**User Guide for R Model “Brainomix\_360\_Stroke\_CEA”**

*Version of this document: V1.2*

*Version of the publicly available model: V1.0*

*Version of R: R version 4.5.1 (2025-06-13 ucrt) -- "Great Square Root". Copyright (C) 2025 The R Foundation for Statistical Computing*

This document provides a guide on how to use the R model, including instructions on running the code, important considerations for inputs and outputs, and best practices. Link to the model: <https://github.com/OIpharmapartners/Brainomix_360_Stroke_CEA>

Contents

[1. Getting Started 1](#_Toc211466333)

[1.1 Prerequisites 2](#_Toc211466334)

[1.2 Installation 2](#_Toc211466335)

[2. How to use 3](#_Toc211466336)

[2.1. Generate the HTML report 3](#_Toc211466337)

[2.2. Run the core model code 4](#_Toc211466338)

[3. Outputs 8](#_Toc211466339)

[3.1 Understanding the Results 8](#_Toc211466340)

[3.2 Interpreting the Results 8](#_Toc211466341)

[4. Troubleshooting 8](#_Toc211466342)

[4.1 Potential issues and fixes to flag 8](#_Toc211466343)

[4.2 Built-in Validation Checks 8](#_Toc211466344)

[4.3 Contact for issues 8](#_Toc211466345)

## 1. Getting Started

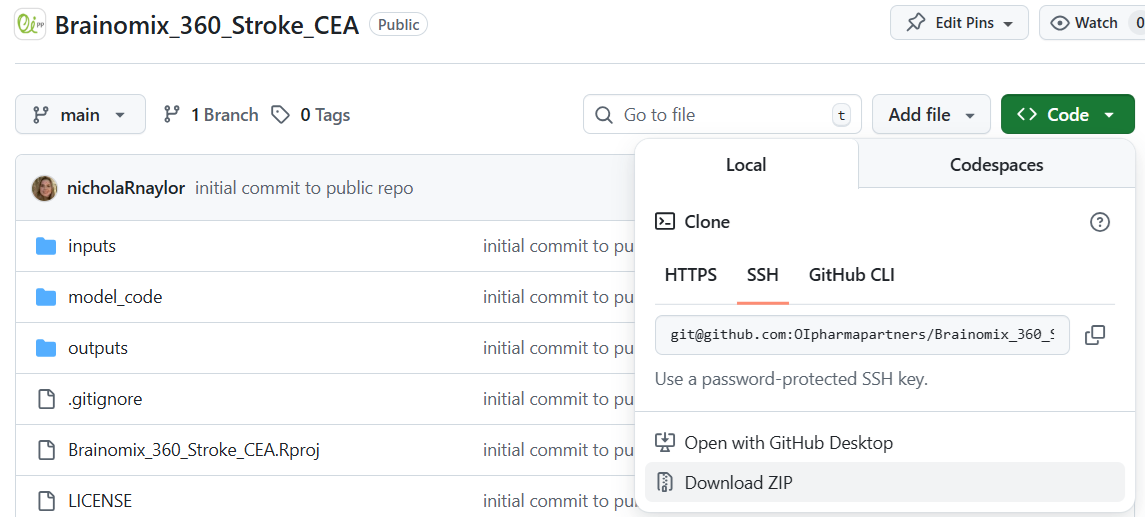
The model estimates health and economic outcomes associated with the use of AI-supported imaging along the stroke care pathway, compared to current standard care, at national, regional (ISDN), and local (hospital) levels.

The analysis includes:

* **Base-case deterministic model**
* **Deterministic sensitivity analysis (DSA)**
* **Probabilistic sensitivity analysis (PSA)**
* **Scenario analyses**
* and supports the generation of **cost-effectiveness planes, CEACs, tornado plots, Sankey diagrams**, and **mRS flow figures**.

[Place holder for manuscript citation]

To download the R project, you can link to Github via SSH or download the whole project as a zip folder by clicking on the “<>Code” button in the right hand corner:



### 1.1 Prerequisites

- Install R (≥ 4.3.0) and RStudio (optional but recommended).

