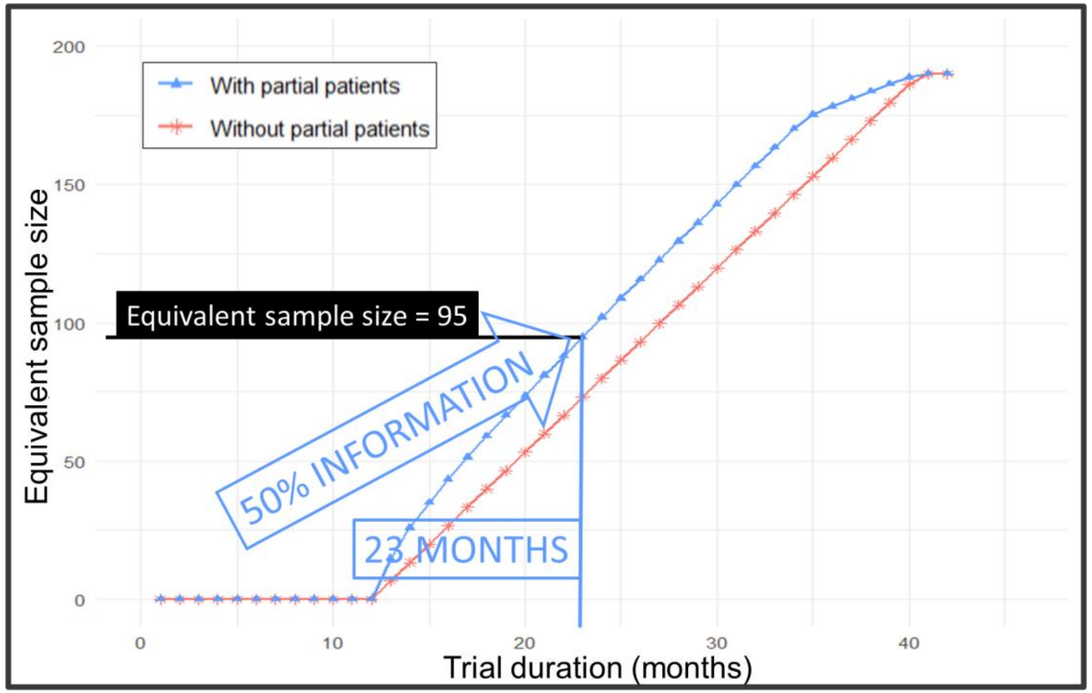
1. Design stage

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- Missing data from staggered entry and dropout $(P(dropout) = 0.05)$

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION? AT 23 MONTHS



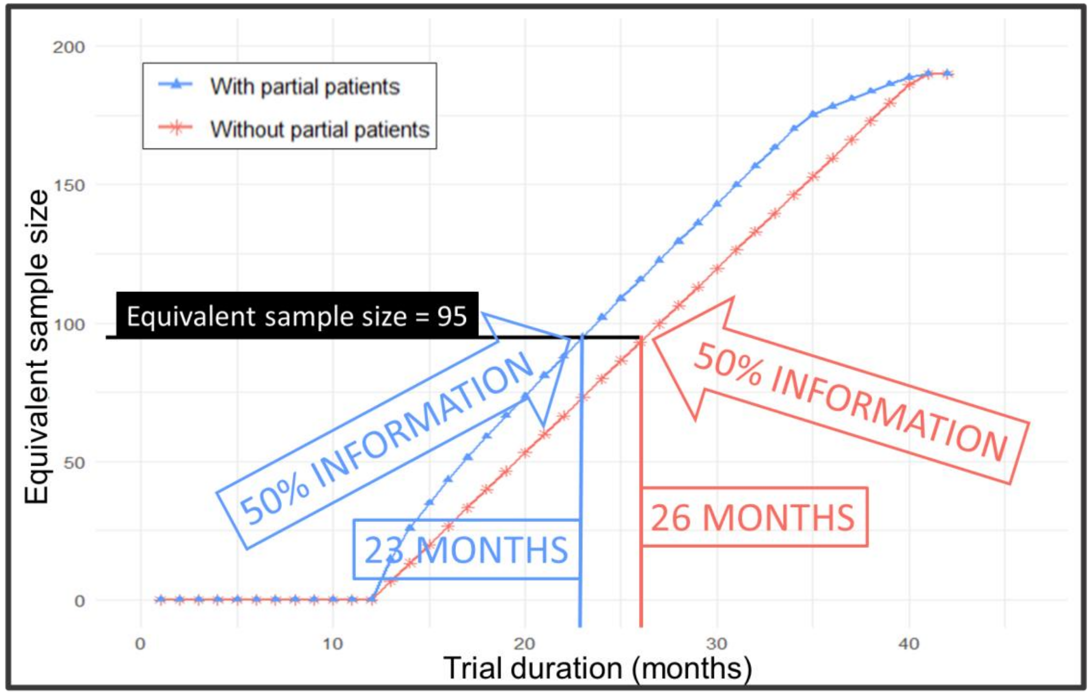
1. Design stage

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Assumptions

- Visit 1: 6 months
- Visit 2: 12 months
- Enrolment rate: 7/month
- Final equivalent sample size = 190
- Correlation between visits $\rho = 0.8$
- Missing data from staggered entry and dropout ($P(dropout) = 0.05$)

