Sample Size considerations for the KIPRUN study

Márton Kiss, MD

2024-11-20

Table of contents

[Executive summary 2](#_Toc182996429)

[Assumptions for sample size calculation 2](#_Toc182996430)

[NPV value of the ABSOLUTE arm 3](#_Toc182996431)

[Enrichment 3](#_Toc182996432)

[Proportion of negative tests on the KINETIC arm 4](#_Toc182996433)

[Remarks 6](#_Toc182996434)

[References 6](#_Toc182996435)

# Executive summary

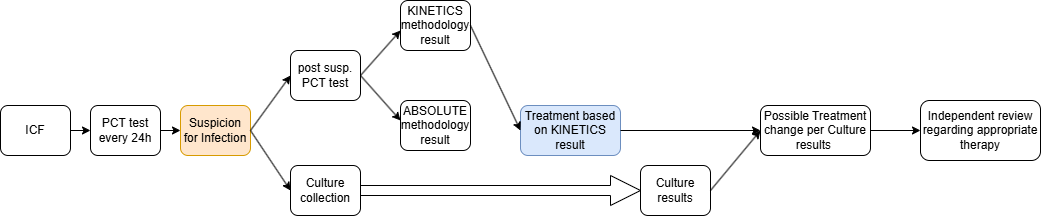
While with certain parameterizations, a sample size may be calculated, but the results are **very sensitive** to the assumptions made.

Since data is available from the previous observational study (n~200), it is *essential* to **review this data** to estimate the parameters of the study. The literature references do not provide information specific enough for our purposes, as in to accurately estimate the parameters of the study.

The review is necessary also to ensure that the test has the desired properties, as in high NPV and a negative rate not (much) greater than the expected ratio of real negative cases.

Thought may be given to **enriching the study population** by focusing on the subgroup most relevant to the research question, which in this case would be the patients with negative test results. This *may* allow for a more efficient use of resources if the gold standard evaluation of a case is more resource intensive relative to the enrollment of a patient.

A natural modification of the study design would be to use a **paired design**, as in, to enroll patients into one cohort only, and then evaluate their results per both methodologies (ABSOLUTE and KINETICS) but treating them as per the KINETICS results only. **If** the chart review section is unaffected by the methodology used to decide on the treatment, this setup would allow for a more direct comparison of the two methodologies while reducing the number of patients needed to enroll approximately by half. (Credit to *Tamás Kói* for the idea.)



Schematic representation for a paired design

As **next steps** we should discuss reviewing the raw data from the observational study and these ideas regarding the study design.

# Assumptions for sample size calculation

The following initial assumptions were made:

* Dropout rate: 10%
* Significance level (two-tailed): 5%
* Arm A (KINETICS) NPV: 85%
* Arm B (KINETICS) NPV: 40%
* Proportion of negative tests on arm A (KINETICS): 25%
* Proportion of negative tests on arm B (ABSOLUTE): 50%
* True proportion of negatives in the population: 25%