- Install required R packages (see section “How to use” for a list)

### 1.2 Installation

Make sure you have installed R (≥ 4.3.0) and RStudio on your computer, see: [RStudio Desktop - Posit](https://posit.co/download/rstudio-desktop/) for more information. If you’re new to RStudio, try these resources:

* [rstudio-ide](https://rstudio.github.io/cheatsheets/rstudio-ide.pdf) [https://rstudio.github.io/cheatsheets/rstudio-ide.pdf]
* [A Installing R and RStudio | Hands-On Programming with R](https://rstudio-education.github.io/hopr/starting.html) [https://rstudio-education.github.io/hopr/starting.html]

Required packages include: ggplot2, data.table, parallel, assertthat (these are downloaded within the model script below so don’t worry you don’t need to do anything yet about this).

Ensure your working directory points to project root and that the zipped folder is unpacked.

Throughout the model and/or the document the following abbreviations will be used:

#### *Table 1. Abbreviations used throughout the model and documentation*

|  |  |
| --- | --- |
| Abbreviation | Meaning / Definition |
| ASC | Acute Stroke Centre – hospital providing immediate stroke assessment and intravenous thrombolysis but not on-site mechanical thrombectomy. |
| CSC | Comprehensive Stroke Centre – tertiary centre providing full stroke imaging and mechanical thrombectomy (MT) capability. |
| IVT | Intravenous Thrombolysis (tPA) – clot-dissolving medication given to eligible stroke patients. |
| MT | Mechanical Thrombectomy – catheter-based procedure to physically remove a clot in patients with large vessel occlusion (LVO). |
| B360S | Brainomix 360 Stroke – the AI imaging software being evaluated in this cost-effectiveness analysis. |
| mRS | Modified Rankin Scale – 7-point functional outcome scale after stroke (0 = no symptoms; 6 = death). |
| CTP | Computed Tomography Perfusion – imaging modality assessing cerebral blood flow and infarct core/penumbra. |
| CTA | Computed Tomography Angiography – imaging modality showing cerebral vasculature to identify LVO. |
| MRI | Magnetic Resonance Imaging – advanced imaging technique used to assess tissue viability or confirm stroke type. |
| NCCT | Non-Contrast Computed Tomography – baseline CT scan used to exclude haemorrhage and confirm ischaemic stroke. |
| PSA | Probabilistic Sensitivity Analysis – Monte Carlo sampling of uncertain parameters to quantify uncertainty in model results. |
| DSA | Deterministic Sensitivity Analysis – one-way or scenario-based parameter variation for robustness testing. |
| ISDN | Integrated Stroke Delivery Network – regional network of hospitals delivering coordinated stroke care in England. |
| NMB | Net Monetary Benefit – calculated as (Incremental QALYs × Willingness-to-pay) − Incremental Costs. |
| QALY | Quality-Adjusted Life Year – outcome measure combining survival and health-related quality of life. |
| WTP | Willingness-to-Pay – monetary value per QALY (commonly £20,000 per QALY, per NICE guidance). |
| LVO | Large Vessel Occlusion – blockage in a major cerebral artery often requiring MT. |
| EARLY / LATE | Presentation time windows – “EARLY” = within the early therapeutic window for IVT/MT; “LATE” = anything greater than the early threshold. |
| EMT | Eligible for Mechanical Thrombectomy – model state for patients triaged and sent for MT assessment. |
| NOEMT | Not Eligible for Mechanical Thrombectomy – model state for patients ruled out of MT after assessment. |
| AIS | Acute Ischaemic Stroke – the initial condition represented in the decision-tree portion of the model. |
| ASC / CSC ratio | Network configuration parameter – proportion of stroke patients initially presenting to ASC vs CSC hospitals. |
| dr | Discount rate – annual rate applied to future costs and QALYs in the Markov model. |

## 2. How to use

To open the project, before doing anything described below, click on the “Brainomix\_360\_Stroke\_CEA.Rproj” R project file. If you just click on individual files it will not recognise the script as being within a project and will have trouble loading sourced data and scripts.

### 2.1. Generate the HTML report

**What this does:** Compiles a HTML that reads precomputed results (no models are re-run here). You can switch between National / ISDN / Hospital level analyses.

#### a) Prerequisites

You already have these files saved by the model runs:

* outputs/base\_case\_results.RData → contains an object base\_case
* outputs/network\_results.RData → contains a named list network\_results (names = ISDNs)
* outputs/hospital\_results.RData → contains a named list hospital\_results (names = Hospitals)

R packages for rendering: you need to make sure you have the following packages downloaded and ready to go!

