**General setup of parameters and functions**

1. Install required packages
2. Run required packages
3. Specify directory containing necessary files
4. Set number of cores for parallel computing
5. Start timer to determine length of simulation
6. Call R script containing required functions (see separate text algorithm for functions)
7. Set inner and outer loop numbers. Outer loop is to sample individual risk factors and needs to be set at ~7million for stable results from the most variable screening approach currently. Inner loop is for random tumour and other variable draws and initially set at 1.
8. Input programme parameters:
   1. Set screening strategy
      1. Procas
      2. Risk tertiles
      3. 3 yearly
      4. 2 yearly
      5. 5 yearly
      6. 2 rounds (50 and 60) other=no screening
   2. Whether supplemental screening offered (0/1)
   3. Set age at start of screening (default 38)
   4. Set time horizon (default 100 years)
   5. Set health and cost discount rates (default 0.035)
9. Set fixed (non PSA varying) parameters
   1. Proportion of cancers detected by screen
   2. Mean and sd doubling rate of screen detected tumours
   3. Set parameters of weibull survival curve for all-cause mortality
   4. Set breast cancer survival for patients with NPI cat 1,2 or 3 tumours based on exponential survival curves
   5. Read in csv sheet for incidence and mortality by age
   6. Read in csv sheet of observations of breast density, 10yr breast cancer risk, and lifetime risk from synthetic dataset based on observations from PROCAS 2 study
   7. Define data.frame of probability of metastatic cancer at different ages
   8. Define proportion of tumours which are ductal carcinoma in situ (baseline 0.21)
   9. Define data.frame which of probability a cancer is different NPI grades depending on size
   10. Define data.frame of baseline utility weights by age
   11. Define mean and standard deviation of tumour doubling when cancers identified clinically
   12. As above but for screened cancers
   13. Set parameters for mammography sensitivity conditional on size of tumour
   14. Set vector of sensitivities by VDG and a single average parameter
   15. Set cancer detection rate per 1000 women for mammography, MRI after mammography and US after mammography
   16. Set parameters (mean and sd) of log normal tumour growth curve
   17. Set maximum diameter of tumour in mm
   18. Set start size of tumour in mm
   19. Set starting volume and maximum volume of tumour based on q and r
   20. Set survival terms of metastatic cancer for different age ranges
   21. Set screening start age and end age
   22. Set intervals for screening depending on whether intervals are 1,2 or 3 yearly for high, medium, or low risk respectively
   23. Set maximum screening sensitivity
   24. Set the risk level cut-offs for the PROCAS and risk tertiles strategies
   25. Set the breast density (VDG) cut-off for supplemental screening
10. Set cost parameters
    1. Cost of stratification mechanism
    2. Cost of mammography
    3. Cost of follow up
    4. Cost of biopsy
    5. Cost of treating ductal in situ carcinoma
    6. Cost of treating NPI good, moderate, poor or metastatic cancer
    7. Cost of US
    8. Cost of MRI
11. Set false-positive and overdiagnosis parameters
    1. Set recall rate
    2. Set biopsy rate
12. Set utility parameters
    1. Set utility levels for metastatic tumours (same in every year)
    2. Set utility level for DCIS
    3. Set vector of utility levels for tumours with different NPIs in year 1
    4. Set vector of utility levels for tumours with different NPIs in following years

**Outer I loop for individual patient variables**

1. Set loop to divide outer loop into 10 smaller loops in case of simulation break
2. Set counters for:
   1. Total screens
   2. Total cancers detected
   3. Total costs
   4. Total US costs
   5. Total MRI costs
   6. Total life years
   7. Total US
   8. Total MRI
   9. Total QALYs
   10. Totals costs of follow up
3. Set up foreach command for parallel processing and to combine results
4. Set up an empty vector called cancer diagnostic to record age, size, mode of detection of each detected cancer
5. For the individual, create a vector which draws an observation of breast density, 10 year, and lifetime risk from the synthetic dataset
6. Set a variable, the individual’s 10 year risk, which is a percentage
7. If a risk based screening strategy is being used, place the individual into a risk group based on their ten year risk and the risk cut-offs for the strategy
8. Assign woman to a Volpara density group based on breast density estimate
9. Determine whether additional screening is taking place based on whether supplemental screening variable is “on” and breast density group
   1. If supplemental screening is off then MRI and ultrasound off
   2. If not, if density is greater than the density cutoff and 10 year risk is greater than 8% then turn MRI screening variable on, if risk less than 8% then turn ultrasound variable on
   3. If density is less than cutoff then set MRI and US to 0
10. Determine screening times
    1. Set up a screening time vector
    2. If Procas screening used then set screening time equal to low risk screen time variable for risk groups 1 or 2, equal to medium risk times for risk groups 3 or 4 or high risk screening times for risk group 5
    3. If risk tertiles used then set screening time equal to low risk screening time for group 1, medium risk screening time for group 2, high risk screening time for group 3
    4. If 3 yearly screening used then use low risk screening times
    5. If 2 yearly screening used then set screening time to medium risk screening times
    6. If 5 yearly screening used then create a sequence of screen times every 5 years from the start of screening age
    7. If 2x screens 10 years apart then set starting age as in f and create a sequence with one additional screen at 10 years and a maximum length of 10 years (i.e. only one more screen)
11. Set counters for I loop
    1. Number of screen detected cancers
    2. Number of cancers detected at first screen
    3. Numbers of cancers detected at last screen
    4. Total number of screens
    5. Number of people who had their last screen
    6. Number of ultrasounds
    7. Number of MRIs
    8. Number of recalls
    9. Total cost
    10. Total life years
    11. Total QALYs
    12. Number of cancers in each NPI group (5 counters)
12. Set variables for j loop
    1. Assign a variable mort\_sample which is an age of death drawn from a Weibull distribution using pre-specified distribution descriptors
    2. Assign a variable clin\_detect\_sample which is a size of clinically detected cancer drawn from a normal distribution described by pre-specified distribution descriptors
    3. Set min and max sizes of clinically detected cancer of 4 and 8.99