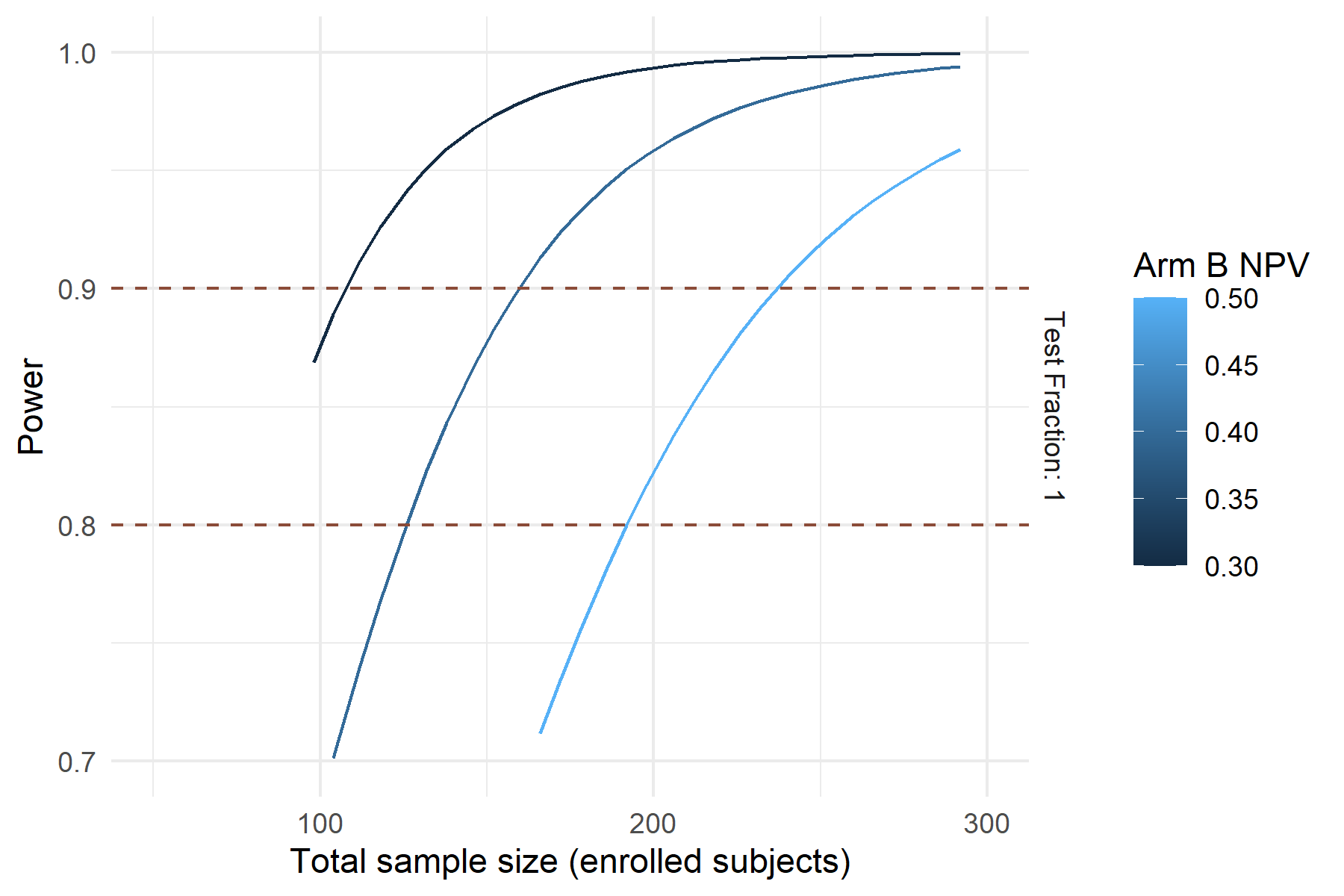
## NPV value of the ABSOLUTE arm

The below graph shows the power of the study (y axis) as a function of the **total** number of subjects enrolled in the study (x axis).

The color of the line represents the NPV of arm B (ABSOLUTE).

The figure shows the impact of the NPV value of arm B on the power of the study highlighting the possible impact of accurately identifying the parameters of the study.

To note, all of these parameters could be estimated from the previous observational study’s raw data.