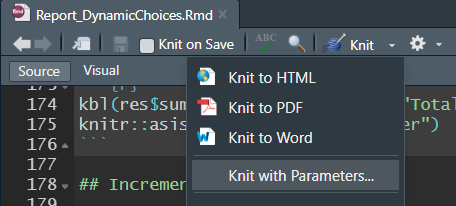
To generate the automated HTML report, the following packages are needed:

install.packages(c("rmarkdown", "knitr", "htmltools",”"kableExtra"))

You only need to do the package installing once on your computer.

#### b) Render from RStudio (interactive)

* Open the Report\_Generation.Rmd
* Click **Knit ▼ → Knit with Parameters…**



* A pop-up box should be created. Choose:
  + **Level:** National, ISDN, or Hospital
  + If **Hospital**, pick one from the dropdown (if this is not picked an error will occur and the html report will not update).
  + If **ISDN**, pick one from the dropdown (if this is not picked an error will occur and the html report will not update).
* Then click “Knit” in the bottom right of the box.
* If you just select “Knit” (not “Knit with Parameters”). The HTML will be created in the project directory (default based on “National” as the selection) & the selections will not be available.
* When complete, the file will be automatically available for you to click and open within the Rproject folder.
* **Report won’t open?** Check that a default browser is set and that the three needed .RData files exist in /outputs.

**Windows note:** The package “RTools” may be required only if R needs to build packages from source (e.g., when a precompiled binary isn’t available). For standard CRAN installs on Windows, R will use binaries and RTools is not typically required.

### 2.2. Run the core model code

**What this does:** Runs the acute decision tree and long-term mRS Markov, plus (optionally) network/hospital aggregations. Saves the objects the report (above)reads. You might want to do this if you update the values in the input files, such as the “inputs/parameters.csv” file.

#### a) Prerequisites

The following packages need to be installed, enter this into the console:

install.packages(c( "data.table", "tidyverse", "assertthat", "truncnorm", "conflicted", "scales", "reshape2", "networkD3", "htmlwidgets"))

#### b) Inputs expected

Table 2 below gives an overview of the files necessary to run the model code.

#### Table 2. Data Inputs Required

|  |  |  |  |
| --- | --- | --- | --- |
| File Name | Description | Variables | Current Reference Material |
| inputs/parameters.csv | A CSV file containing the parameter inputs for the model, including costs, probabilities, and other relevant data | Presentation/Setting; Intervention; Description; Reference(s); Calculations; model\_param; base\_case; PSA\_distribution; PSA\_low; PSA\_high; references/calculations for distributional information; DSA\_flag; DSA\_low; DSA\_high; mrs; age \*see below for details on each | Multiple data and literature sources (see file for individual references). |
| inputs/csc\_key.csv | A CSV file listing the names of CSCs, used to match to hospitals | Hospital (a list of the hospitals deemed CSCs) | Internal document |
| inputs/hospital\_names\_old.csv | A CSV file with the names of the hospitals used in the regional data | Hospital (hospital name) | Internal document |
| inputs/isdn\_key.csv | A CSV file containing the names of the ISDNs and the hospitals within them | ISDN (integrated stroke network name, e.g. London), site.name (hospital site name, e.g. Barking, Havering and Redbridge University Hospitals NHS Trust), Hospital (e.g. Queens Hospital Romford HASU), date (date range for stroke recordings), n.stroke (number of strokes recorded at that hospital for the stated time frame) | SSNAP “Apr2023Mar2024 AnnualResultsPortfolio.xlsx” |
| inputs/regional/isdn\_ivt.csv | A CSV file with the number of IVT patients from eligible IVT patients in each ISDN, used to calculate the number of patients per hospital at the national level | ISDN (integrated stroke network name, e.g. London), site.name (hospital site name, e.g. Barking, Havering and Redbridge University Hospitals NHS Trust), Hospital (e.g. Queens Hospital Romford HASU), date (date range for stroke recordings), p\_eivt2ivt( represents the percentage of eligible IVT patients that received IVT) | SSNAP “Apr2023Mar2024 AnnualResultsPortfolio.xlsx” |

These are the input csv files needed to run the model. These are used, by the model code, to create inputs which feed into the "created\_inputs" subfolder.

Note all data included and used in this repository are data that are publicly available or published data.

