Mount Sinai
Version: 1.0.00
Released: 2025-09-30

# Mount Sinai Uterine Cancer Model (MUSIC): Model Profile

## Icahn School of Medicine at Mount Sinai



### Contact

Chung Yin (Joey) Kong (chungyin.kong@mountsinai.org)

### Funding

The development of this model was supported by the NIH/NCI CISNET Uterine Cancer Grant (U01CA265739).

### Suggested Citation

Kong CY, Frotscher A, Azzis A, Bickell N, Blank S, Gwalani P, Lane T. Mount Sinai Uterine Cancer Model (MUSIC): Model Profile. [Internet] Sep 30, 2025. Cancer Intervention and Surveillance Modeling Network (CISNET). Available from: https://cisnet.cancer.gov/resources/files/mpd/uterine/CISNET-uterine-music-model-profile-1.0.00-2025-09-30.pdf

### Version Table

| Version | Date | Notes |
|---|---|---|
| 1.0.00 | 2025-09-30 | Initial release |

*Mount Sinai*
Readers Guide

Icahn School of Medicine at **Mount Sinai**



# Reader's Guide

**Core Profile Documentation**

These sections will provide an overview of the model without the burden of excessive detail. Each can be read in about 5-10 minutes. Each contains links to more detailed information if required.

**Model Purpose**
This document describes the primary purpose of the model.

**Model Overview**
This document describes the primary aims and general purposes of this modeling effort.

**Assumption Overview**
An overview of the basic assumptions inherent in this model.

**Parameter Overview**
Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

**Component Overview**
A description of the basic computational building blocks (components) of the model.

**Output Overview**
Definitions and methodologies for the basic model outputs.

**Results Overview**
A guide to the results obtained from the model.

**KeyReferences**
A list of references used in the development of the model.

*Mount Sinai*
Model Purpose

Icahn School of Medicine at **Mount Sinai**



# Model Purpose

## Summary

This page describes the purpose of the Mount Sinai Uterine Cancer Model.

## Purpose

The MUSIC model is designed to project uterine cancer incidence and mortality over time, considering key factors such as stage at diagnosis, histology, race (non-Hispanic White and non-Hispanic Black), age, and various screening and treatment strategies. By incorporating birth cohort effects, the model captures how generational differences in risk factors and healthcare access influence disease patterns over time. Additionally, the model explicitly accounts for cancer recurrence: this provides insight into which combinations of histology and racial background are more likely to experience recurrence, further informing targeted interventions.

The model does not provide individualized risk predictions, but rather is designed to guide population-level decision-making by identifying trends and intervention opportunities that could have the greatest impact. It is continuously evolving and can be updated with data from more recent calendar years, ensuring that projections remain relevant and reflect changing healthcare trends.

# Model Overview

## Summary

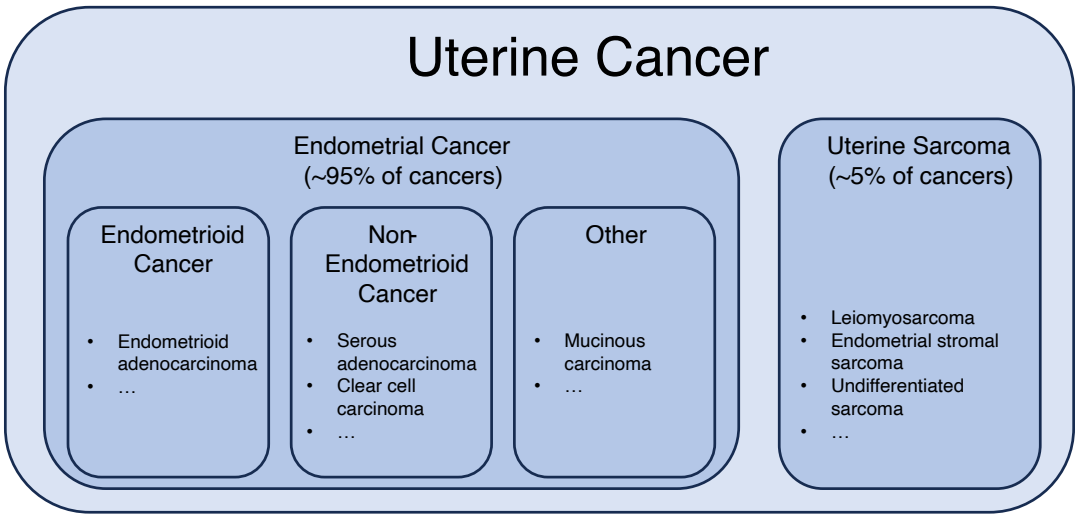
## Purpose

The purpose of MUSIC is to simulate uterine cancer incidence and mortality while explicitly modeling age, race, calendar year, obesity, and birth cohort effects.