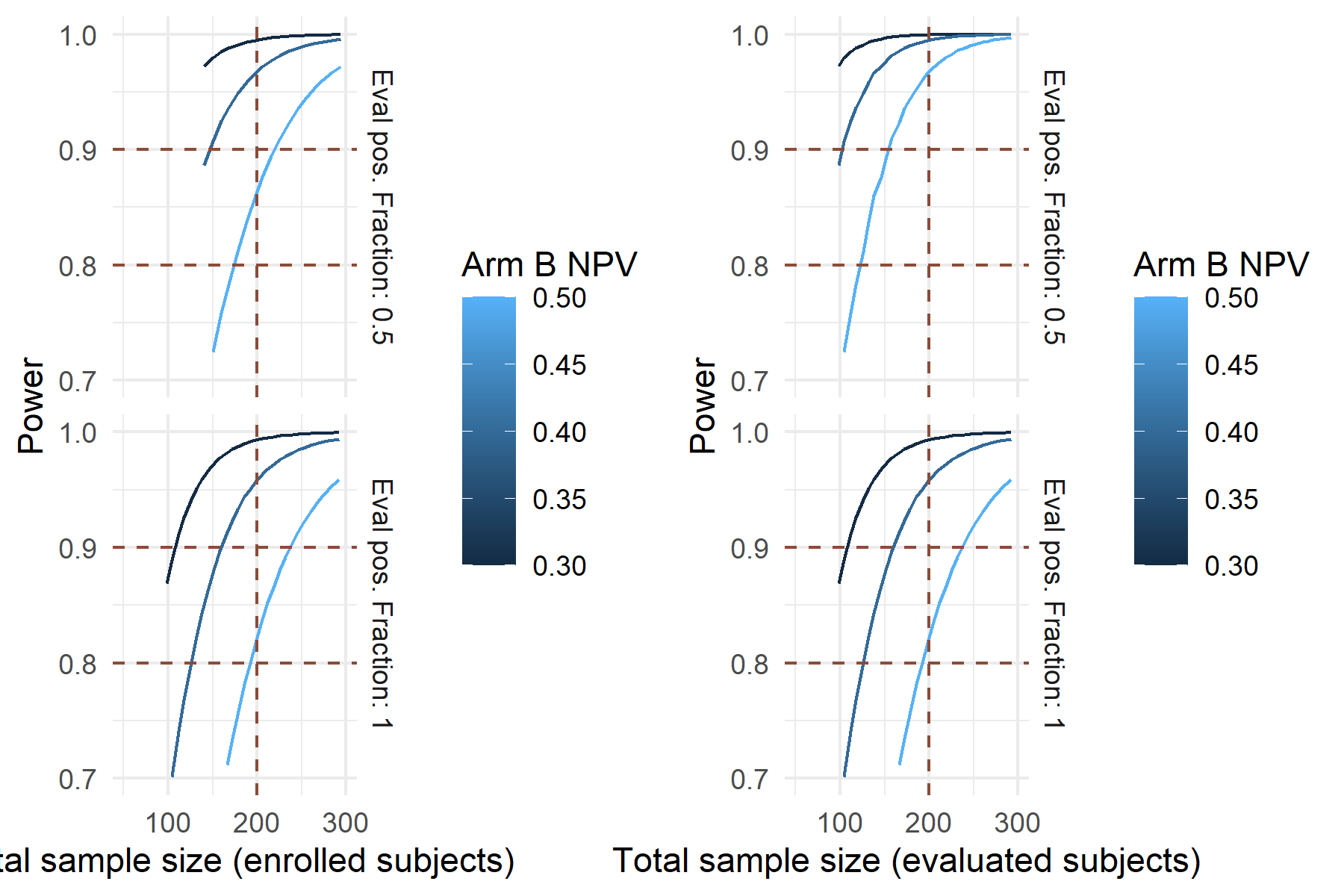


## Enrichment

An **enrichment design** in your context involves strategically focusing your study resources on the subgroup most relevant to your research question—the patients with negative test results, which are expected to be about 20% of your cohort. Since you’re interested in evaluating the **negative predictive value (NPV)** of the test, you’d include **all patients with negative test results** in your analysis to maximize data on this critical group. For the more prevalent positive test results (approximately 80% of cases), you might choose to evaluate only **50% of these patients**. By doing so, you enrich your study population with a higher proportion of negative cases, enhancing the efficiency and statistical power related to the NPV. This selective evaluation allows you to allocate resources more effectively, concentrating on the less common but most informative negative outcomes while still maintaining adequate data on positive cases to assess the overall test performance.

If the most labor-intensive part of the study is the evaluation of the treatment’s appropriateness, it may be more efficient to evaluate only a subset of the patients with positive test results.

The below graphs show that if the sample size is established as the number of patients who are *evaluated* rather than enrolled and the study population is enriched, then the study’s power would be greater (n=200 is highlighted for ease of comparison). This method is not usable if a paired design is used.