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION? AT 23 MONTHS

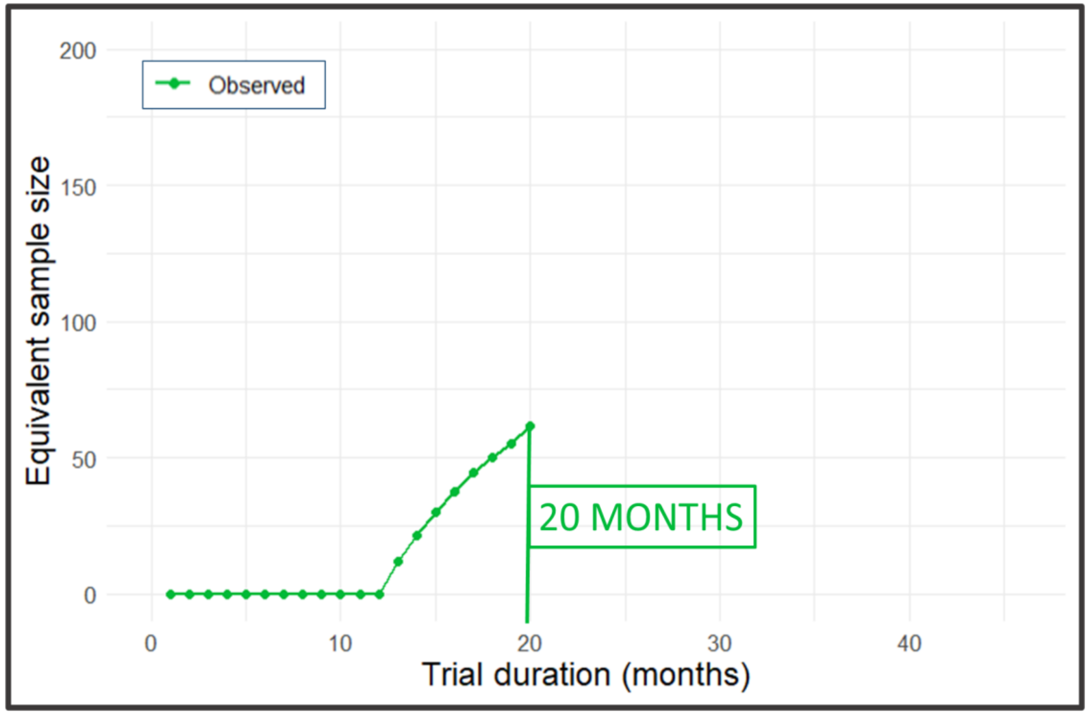


Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

2. After observing some data

Observations at 20 months

- Lower enrolment rate (6/month)
- Unforeseen missingness $P(miss\ Visit\ 1) = 0.11, P(miss\ Visit\ 2) = 0.02$
- $n_1 = 42, n_2 = 5, n_3 = 33$

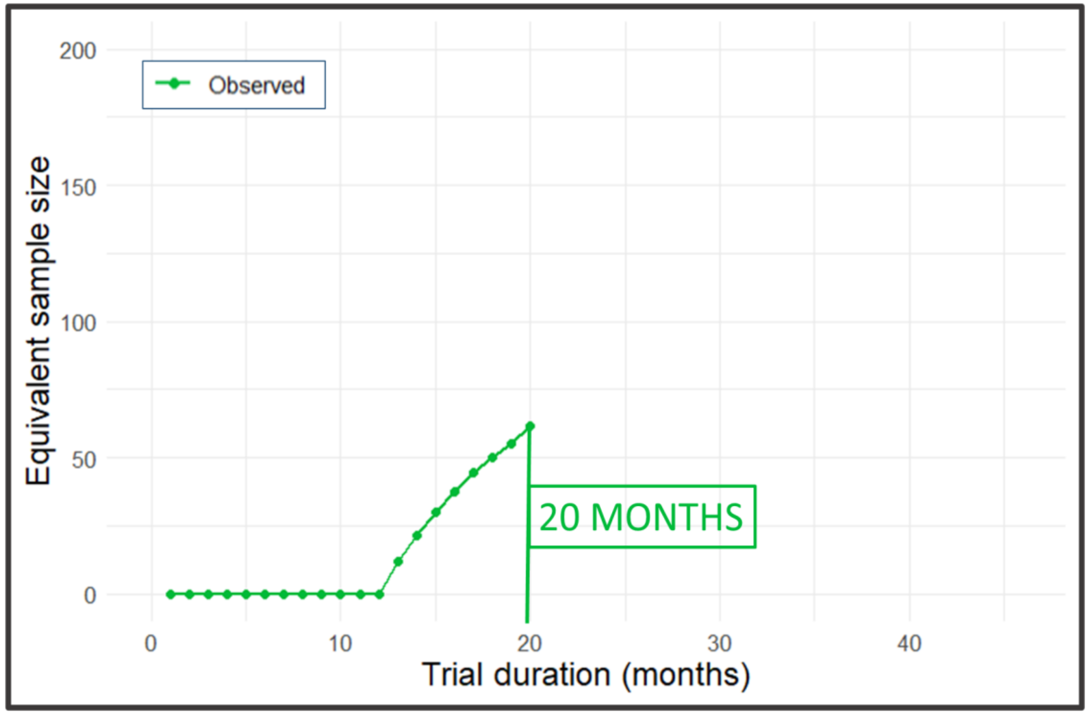


2. After observing some data

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Observations at 20 months

- Lower enrolment rate (6/month)
- Unforeseen missingness $P(miss\ Visit\ 1) = 0.11, P(miss\ Visit\ 2) = 0.02$
- $n_1 = 42, n_2 = 5, n_3 = 33$