## Background

As of 2020, uterine cancer was the sixth most-diagnosed cancer in women worldwide, with the incidence being highest in North America [1]. In the 30-year period between 1990 and 2019, high-income North America experienced a 45.3% increase in the age-standardized incidence rate, compared to a 15.3% increase globally, with an estimated annual percent-change of 1.3-1.44% [2,3]. This trend has been attributed in part to the increasing prevalence of obesity; a meta-analysis indicated that the risk ratio between a $5\,\mathrm{kg/m^2}$ increase in body mass index (BMI) and uterine cancer is 1.59, with this ratio increasing for BMIs above $28\,\mathrm{kg/m^2}$ [4,5]. Hormone replacement therapy consisting of unopposed estrogen has also been shown to increase the risk of uterine cancer, but this was mitigated to an extent by an increased use of progesterone in combination with estrogen starting in the 1980s [6,7].

There are significant racial disparities in the United States regarding the incidence, prognosis, and mortality of uterine cancer. A systematic review of studies performed between 1997 and 2023 indicated that while the incidence of uterine cancer over that time period has risen in African-American women to be comparable to that of white American women, mortality as of 2019 is double for African-American women and 5-year survival is worse [8]. Additionally, African-American women are more likely to be diagnosed at later stages and with non-endometrioid subtypes, which have worse prognoses than endometrioid subtypes [8].




## Model Description

The MUSIC model is an empirically calibrated, stochastic, continuous-time Markov chain (CTMC) model of uterine cancer with four categories: endometrioid `(EM)`, non-endometrioid `(Non-EM)`, carcinomasarcoma, leiomyosarcoma, and dedifferentiated sarcoma `(sarcoma)`, and all other subtypes `(other)`. We are using the following ICD-O-3 codes for each category:

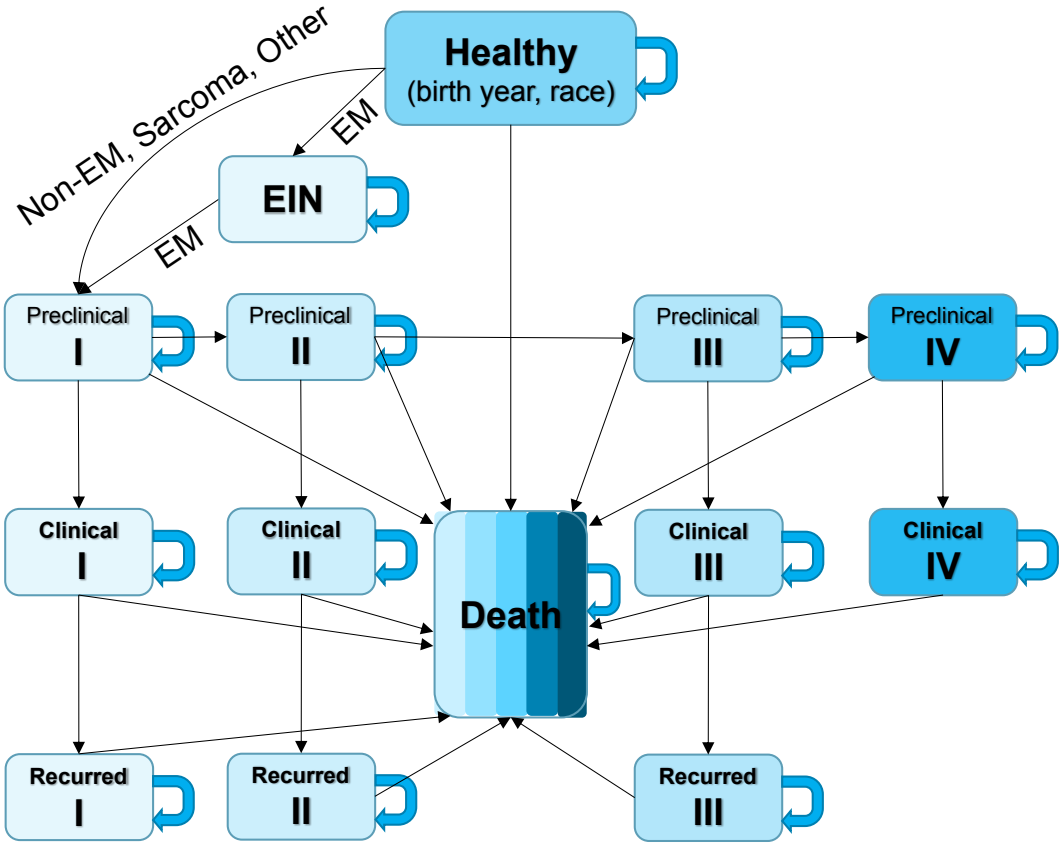
| Category | ICD-O-3 code |
|---|---|
| Endometroid | 8050, 8141, 8210, 8260-8263, 8380-8383, 8440, 8480-8481, 8560, 8570 |
| Non-Endometrioid | 8255, 8310, 8323, 8441, 8460-8461, 8950-8951, 8980-8981 |

