**Description of initial data for CMOST simulation**

Write all data on the right side of “=” without a space. While using numbers write the ending .0 and write always “.” rather than coma and separate numbers in one vector with “,” without a space. At the end of vector do not write anything.

Please be aware that the program is not thoroughly tested for setting much different then those in default .ini files.

1. **Settings.ini**

**[general]**

* PopulationSize

Positive integer which defines how many people are in the population we want to simulate.

* YearsToSimulate

The maximum age of individual (should be not more than 100, if much less then program may not work).

* Dt

Timestep for the simulation – fraction of a year.

* RandomNumberSeed

Positive integer used to generate random numbers.

* nCPUs

Positive integer, number of processor cores to be used for simulations.

* maxCPUsPerEvaluation

Setting for multi-socket machines. Typically same as nCPUs.

**[evaluation]**

* CalculateIncidence

0-no, 1-yes

If to report incidence in the output files.

* CalculatePolypsPrevalence

0-no, 1-yes.

If to report polyps prevalence in the output files.

* EvalIniFile

Give a name and path of file containing information for costs evaluation. The program takes into account variables given in this file such as costs of colonoscopy, costs of colonoscopy’s complications etc. The name must be on the right side of symbol "=", first folder followed by “/” and the name of file. This file must be “.ini” type.

* resultsFile

Give a name and path of file where the baseline results data have to be located. These are the results where there is no screening in population. The name must be on the right side of symbol "=", first folder followed by “/” from the name of file. This file must be “.out” type.

**[stratification]**

* StratifyPopulation

0-no, 1-yes.

When this option in turned on the population is divided into 2 pieces: high and low risk of getting cancer. When there is no stratification than each output file (except baseline results) will have prefix “strata0”. On the contrary, with stratification turned on two different prefix will be added to output files: “strata0” and “strata1” according to high or low risk group in order.

* StratificationIniFile

Strafitication Initial File. Give a name and path of file with settings for stratification. The name must be on the right side of symbol "=", first folder followed by “/” and the name of file. This file must be “.ini” type.

**[optimization]**

* performOptimization

0-no 1-yes;

This option allows to run optimization to find screening moments that minimize mortality, incidence, life years lost or total costs (can have more than one at the same time).

* optimizationINIfile

Optimization Initial File. Give a name and path of file with initial data of optimization parameters. The program takes into account variables given in this file such as which function to minimize, search space, optimizer settings etc. There is also an option to generate additional output files: Screen and Solution. The name must be on the right side of symbol "=", first folder followed by “/” and the name of file. This file must be “.ini” type.

* optimizationOutFile

Give a name and path of file where the optimization results have to be located. The name must be on the right side of symbol "=", first folder followed by “/” from the name of file. This file must be “.out” type.

**[screening]**

* evaluateScreeningScenarios

0-no, 1-yes.

This option allows to evaluate screening scenarios.

* screeningINIfile

Screening Initial File. Give a name and path of file with initial data of screening parameters. The program takes into account variables given in this file which are e.g. 20 different screening scenarios with given colonoscopy moments, assumed patient compliance to guidelines and minimal time between two screening colonoscopies. The name must be on the right side of symbol "=", first folder followed by “/” and the name of file. This file must be “.ini” type.

* screeningOutFile

Give a name and path of file where the screening strategy results have to be located. The name must be on the right side of symbol "=", first folder followed by “/” from the name of file. This file must be “.out” type.

**[flags]**

* correlation

0-no 1-yes.

If this option is turned on then advanced and early polyp progression will be correlated (if fast progressing early polyp then fast progressing advanced polyp).

* PolypSurveillance

0-no, 1-yes.

If to assume standard polyp surveillance after removal.

* CancerSurveillance

0-no, 1-yes

If we assume standard cancer surveillance.

* AllPolypFollowUp

0-no 1-yes

If to assume follow-up after any number and kind of polyps.

**[population\_parameters]**

* fractionFemale

Number from range 0-1 which indicates which fraction of females in population.

* LifeTableMales

Vector of probabilities of dying within next year at given age for males (position within vector corresponds to given age). Length >= YearsToSimulate

* LifeTableFemales

Vector of probabilities of dying within next year at given age for males (position within vector corresponds to given age).

**[polyp\_related\_parameters]**

* initialPolypStage

Initial stage of new polyps.

* NumPolypStages

Nonnegative integer which represents the maximum possible stage of polyp. When the polyp’s stage number reaches value more than “NumPolypStages” then it becomes a cancer.

* advancedPolypStageTransition

Nonnegative integer. When the polyp stage value is equal to “advancedPolypStageTransition” or bigger it is taken as advanced polyp, when smaller then it is an early polyp.

* GeneralNewPolypsRisk

Vector describing risk of polyp creation according to age (each element is risk at given age). If this vector has less numbers than “YearToSimulate” value program will not work properly.