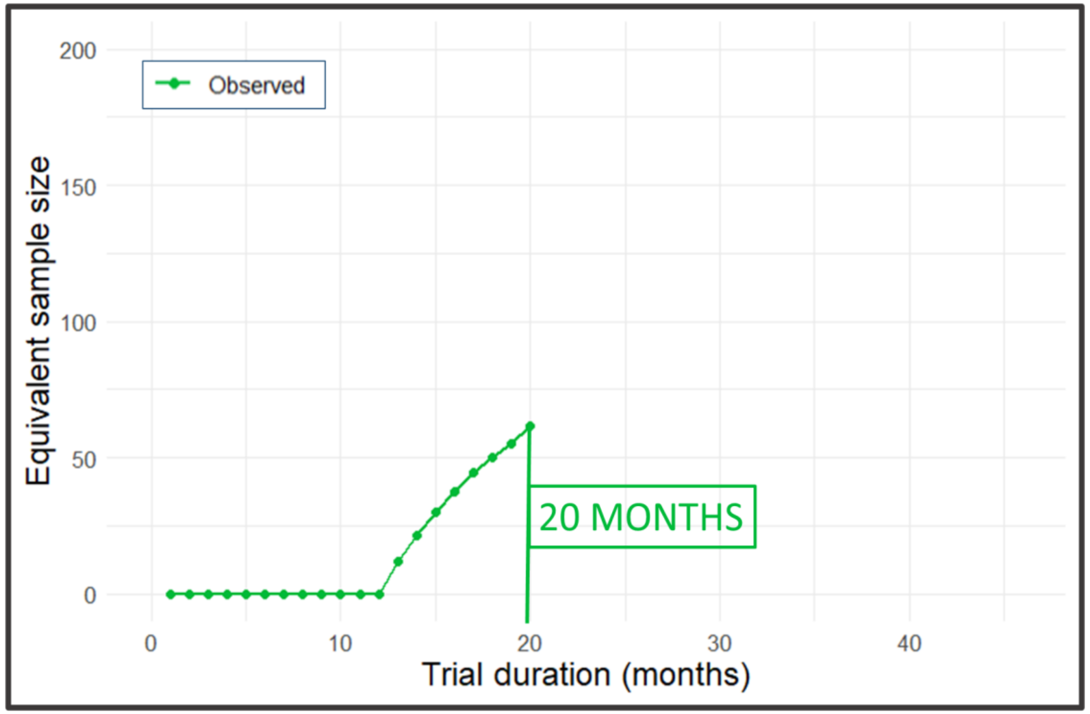
Assess current information level

2. After observing some data

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Observations at 20 months

- Lower enrolment rate (6/month)
- Unforeseen missingness $P(miss\ Visit\ 1) = 0.11, P(miss\ Visit\ 2) = 0.02$
- $n_1 = 42, n_2 = 5, n_3 = 33$



Assess current information level

→ incorporate prior belief

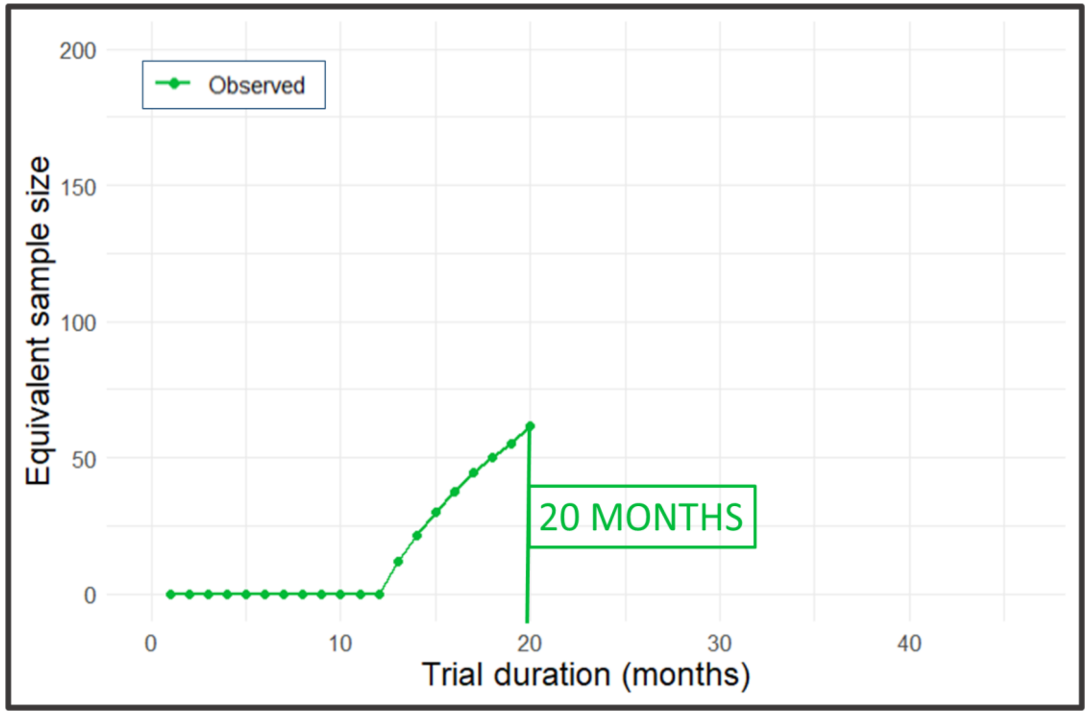
→ obtain updated probabilities

2. After observing some data

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Observations at 20 months

- Lower enrolment rate (6/month)
- Unforeseen missingness $P(miss\ Visit\ 1) = 0.11, P(miss\ Visit\ 2) = 0.02$
- $n_1 = 42, n_2 = 5, n_3 = 33$



Assess current information level

→ incorporate prior belief

→ obtain updated probabilities

→ Calculate:

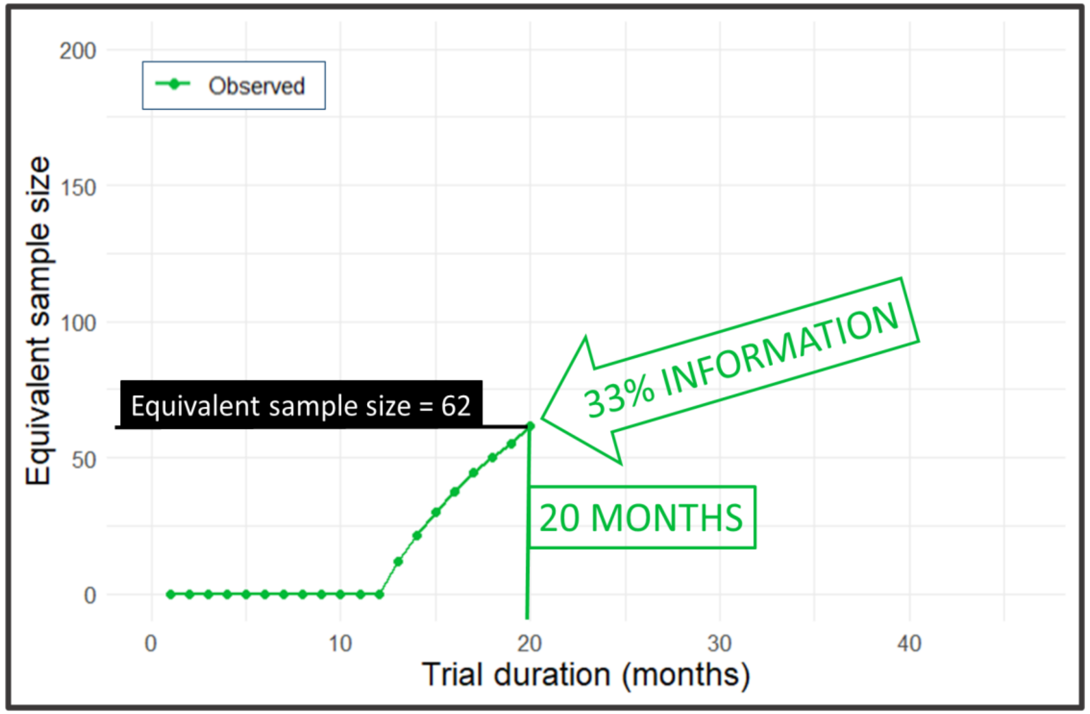
- current equivalent sample size = 62
- projected final equivalent sample size = 188

2. After observing some data

Pattern	Visit 1	Visit 2	
1	$\bar{X}_{11}$	$\bar{X}_{12}$	n_1
2	NA	$\bar{X}_{22}$	n_2
3	$\bar{X}_{31}$	NA	n_3
0	NA	NA	n_0

Observations at 20 months

- Lower enrolment rate (6/month)
- Unforeseen missingness $P(miss\ Visit\ 1) = 0.11, P(miss\ Visit\ 2) = 0.02$
- $n_1 = 42, n_2 = 5, n_3 = 33$



Assess current information level

→ incorporate prior belief

→ obtain updated probabilities

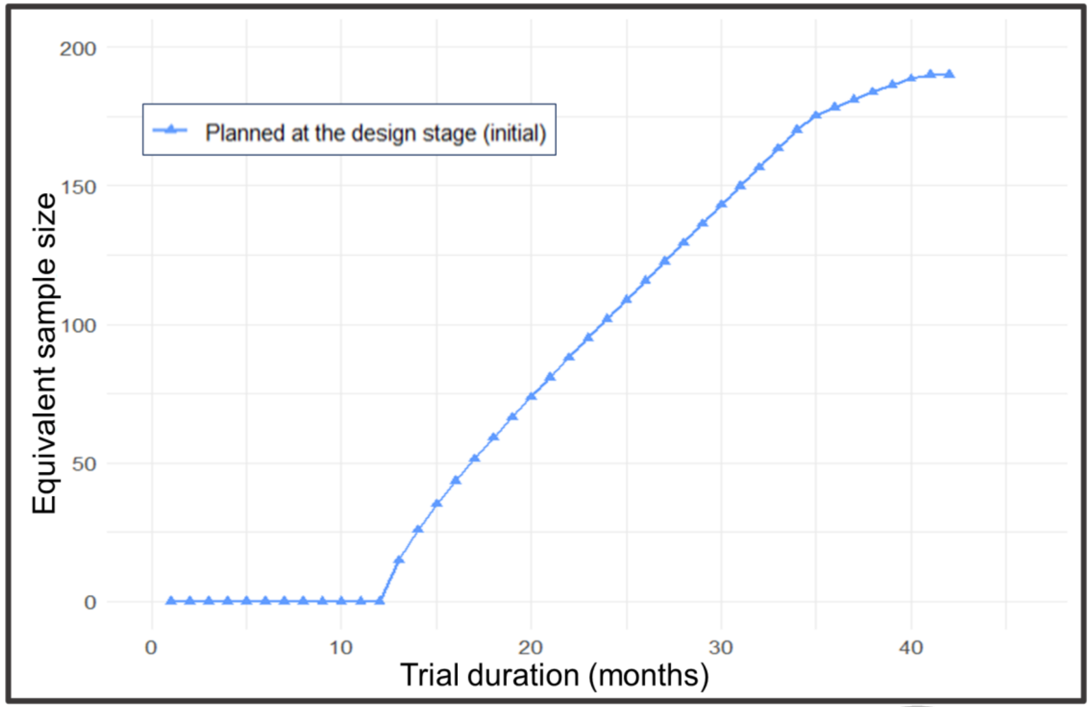
→ Calculate:

- current equivalent sample size = 62
- projected final equivalent sample size = 188