| Category | ICD-O-3 code |
|---|---|
| Sarcoma | 8800-8802, 8804-8805, 8840, 8850, 8890-8891, 8895-8896, 8900-8902, 8910, 8912, 8920, 8930-8931, 8933, 8935, 9120 |
| Other | all other codes |

This model is designed to follow a synthetic birth cohort with specific ages, ranging from 35 to 85 years. The cycle length is variable but set to `1 month`. At the start of the simulation, all subjects start in the healthy state, as in Fig. 1. Subjects can transition multiple times per cycle, but only to 'more advanced' states: Healthy `h` → Preclinical Cancer `Pre` → Clinical Cancer `Clinical` → Recurred Cancer `Recur` → Death `Death`. The computational core is a transition rate matrix $A$, which depends on age, race, BMI, and time since cancer diagnosis (for subjects with cancer only). The transition probability vector $\vec{p}$ is calculated as follows (with simplified transition rate matrix $A$ and the initial state $\vec{s}_0$):

$$\vec{p}(t) = \mathrm{e}^{A \cdot t} \cdot \vec{s}_0 \qquad \vec{s}_0 = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \qquad A = \begin{bmatrix} -\sum t_{i0} & 0 & 0 & 0 & 0 & 0 \\ t_{h \to \mathrm{Pre}} & -\sum t_{i1} & 0 & 0 & 0 & 0 \\ 0 & t_{\mathrm{Pre} \to \mathrm{Clinical}} & -\sum t_{i2} & 0 & 0 & 0 \\ 0 & 0 & t_{\mathrm{Clinical} \to \mathrm{Recur}} & -\sum t_{i3} & \dots & 0 \\ 0 & 0 & t_{\mathrm{Clinical} \to \mathrm{C.Death}} & t_{\mathrm{Recur} \to \mathrm{C.Death}} & \ddots & \vdots \\ t_{h \to \mathrm{Death}} & t_{h \to \mathrm{Death}} & t_{\mathrm{Clinical} \to \mathrm{Death}} & t_{\mathrm{Recur} \to \mathrm{Death}} & \dots & 0 \end{bmatrix}$$

The subcomponents of natural history, incidence, and survival/mortality are described in the component overview section; however, they all work together to function.




# References

1. Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. Wiley Online Library; 2021;71(3):209–249.