#### c) How to run the code

* Users can either open each file and run, or use “source(<insert name of file>)”.
* Files should be run in the order as listed in the hierarchy below.

The folder hierarchy is:

Brainomix\_360\_Stroke\_CEA/

├─ Rscripts/

│ ├─ model\_functions.R # core acute-tree + long-term Markov model

│ ├─ model\_functions\_mortality\_sc.R # scenario variant (mortality RR year 1)

│ ├─ model\_functions\_MTTPA\_sc.R # scenario variant (additive LT effects)

│ ├─ 1\_hospital\_data\_processing.R # prepares ISDN hospital input

│ ├─ 2\_probablistic\_sampling.R # parameter sampling for PSA

│ ├─ 3a\_deterministic\_model.R # base case + DSA

│ ├─ 3b\_probablistic\_model.R # PSA driver

│ ├─ 4\_scenario\_analysis.R # scenario analyses

│ ├─ regional\_level.R # ISDN-level analysis

│ ├─ local\_level.R # hospital-level analysis

│ ├─figure\_creation/ [contains R scripts creating figures, no order/hierarchy to run these]

├─ inputs/ [see relevant section for more information]

└─ outputs/ [see relevant section for more information]

#### d) Other notes on data and use regarding the model:

Provide guidelines on how to prepare and preprocess the data before feeding it into the model.

* In “1\_hospital\_data\_processing.R” :
* Dates: The ISDN datasets include a date label such as "Apr 2022–Mar 2023". Retain this exact format (month name plus year range, using a consistent en-dash or hyphen). If you use a different period label, update the filter in the script accordingly.
* Hospital naming: Hospital names in the input files must match exactly, including punctuation and apostrophes (e.g., Queen's Medical Centre - Nottingham). Avoid extra spaces and ensure files are encoded in UTF-8.
* Rehab units: Any rows with hospital names containing "Rehabilitation" or "Rehab" are removed. If your site list uses different wording, update the pattern in the script or source file.
* Region filtering: ISDN values "Wales" and "Northern Ireland" are excluded. Use these exact strings if you intend to drop those records.
* CSC flags:
* csc\_key.csv must include one row per hospital with a Hospital column.
* CSCs are flagged in the script, and any hospital whose name contains "HASU" is force-flagged as CSC.
* If your HASU naming convention differs, adjust the "HASU" pattern in the script.
* Percent vs proportion: isdn\_TPA\_E.csv must contain percent values in p\_etpa2tpa (e.g., 37.5, not 0.375). The script checks that the maximum value is greater than 1 as a sanity check and then divides by 100 when writing parameters.
* Missing/suppressed values: The script drops "Too few to report" and "." in numeric fields. Use these exact tokens for suppressed data; otherwise, adjust the filters accordingly.
* In “2\_probablistic\_sampling.R”
  + Distributions used are below, any other distributions would need additional flagging in the parameters file AND additional coding in the sampling files
  + Utilities (by mRS): truncated normal on [0,1] from mean (base\_case) + 95% CI (PSA\_low and PSA\_high)
  + Costs: gamma fit from mean (base\_case) + 95% CI (PSA\_low and PSA\_high)
  + RR mortality (by mRS): lognormal from mean (base\_case) + 95% CI (PSA\_low and PSA\_high)
  + Control probabilities: Beta(mean, k) — here k is stored in the parameter file under PSA\_low for Beta rows (not a CI bound).
  + Treatment probabilities: Intervention-arm probabilities are derived from control probabilities and sampled ORs using the standard odds transformation (p\_treat = OR \* odds\_control / (1 + OR \* odds\_control)), with bounds to keep probabilities strictly (0,1).
  + Deterministic vs PSA values: The PSA creates full sample sets for utilities, costs, and mortality by mRS; deterministic values used by the base model come directly from the parameter file (base\_case), not from PSA medians.
  + mRS conventions: mRS is treated as an integer key (0–6) across parameter files and PSA outputsUtilities and costs for mRS 6 are set to 0; RR for mRS 6 is NA (dead state, not used). Ensure your parameter file has mRS recorded as integers 0–6.
  + Log-normal PSA (mortality RRs): We parameterize the log-normal to match the reported 95% interval exactly (2.5% and 97.5% quantiles). The implied mean may differ if the published mean and CI are not exactly log-normal–consistent; the script warns when the discrepancy exceeds 5%.