* PolypStage1AgeProgressionRate
* PolypStage2AgeProgressionRate
* PolypStage3AgeProgressionRate
* PolypStage4AgeProgressionRate
* PolypStage5AgeProgressionRate
* PolypStage6AgeProgressionRate

These are vectors of real numbers from range 0-1. These 6 vector above should be the same length which is equal to “YearsToSimulate” value. Indicates the rate at which the polyp will progress to next stage according to age (location in the vector reflects the age starting from 0).

* IndividualPolypRisk

Vector describing distribution of individual risks of polyp development at each time step. For each individual in the population risk is randomly selected from this vector.

* femaleNewPolypRisk

Scaling factor of individual risk of polyp creation for females.

* HealingRates

Vector of real numbers from range 0-1. Should be at least "NumPolypStages" long. Each place in this vector reflects current polyp stage. These numbers represents the probability that at given stage the polyp will get smaller by itself.

* FastCancerRates

Vector of real numbers describing the rates at which polyp of a given stage can directly progress to cancer (multiplied by general rate at which polyp progresses to the next stage).

* DwellSpeed

0-no 1-yes.

When this option in turned on then rate at which polyp directly progresses to cancer is additionally multiplied by general polyp progression rates.

**[cancer\_related\_parameters]**

* GeneralDirectCancerRiskMale
* GeneralDirectCancerRiskFemale

These are vectors of real numbers which indicate general risk of direct cancer according to the age and gender. This vector should be at least “YearToSimulate” long.

* DirectCancerSpeed

Real number multiplying GeneralDirectCancerRiskFemale and GeneralDirectCancerRiskMale.

* StageDurationStageIIDiagnosis
* StageDurationStageIIIDiagnosis
* StageDurationStageIVDiagnosis

Vectors of real numbers: stageII: 1 number, stageIII: 2 numbers, stageIV:3 numbers. They describe how long the cancer is particular stage before symptoms development at a given stage.

* survivalCutoff

Length of the survival curves to be considered (e.g. 5 years).

* osByGenderAgeStage

Survival data by gender, age group and stage (up to survival cutoff value).

* nAgeGroups

Nonnegative integer describing the number of age groups in survival data.

* osAgeRanges

Vector of nonnegative integers which indicates the maximum age in given age group. Has to include 1 number less than "nAgeGroups” value. e.g. here age groups are: 1-51, 52-60, 61-70, 71-80, 80-100.

* numDataPointsPerSurvCurve

Number Data Points per Surveillance Curve. Nonnegative integer smaller than 255. The value depends on the intervals at which survival is given (months, days, years).

* fractionBySexAndAgeAtDiagnosis\_Stage1
* fractionBySexAndAgeAtDiagnosis\_Stage2
* fractionBySexAndAgeAtDiagnosis\_Stage3
* fractionBySexAndAgeAtDiagnosis\_Stage4

These are vectors of real numbers from rage 0-1. Indicate the fraction of cancers at a given stage by sex and age group.

* sojournTimeStage1AtDiagnosisCDF
* sojournTimeStage2AtDiagnosisCDF
* sojournTimeStage3AtDiagnosisCDF
* sojournTimeStage4AtDiagnosisCDF

Cumulative distribution function describing the distribution of time from cancer initiation to symptoms development.

**[progression\_risk\_parameters]**

* femaleEarlyProgression

Real number. If more than 1 then it makes early polyps progress faster in case of woman.

* femaleAdvancedProgression

Real number. If more than 1 then it makes advanced polyps progress faster in case of woman.

* EarlyPolypProgression
* AdvancedPolypProgression

Vectors of real numbers. These two should be the same length. New polyp has a value of early/advanced progression risks taken randomly from those vectors. If “correlation” is turned on then the number will be taken in the same place from each vectors.

**[location\_related\_parameters]**

* numLocation

Nonnegative integer smaller than 255. It indicates numbers of possible polyps’ locations.

* NewPolypLocation

Vector describing the cumulative distribution function for location of newly developed polyp.

* NewCancerLocation

Analogous to “NewPolypLocation” for direct cancers.

* LocationEarlyPolypProgression
* LocationAdvancedPolypProgression

Vectors of real numbers describing risk of progression depending on location. Should be “numLocations” long.

**[colonoscopy\_related\_parameters]**

* ColoReach

Vector describing the probability distribution defining to which location performed colonoscopy can reach.

* ColoDetectionPolyp

Vector describing detection probability of polyp at a given stage (need to by “NumPolypStages” long).

* ColoDetectionCancer

Vector describing detection probability of cancer at a given stage.

* ColoDetectionLocation

Vector describing the probability of polyp/cancer detection at a given location.

**[risks]**

* ColonoscopyRiscPerforation

Probability of perforation during colonoscopy.

* DeathPerforation

Probability of dying because of perforation

* RectosigmoPerforation

Probability of perforation during rectosigmoidoscapy (not currently used by the program).

* ColonoscopyRiscSerosaBurn

Probability of serosa burn during colonoscopy.

* ColonoscopyRiscBleeding

Probability of bleeding during colonoscopy.

* ColonoscopyRiscBleedingTransfusion

Probability of bleeding during colonoscopy that requires transfusion.

* DeathBleedingTransfusion

Probability of dying during colonoscopy after bleeding that requires transfusion.