2. Baoxia Gu, Xiaogai Shang, Mengqing Yan, Xiao Li, Wei Wang, Qi Wang, et al. Variations in incidence and mortality rates of endometrial cancer at the global, regional, and national levels, 1990–2019. Gynecologic Oncology [Internet]. 2021;161(2):573–580. Available from: https://www.sciencedirect.com/science/article/pii/S0090825821000962
3. Liu Yang, Yue Yuan, Rongyan Zhu, Xuehong Zhang. Time trend of global uterine cancer burden: an age-period-cohort analysis from 1990 to 2019 and predictions in a 25-year period. BMC women's health. Springer; 2023;23(1):384.
4. Gillian K Reeves, Kirstin Pirie, Valerie Beral, Jane Green, Elizabeth Spencer, Diana Bull. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj. British Medical Journal Publishing Group; 2007;335(7630):1134.
5. Andrew G Renehan, Margaret Tyson, Matthias Egger, Richard F Heller, Marcel Zwahlen. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. The lancet. Elsevier; 2008;371(9612):569–578.
6. Deborah Grady, Tebeb Gebretsadik, Karla Kerlikowske, Virginia Ernster, Diana Petitti. Hormone replacement therapy and endometrial cancer risk: a meta-analysis. Obstetrics & Gynecology. Elsevier; 1995;85(2):304–313.
7. Harry K Ziel, William D Finkle, Sander Greenland. Decline in incidence of endometrial cancer following increase in prescriptions for opposed conjugated estrogens in a prepaid health plan. Gynecologic oncology. Elsevier; 1998;68(3):253–255.
8. Michael L Hicks, Maya M Hicks, Roland P Mathews, Dineo Khabele, Camille A Clare, Onyinye Balogun, et al. Racial disparities in endometrial cancer: Where are we after 26 years? Gynecologic oncology. Elsevier; 2024;184:236–242.



Icahn School of Medicine at **Mount Sinai**



# Assumption Overview

## Summary

An overview of the basic assumptions inherent to the Mount Sinai MUSIC model.

## Background

The Mount Sinai MUSIC model is a microsimulation model, which is based on a continuous-time Markov chain. As such, for each cycle all transition rates need to be known/estimated for the model to function, including cycles which relate to future calendar years.

## Assumption Listing

Model Input Data

- We can combine BMI data from the Uterine BMI history generator together with cancer mortality and cancer incidence data from SEER.
- We can also add all-cause mortality data from CDC Wonder / Columbia and combine it.
- Obesity is an important factor for uterine cancer incidence. More specifically, it is assumed to affect only the endometrioid subtype. There are 5 groups (see Component Overview), and there is an average risk ratio of 1.72 between all of those groups.

Natural History of Uterine Cancer

- A endometrial intraepithelial neoplasia state exists. It is the sole precursor state to the endometrioid cancer state.
- All cancers grow without skipping stages, before being detected, i.e. Preclinical I $\rightarrow$ II $\rightarrow$ III $\rightarrow$ IV. They start from an undetected state.
- All subjects in clinical cancer stages I, II and III can recur, after having spent 1 year in their state. They cannot recur after more than 10 years.
- Subjects who did recur cannot recur again, and can only die from other causes or due to uterine cancer.
- Transition rates between states remain constant within each cycle but can change in between cycles.

# Parameter Overview

## Summary

This section lists the most important parameters for the MUSIC model.

## Background

Parameters affect the transition rates from one state to another in the model, and are thus crucial to its predictions. They can be divided into "Model Inputs" (parameters that are fixed and derived from the properties of uterine cancer), "Calibration Targets" (parameters the model will be tuned to reproduce), and "Internal Parameters" (parameters the model uses for tuning).

## Parameter Listing Overview

General Parameters:

- startAge: age of subjects at the start of the simulation
- stopAge: age of subjects (that didn't die) at the end of the simulation
- cycleFrequency: number of cycles/yr that are calculated
- numberOfPatients: number of subjects that are simulated
- startYear: calendar year in which simulation starts
- raceDistribution: (share of white subjects, share of black subjects)