## 3. Outputs

### 3.1 Understanding the Results

After you run the model scripts, results are written to the /outputs folder. At a glance:

* Core base-case bundle (from 3a\_deterministic\_model.R):

- outputs/base\_case\_results.RData → contains base\_case, a list with:  
 - summary\_data\_all: totals by strategy (costs, QALYs)  
 - incremental\_results: key incremental outcomes and Net Monetary Benefit (NMB)  
 - process\_results: pathway/process counts and component costs/QALYs  
 - mrs.trace: Markov trace (mRS0–mRS6 over cycles)  
 - trace.standard / trace.intervention: decision-tree traces  
 - tm.standard / tm.intervention: transition arrays (per cycle)  
- Convenience CSVs: population\_results.csv, NMB.csv, proc\_results.csv, mrstrace\_results.csv

- DSA\_results: results from the deterministic sensitivity analysis

- scenario\_incremental\_results.csv: Scenario results

* Regional & local aggregation:

- outputs/network\_results.RData (ISDN-level list of model results) and outputs/ISDN\_incremental\_results.csv  
- outputs/hospital\_results.RData (hospital-level list of model results) and outputs/hospital\_incremental\_results.csv

* Probabilistic Sensitivity Analysis (PSA):

- outputs/psa\_outputs.csv: per-simulation totals (costs/QALYs by arm) plus incremental values  
- outputs/CEAC.long.csv and outputs/plot.ceac.all.png: CEAC table and plot derived from PSA

* Figures (generated from figure scripts using the above files):

- CE plane: outputs/plot.ce.plane.png (incremental cost vs QALY)  
- CEAC: outputs/plot.ceac.all.png (probability cost-effective vs WTP)  
- mRS flow: outputs/mrs\_plot.png  
- Pathway Sankey: outputs/standard\_sankey\_plot.html, outputs/intervention\_sankey\_plot.html

* Report inputs:

- The HTML report (Report\_Creation.Rmd) reads base\_case\_results.RData, network\_results.RData, and hospital\_results.RData.

### 3.2 Interpreting the Results

This section focuses on the two metrics most commonly requested by decision-makers: incremental results and Net Monetary Benefit (NMB).

• Incremental outcomes:

The model computes incremental QALYs and incremental costs as:

ΔQALY = QALY\_Intervention − QALY\_Standard

ΔCost = Cost\_Intervention − Cost\_Standard

Positive ΔQALY indicates health gain; negative ΔCost indicates cost savings.

• Net Monetary Benefit (NMB):

NMB translates QALYs and costs into a single monetary metric given a Willingness-to-Pay (WTP) threshold (e.g., £20,000/QALY):

NMB = (ΔQALY × WTP) − ΔCost

Interpretation:

- NMB > 0 → intervention is cost-effective at that WTP

- NMB < 0 → not cost-effective

• From PSA plots:

- CE plane: each point is a PSA draw of (ΔQALY, ΔCost); the line with slope = WTP (e.g., 20,000) separates cost-effective from not.

- CEAC: shows the probability that NMB > 0 across WTP values; e.g., CEAC of 0.85 at £20k/QALY means 85% chance cost-effective.

## 4. Troubleshooting

### 4.1 Potential issues and fixes to flag

* Throughout the code “!!!” is added to try and flag where potential issues could arise or where things should be checked if the model is being rerun on other data.
* In “1\_hospital\_data\_processing.R” it is hard-coded which year filter is used with “isdn <- isdn[date=="Apr 2022-Mar 2023"]” . If a later refresh lands, you’ll unknowingly drop newer data.
* In “1\_hospital\_data\_processing.R” it overrides csc flags in the csv file based on criteria such as “HASU”, so please check csv and edit codes used if hospital names change, networks change or descriptors change.

### 4.2 Built-in Validation Checks

The model includes defensive assertions for:

* **n.stroke** is a single non-negative integer.
* **Probabilities** are within [0,1].
* **Costs** and **utilities** are non-negative.
* **Mortality tables** extend beyond the specified start age.
* **Transition matrices** are checked for row sums = 1 and for any out-of-range cell values.

### 4.3 Contact for issues

* Please feel free to raise an issue on this GitHub, using the Issues functionality on the GitHub Repository, though the funding this project has been finished so responses and capacity for changes may be limited.
* Alternatively please contact Nichola.naylor@oipharmapartners.com