→ 62/188 = 33% information accrued

3. New projection

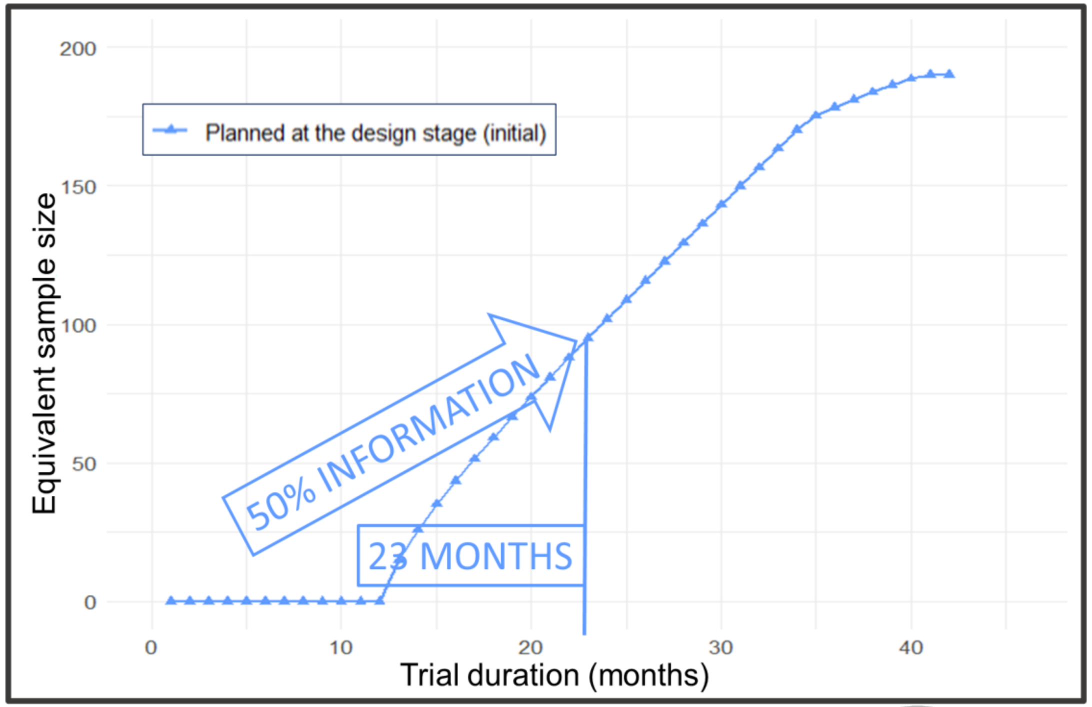
INITIAL PROJECTION



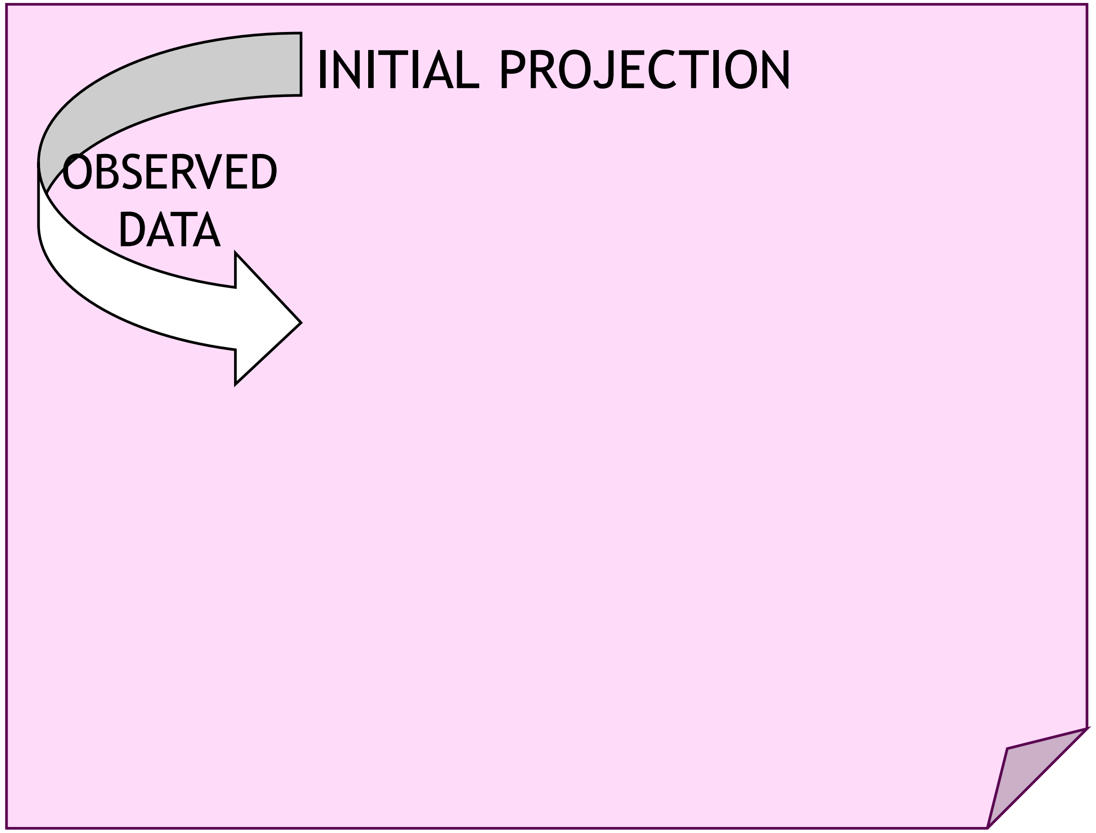
3. New projection

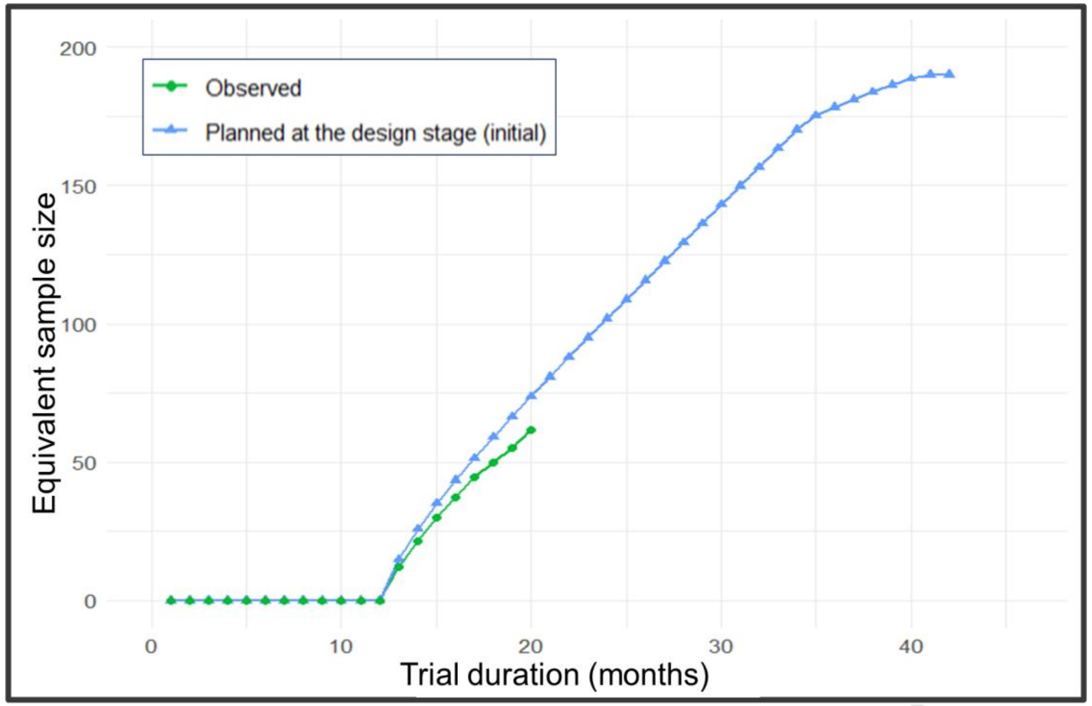
INITIAL PROJECTION