| | Implementation | Data Source | Example |
|---|---|---|---|
| **Model inputs** | | | |
| BMI groups | ranges | Pfeiffer *et al.*[1] | $20-25\,\mathrm{kg/m^2}, >40\,\mathrm{kg/m^2}$ |
| BMI group risk | risk ratio between groups | Pfeiffer *et al.*[1] | $1.72/\mathrm{group}$ |
| BMI risk follow up period | conversion risk -> hazard | Pfeiffer *et al.*[1] | 10 years |
| Cancer dwell times | fixing progression between preclinical cancer states | Sasiensi *et al.*[2] | Stage I median time: 4 years |
| Relative uterine death | relative mortality rates by age | Global Burden of Disease[3] | 0.49% for women age 40-45 |
| All-cause mortality | monthly, mortality rates by race, age and birth cohort | CDC Wonder [4] | Mortality rate for 1970 born black women, age 34.5-44.5 is 530/yr/1E5 |
| Cancer-specific survival | cause-specific surival by months since diagnosis, race and histology for Stage IV | SEER (2000-2018) | cause-specific survival for white women (age 40-44 at diagnosis) with endometrioid cancer stage IV after 100 months is 37.1% |
| **Calibration targets** | | | |
| Age-adjusted cancer incidene | rates by histology, race, stage, birth cohort | SEER (2000-2018) | Incidence 2000 for NH-black women, Non-endometrioid cancer 9.8/yr/1E5 |
| **Internal parameters** | | | |
| Cancer detection rates | Stage-, histology-, birth-cohort- and race-specific transition rate between preclinical cancer and clinical cancer | - | transition rate for 1980-1990 birth cohort, white, sarcoma, Stage I Preclinical to I Clinical is 0.099/yr |

Additional input parameters are cancer incidences by year, age, race, histology, and stage that have been processed with multiple imputation by Columbia University ("YARHA").

# References

1. Ruth M Pfeiffer, Yikyung Park, Aimée R Kreimer, James V Lacey Jr, David Pee, Robert T Greenlee, et al. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS medicine. Public Library of Science San Francisco, USA; 2013;10(7):e1001492.
2. Peter Sasieni, Rebecca Smittenaar, Earl Hubbell, John Broggio, Richard D Neal, Charles Swanton. Modelled mortality benefits of multi-cancer early detection screening in England. British Journal of Cancer. Nature Publishing Group UK London; 2023;129(1):72–80.
3. Global Burden of Disease Collaborative Network. Global Burden of Disease Study. Seattle, United States: Institute for Health Metrics; 2021;
4. Centers for Disease Control, National Center for Health Statistics Prevention. National Vital Statistics System, Mortality 1999-2020 on CDC WONDER Online Database. 2021;

# Component Overview

## Summary

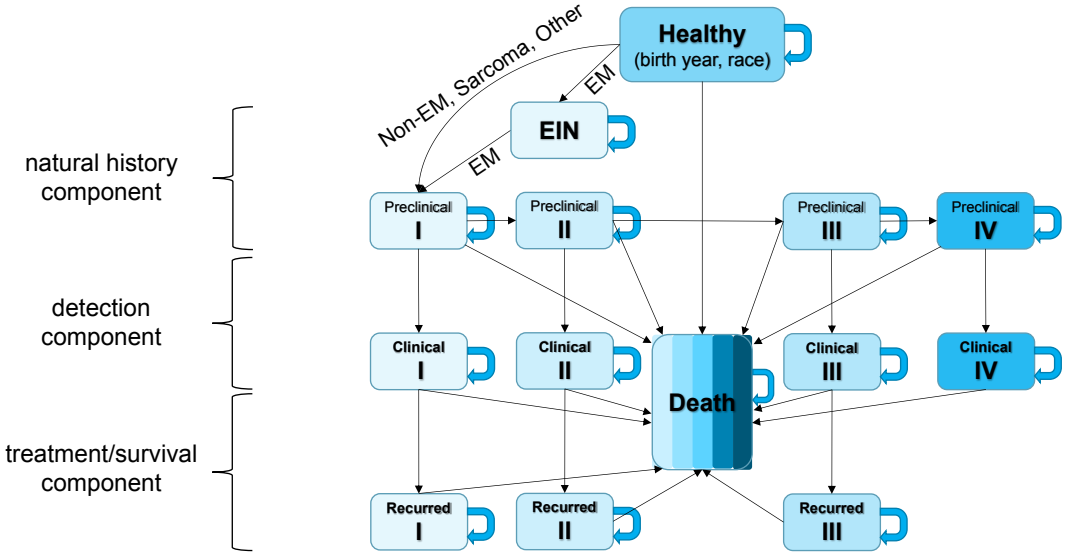
This page contains a description of the different computational elements which make up the Mount Sinai Uterine Cancer Model.