**J Loop for cancer screening simulation**

1. Open j loop
2. Set j loop counters at zero:
   1. Number of screens
   2. Number of recalls
   3. Cancer detected at last screen
   4. Last screen has occurred
   5. Cancer detected at first screen
   6. NPI category
   7. Number of MRIs
   8. Number of US
   9. Age at cancer occurrence
   10. Costs
   11. US costs
   12. MR costs
   13. Follow up costs
3. Determine if patient will have cancer by comparing a random number from a uniform distribution with the lifetime risk for the person. If rand number is lower than lifetime risk:
   1. Mark a variable names cancer case as 1
   2. Determine the growth rate for the cancer by randomly drawing from a log normal distribution described by pre-specified parameters
   3. Call the cancer incidence function to determine when a cancer occurs, whether it is screen or clinically detected, what size it is, age of death due to all-cause mortality
   4. Define a variable cancer\_incidence\_age which is assigned from the results of c and represents the age at which cancer diagnosed
   5. As above but for cancer\_size
   6. If the cancer is screen detected then set the clinical detection age at the cancer incidence age
   7. If the cancer was not screen detected set clinical detection age as the cancer incidence age + the time required for the cancer to grow to the clinical detection size
   8. Record the clinically detected size in the cancer\_diagnostic results vector
   9. Assign a parameter time of genesis which is calculated by determining how long it would take for the cancer to reach the size at clinical detection
   10. Calculate a variable for the age of genesis by and subtracting i from the age of clinical detection
   11. If the person did not have a cancer then set cancer case variable as 0
   12. Set cancer incidence age as 999 and clinical detection age as 999 (this is redundant but ensures that both occur after the end of the simulation if they are called)
4. Define a variable mort\_age which is the age of all cause mortality taken from mort\_sample for person j
5. Make sure mortality age isn’t lower than starting age of the sample and if so reassign based on the tail of the distribution above the starting age
6. If the individual will get cancer but they are expected to die before their cancer will start then reassign the death age based on the tail of the distribution above the cancer start age
7. If individual will die after time horizon of model set age of death to 99.99
8. Assign death age to the 7th value in the cancer diagnostic vector
9. Set a variable age equal to start\_age
10. Set a variable interval\_ca as equal to 0 for a cancer detected between screens
11. Set a variable screen\_detected\_ca to 0 for cancer detected at screen

**DES component**

1. Set the time to next screen (new variable) as the first value in the screen times vector minus age
2. Set the time to death (new variable) as the mortality age minus age
3. Set the time to cancer diagnosis (new variable) as the cancer diagnosis age minus age
4. Open while loop for while age is less than death age, screen and interval detected cancers are 0 (i.e. person is alive and cancer free):
   1. Assign a variable event list which is a vector of time to next screen, time to death and time to cancer diagnosis
   2. Assign a variable event\_place which is the next event (nearest in time) i.e. the min of the event\_list
   3. Assign a variable next event time which is the time until the next event
   4. Assign a variable current discount which calculates the current discount rate for the next event given the current age, time to next event and starting age of the woman
   5. If event place is 1 (the next event is a cancer screen):
      1. Add one to the screen count counter
      2. Add the cost of a screen (discounted at current time) to the costs counter
      3. If the screen count variable is equal to the length of the screen times variable then add one to the last screen counter
         1. If US screening is 1 then:
            1. Add 1 to the US screen counter
            2. Add the discounted cost of a US to the costs counter
            3. Add the discounted cost of a US to the US costs counter
         2. If MRI screening is 1 then:
            1. Add 1 to the MRI screen counter
            2. Add the discounted cost of an MRI to the costs counter
            3. Add the discounted cost of an MRI to the MRI costs counter
      4. If the event place variable is 1 (next event is a cancer screen) and the individual will develop a cancer at some point then:
         * 1. Determine if tumour is present at this time with a new variable t which is age plus next\_event time – gen\_age
           2. If t is greater than 0 (the tumour is present then):

Define cancer size based on a growth formula which predicts tumour volume at time t

Redefine size based on second formula

Call screening result function with inputs cancer size, breast density, MRI and US screening