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION? AT 23 MONTHS



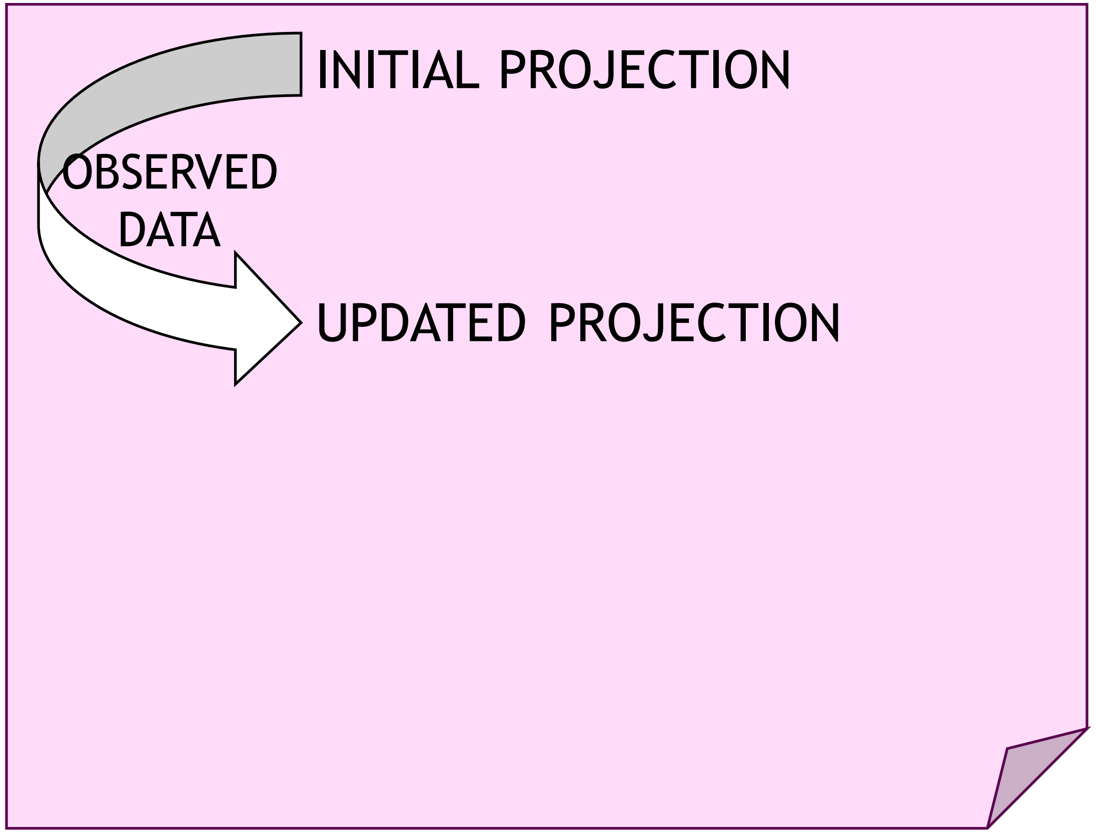
3. New projection

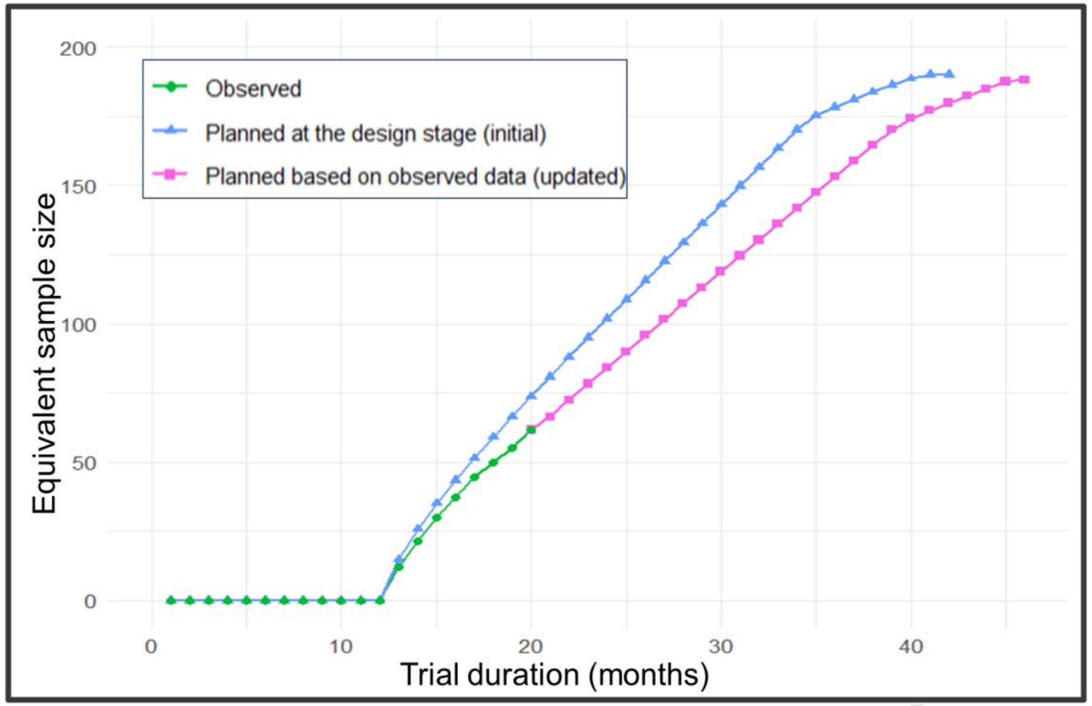




UNIVERSITY OF BATH

3. New projection





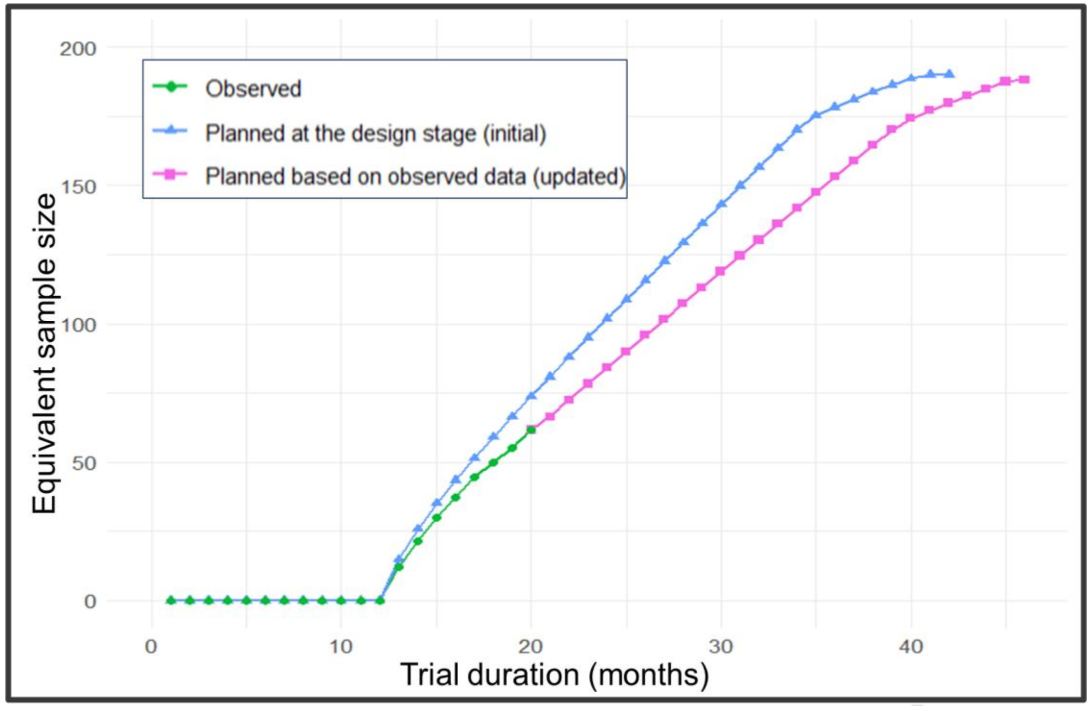
3. New projection

INITIAL PROJECTION