## Overview

The subjects are first generated in the population component. Their birth year and race is decided, and a BMI trajectory that they will follow is picked. The natural history component takes care of transitioning subjects from healthy to preclinical cancer states. Note that higher preclinical cancer states are always preceeded by their lower states.
The detection component is responsible for transitioning subject to clinical cancer states. It takes into account the lag between the development and detection of cancer.
The treatment and survival component handles cancer-specific mortality and recurrence. It keeps track of the time spent in those states to generate accurate transition rates.

## Component Listing

- Natural History Component
- Calibration Component
- Detection Component
- Treatment/Recurrence Component
- Survival Mortality Component




### Natural History Component

Transition rates of subjects from the healthy state to a preclinical cancer Stage I (for non-EM, sarcoma and other uterine cancer) or to the endometrial intraepithelial neoplasia (EIN) state (for EM uterine cancer) must be calculated. The benchmark, incidence data, is taken from SEER and depends on calendar year (2000-2021), subtype (EM, non-EM, sarcoma, other), race (NH white, NH black), age group (5yr blocks) and AJCC stage. Each preclinical state (and the EIN state) is associated with a stage-specific dwell time $T_{\text{dwell}}$, taken from the supplemental material of Sasieni *et al.* [1]. For the current age, linear interpolation is applied to get specific incidences for each cycle. The dwell time used is tabulated as follows:

| Uterine Cancer Subtype | Stage I | Stage II | Stage III | Stage IV |
|---|---|---|---|---|
| Endometrioid | 4 yr | 2 yr | 1 yr | 1 yr |

| Uterine Cancer Subtype | Stage I | Stage II | Stage III | Stage IV |
|---|---|---|---|---|
| Non-Endometrioid | 4 yr | 2 yr | 1 yr | 1 yr |
| Sarcoma | 4 yr | 2 yr | 1 yr | 1 yr |
| Other | 4 yr | 2 yr | 1 yr | 1 yr |

From it, transition rates can be calculated:

$$\frac{1}{T_{\text{dwell}}} = \sum_{r_{\text{out}}} r = r_{\text{detection}} + r_{\text{Nat. death}} + r_{\text{advance}}$$

with the transition rate to get detected denoted as $r_{\text{detection}}$, the transition rate to die from other causes denoted as $r_{\text{Nat. death}}$ and the transition rate to a higher preclinical stage (except of subject is already at Stage IV) denoted as $r_{\text{advance}}$.

## Calibration Component

The calibration component is responsible for adjusting the detection rates, i.e. preclinical cancer $\rightarrow$ clinical cancer. Those rates depend on race, birth cohort, age, stage (dwell times) and histology. Clinical cancer is not reversible, so all rates are positive. The calibration process is performed separately for each birth cohort. The actual process closely resembles a fixed-point iteration with relaxation (see Epperson[2] Chapter 3.9 and Theorem 7.18). Transition rates are deemed optimal if the transition rate scales $s$ (see Detection Component) averaged over all cycles are within 1% of each other. If necessary, manual fine-tuning will complete the calibration.

## Detection Component (preclinical to clinical states)

The incidences given by SEER must be matched with the transition rates from the natural history component. The delay, which originates from the dwell times of the preclinical cancer states, has to be taken into account. Transition rates and dwell times do not depend on the cycle number, such that the response function $f$ of 'subjects transition to preclinical cancer' to 'subjects transition from preclinical cancer to clinical cancer' does not change over time. It is calculated with an arbitrary rate ($r = 10^{-3}$), which needs to be scaled later, using a limited transition rate matrix in which all clinical cancer states are also absorbing states. It includes the possibility of death before diagnosis as well. The response function is cut off at 30 years past initial transition to preclinical cancer.