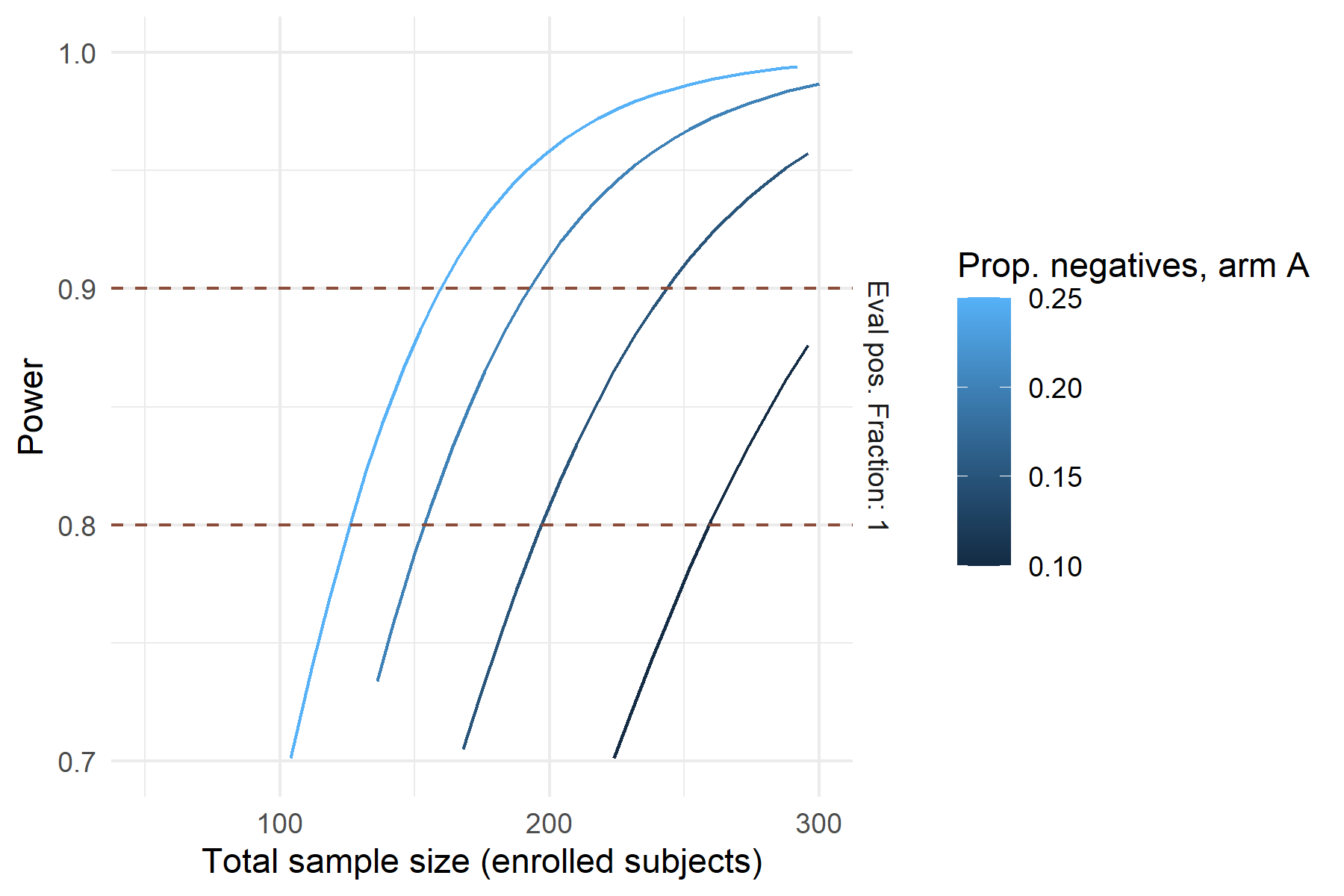


## Proportion of negative tests on the KINETIC arm

The below graph shows the power of the study (y axis) as a function of the **total** number of subjects enrolled in the study (x axis).

The different colors represent a different proportion of negative tests on arm A (KINETICS).

To note, there are theoretical upper limits on this variable if the NPV is to be kept at a constant level (here: 85% as discussed in the assumptions section).



# Remarks

### Information regarding the document’s compilation

Analyses were conducted using the R Statistical language (version 4.4.1; R Core Team, 2024) on Windows 11 x64 (build 22631), using the packages ggplot2 (version 3.5.1; Wickham H, 2016) and dplyr (version 1.1.4; Wickham H et al., 2023).

## References

* R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
* Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
* Wickham H, François R, Henry L, Müller K, Vaughan D (2023). dplyr: A Grammar of Data Manipulation. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.

### Time of compilation

2024-11-20 11:59:14.348472