If screening result =1 (cancer diagnosed at screen) then:

Add one to the screening result counter

Add age of screening diagnosis to cancer diagnostic vector position 1

Add size at diagnosis to cancer diagnostic vector position 3

Add a value of 1 to cancer diagnostic vector position 4

Add the value at the 4th position of the screen result function output to cancer diagnostic vector position 5

Add the value at the 3rd position of the screen result function output to cancer diagnostic vector position 6

Add the screen count to the cancer diagnostic position 10

Create a parameter incidence time which is equal to age plus time to screen

Add the discounted cost of follow up to the cost counter

Add the discounted cost of follow up to the cost of follow up counter

If screening result equals 1 and screen count equals 1 then add one to the detected at first screen counter

If screening result equals 1 and screen count is equal to the length of screening times then add one to the detected at last screen counter

If screening result equals 0 then assign a value of 0 to the screen detected counter

* + 1. If event place is 1 (next event is screen) and screen detected equals 0 (cancer not detected at screen) then check to see if a false-positive occurred by comparing a random number to the recall rate. If it does then
       1. add one to the recall count counter
       2. Add a cost of follow up plus an expected cost of biopsy (cost of biopsy\*biopsy rate) discounted at current rate to the costs and costs of follow up counters
  1. If event place =3 (the next event is a cancer diagnosis)
     1. Add one to the interval cancer counter
     2. Set the incidence age record as age+ time to cancer diagnosis
     3. Add the discounted cost of follow up to the cost counter
     4. Add the age of cancer diagnosis to the cancer diagnostic vector position 1
     5. Record clinically detected size to cancer diagnostic vector position 3
  2. If a cancer has been detected by screen or clinically:
     1. Update age to age plus next event time
     2. If an interval cancer detected then recode ca size as clinically detected size
     3. Set an NPI category by calling the NPI\_by\_size function with inputs cancer size and whether cancer screen detected
     4. If NPI category equals 1:
        1. Add 1 to NPI cat 1 counter
        2. Add discounted NPI good cost to costs counter
     5. If NPI category equals 2:
        1. Add 1 to NPI cat 2 counter
        2. Add discounted NPI moderate cost to costs counter
     6. If NPI category equals 3:
        1. Add 1 to NPI cat 3 counter
        2. Add discounted NPI poor cost to costs counter
     7. If NPI category equals 4:
        1. Add 1 to NPI cat 4 counter
        2. Add discounted DCI cost to costs counter
     8. If NPI category equals 5:
        1. Add 1 to NPI cat 5 counter
        2. Add discounted metastatic cost to costs counter
     9. Generate a cancer specific survival time (as a new variable) by calling the ca\_survival\_time function with inputs NPI category, age of all cause death, age, clinical diagnosis age
     10. Replace age of death with cancer age of death
  3. Else (if no cancer detected) replace age with age + next event time
  4. Update times for next event:
     1. If screen count is less than the length of screen times then replace time to screen with the next screening time from the screen times vector minus current age. Else set next screen time to 101 (i.e. no more screens in time horizon)
     2. Replace time to death with mortality age minus age
     3. Replace time to clinical diagnosis with cd\_age minus age
  5. End while loop

1. If no cancers detected assign age of death to cancer diagnostic vector in position 1
2. Add screen detected cancer to total screen detected count
3. Add screen count to total screen counter
4. Add US count to total US counter
5. Add MRI count to total MRI counter
6. Add recalls to total recall counter
7. Add cancer detected at first screen to total first screen detected counter
8. Add cancer detected at last screen to total last screen detected counter
9. Add last screens to total last screen counter
10. Add life years (death age minus start age) to total life year counter
11. QALY counter:
    1. Define QALY length as the ceiling for age of death minus start\_age
    2. Make sure QALY\_length is at least 1
    3. Make sure QALY\_length is not greater then time\_horizon minus start age
    4. Create a QALY vector of length QALY length, filled with 0’s
    5. Open y loop:
       1. Assign in QALY vector position y, the utility value matching that age band, discounted to the current year
       2. Close loop
    6. If a cancer occurred:
       1. Replace the utility values in the QALY vector in the year after incidence age with the utility values corresponding to the NPI category of the tumour (current utilities \* NPI utility odds ratios)
       2. If individual survived more than 1 year after diagnosis, replace values in QALY vector in the years between this year and death with current utilities \* NPI following year utility value vector
12. Add QALYs (sum of QALY vector) to the total QALY counter
13. End of J loop
14. Combine individual life year counter, QALY counter, costs, screen count, cancer detection, and cancer diagnostic vector in a single results vector (the I loop is a “for each” and binds these results vectors together)
15. End of I loop
16. Create a results data frame using these results
17. Rename columns of results data frame
18. Save results in R file (will result in 10 R files)
19. End of 10 grouping loop
20. Create a matrix for merged results with 5 columns and 10 rows
21. Set up new loop:
    1. Load R file
    2. Copy QALYs into first column
    3. Copy cost in second column
    4. Copy screens into third column
    5. Copy cancers into 4th column
    6. Copy screen detected into fifth column
    7. Close loop
22. Save results as csv