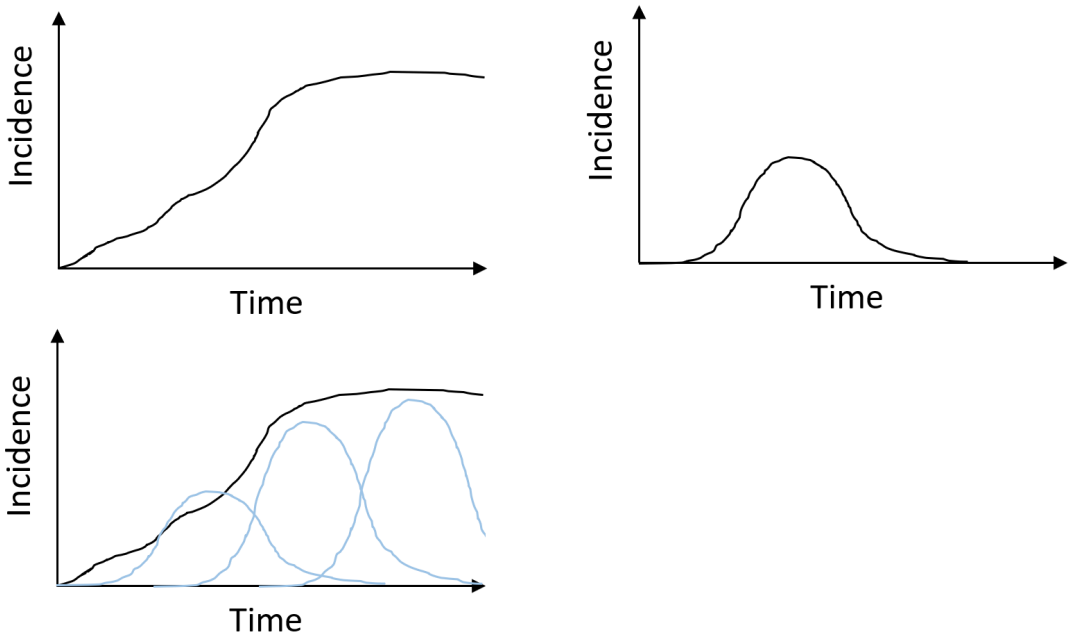
$$f_I(t) = P(\text{Preclinical I} \rightarrow \text{Clinical I} \mid [\text{Healthy} \rightarrow \text{Preclinical I}]_{t=\text{Age}}) \quad \text{Age} + 0\text{yr} \leq t \leq \text{Age} + 30\text{yr}$$

Thus, a simple deconvolution of each incidence function (which depends on cancer subtype, stage, calendar year, and race) over time is performed. The result is an optimal transition rate scale $s$ from healthy state to preclinical cancer stage I for each cycle,

$$s = \frac{1}{I_{\text{tot}}^2} \cdot \sum_{\text{cycle}j=0}^{n} I_j \cdot I_{\text{SEER}}(j)$$

with $I_{\text{tot}} = \int_{t=0}^{30} f(t)\mathrm{d}t$ being the total incidence of the response function, $j$ the current relative cycle of the response function, and $I_{\text{SEER}}$ the expected incidence given by SEER.

However, after a subject transitions to a preclinical cancer state, it is undetermined at which (if any) stage it gets diagnosed.
To account for all options, the actual applied transition rate from healthy to a specific preclinical cancer state is calculated using a weighted sum of the obtained parameters for each stage. The weights are calculated from the relative chance of being detected in any particular stage during the duration of the simulation. Fig. 2 illustrates a simplified matching process: the incidence function from SEER, the response function of the Markov model, as well as the matched incidence function with multiple response functions.




Cancer incidence also depends on BMI. In this model, 5 groups are established.

| BMI Group # | Lower Limit (kg/m²) | Upper Limit (kg/m²) |
|---|---|---|
| 0 | - | 25 |
| 1 | 25 | 30 |
| 2 | 30 | 35 |
| 3 | 35 | 40 |
| 4 | 40 | - |

BMI trajectories are sampled from the uterine obesity history calculator. Subjects will select one BMI trajectory at random at the start of the simulation and then follow it throughout the simulation. The incidences will then be rebalanced (at each cycle, per race), so that:

1. the total expected number of cases is unchanged and
2. the hazard ratio (HR) to transition to preclinical cancer I between two adjacent groups (i.e., HR (II vs. I), …) is constant, and >1.