OBSERVED DATA

UPDATED PROJECTION

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION?



3. New projection

INITIAL PROJECTION

OBSERVED DATA

UPDATED PROJECTION

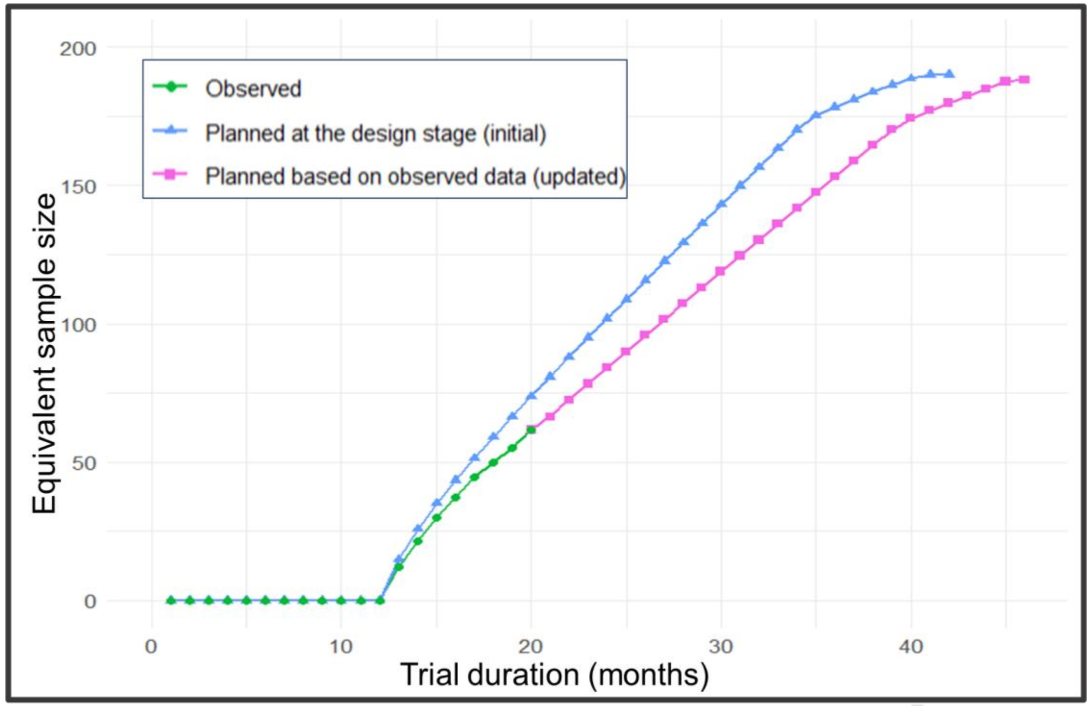
- Projected final equivalent sample size = 188

50% information

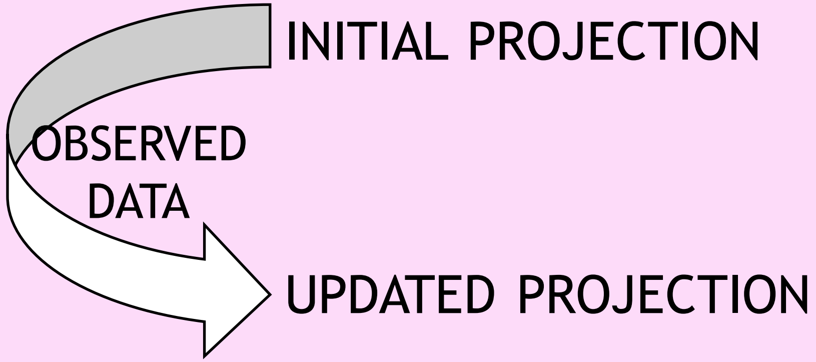
⇕

Equivalent sample size = 188/2 = 94

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION?



3. New projection



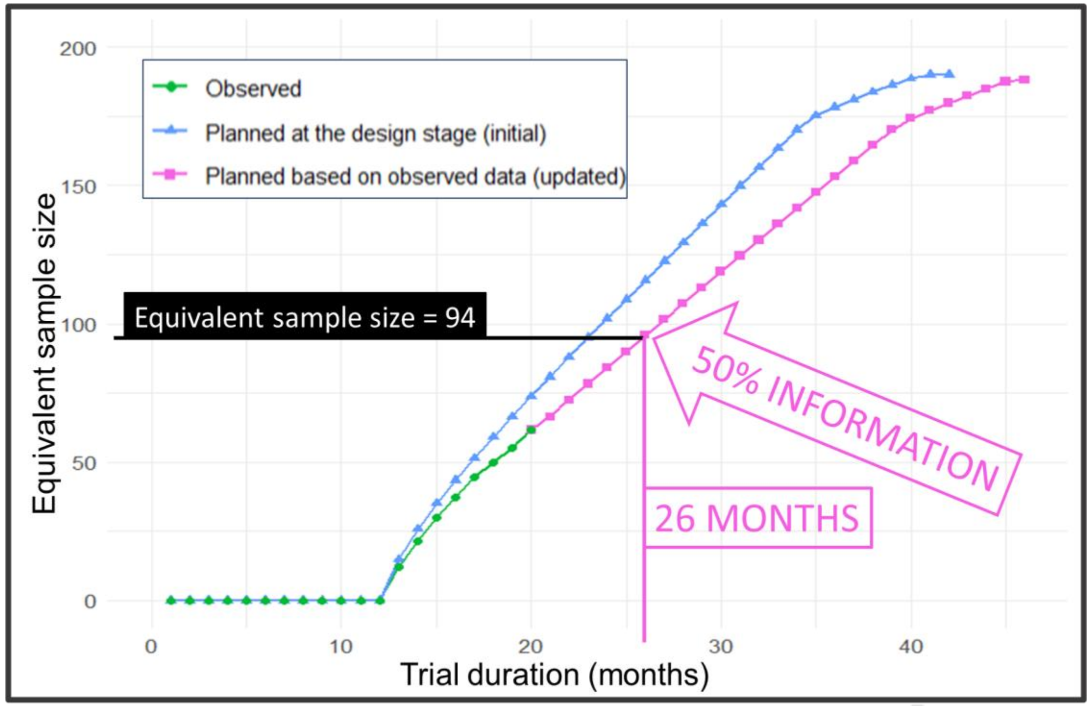
- Projected final equivalent sample size = 188

50% information

⇕

Equivalent sample size = 188/2 = 94

WHEN WILL WE REACH 50% OF THE FINAL INFORMATION? AT 26 MONTHS



UNIVERSITY OF BATH

Summary

Group individuals with the same missingness patterns and apply the framework to:

- make timeline projections based on assumptions at the design stage
- assess information level during the trial
- update timeline projections based on observed data

UNIVERSITY OF BATH

Summary

Group individuals with the same missingness patterns and apply the framework to:

- **make timeline projections** based on assumptions at the design stage
- assess information level during the trial
- update timeline projections based on observed data

Summary

Group individuals with the same missingness patterns and apply the framework to:

- make timeline projections based on assumptions at the design stage
- assess information level during the trial
- update timeline projections based on observed data

Summary

Group individuals with the same missingness patterns and apply the framework to:

- make timeline projections based on assumptions at the design stage
- assess information level during the trial
- update timeline projections based on observed data

UNIVERSITY OF BATH

Summary

Group individuals with the same missingness patterns and apply the framework to:

- **make timeline projections** based on assumptions at the design stage
- **assess information level** during the trial
- **update timeline projections** based on observed data

Extensions

- more visits → approximation of equivalent sample size
- different target information level with any type of interim analysis
- different assumptions at the design stage (enrolment rate, dropout rate, etc.)
- single-patient contribution to the information level

UNIVERSITY OF BATH

THANK YOU!

UNIVERSITY OF BATH