We do have risk ratios for endometrial cancer only, which is why this rebalancing is done for the sum of all cancer subtypes. Endometrioid cancer is thought to be the only subtype depending on BMI, so all changes are projected to this subtype only. However, we do need adapted transition rates, so to convert them, we first calculate the average non-cancer chance $x$:

$$x = \exp\left(-\lambda_{\text{total}} \cdot t_{\text{study}}\right)$$

with $\lambda$ being the target transition rate and $t_{\text{study}}$ the average follow-up time for which the risk ratio was calculated. Next, we calculate the non-cancer chance $x_0$ for subjects in BMI group 0 (<25 kg/m²) after time $t_{\text{study}}$:

$$x_0 = 1 - \frac{1 - x}{\alpha_0 RR^0 + \alpha_1 RR^1 + \alpha_2 RR^2 + \alpha_3 RR^3 + \alpha_4 RR^4}$$

with $\alpha_i$ being the distribution of subjects within the BMI groups at that time and its sum being 1. The modified transition rates $\lambda_i^{\text{EM}}$ for EM cancer can now be calculated:

$$\lambda_i^{\text{EM}} = -\frac{\log(1 - (1 - x_0) \cdot RR^i)}{t_{\text{study}}} - \lambda_i^{\text{nonEM}} - \lambda_i^{\text{sarcoma}} - \lambda_i^{\text{other}}$$

In case that $\lambda_i^{\text{EM}} < 0$ the rate will be fixed to 0, and the remainder will be discounted from the other rates $\lambda_i^{\text{nonEM}}, \lambda_i^{\text{sarcoma}}$ and $\lambda_i^{\text{other}}$, i.e:

$$\lambda_i^{\text{nonEM}'} = (1 - \frac{\lambda_i^{\text{EM}}}{\lambda_i^{\text{nonEM}} + \lambda_i^{\text{sarcoma}} + \lambda_i^{\text{other}}}) \cdot \lambda_i^{\text{nonEM}}$$

Lastly, birth cohort specific-effects are implemented using scaling factors, which are between 0.8 and 1.1.

## Recurring component

Starting 1 year after transitioning into a clinical cancer state, the primary cancer of subjects can recur. The transition rates are being calculated from a Fine & Gray analysis, which yields a cumulative incidence function (CIF). The reference group is White, 66-69 years old, had Stage I endometrioid cancer initially. Subdistribution hazard ratios were generated to account for all other combinations of age at recurrence, race, stage, and histology. Stage IV cancer is typically not treated with intent to cure, as such subjects cannot recur after having been diagnosed with Stage IV cancer. Recurred subjects have higher mortality rates compared to subjects which do not recur. The data was obtained form Pranav Gwalani who used SEER-Medicare linked databases.

The transition rate from state *i* to *k* $\lambda_{ik}$ is calculated from the CIF [3]:

$$\text{CIF}_{ik}(t) = \int_0^t \lambda_{ik}(s) S(s) \mathrm{d}s$$

Here, $S(t)$ represent the survival function, that is the probability of being in the initial state $i$ at time $t$. It can be calculated from all $L$ outgoing transition rates $\lambda_{il}(t)$:

$$S(t) = \exp\left(-\sum_{l=1}^{L} \Lambda_{il}(t)\right) \qquad \Lambda_{il}(t) = \int_0^t \lambda_{il}(s) \mathrm{d}s$$

Transition rates originating from the recurred state (state 1) are assumed to be constant within each cycle. Only death from other causes (state 3) and death due to uterine cancer (state 2) are simulated during recurrence. If the cycle length is set to 1 (causing CIF and $\lambda_{13}$ to have units $[\text{cyclelength}^{-1}]$), the CIF for each cycle $n$ can then be obtained: $\Delta\text{CIF} = \text{CIF}(n+1) - \text{CIF}(n) \equiv \text{CIF}$. Using this simplification, we can carry out the integration to be:

$$\text{CIF}_{12} = \frac{\lambda_{12}}{\lambda_{12} + \lambda_{13}} \times (1 - \mathrm{e}^{-(\lambda_{12} + \lambda_{13})}) \approx \frac{\lambda_{12}}{\lambda_{12} + \lambda_{13}} \times \left(\lambda_{12} + \lambda_{13} - \frac{(-\lambda_{12} - \lambda_{13})^2}{2}\right)$$

The Taylor-expansion of the exponential function to the third term gets its validity from the low maximum change per cycle for the reference group of 5.9%. As such, the transition rate is:

$$\lambda_{12}^{\pm} = \frac{1}{2}\left(\pm\sqrt{(\lambda_{13} - 2)^2 - 8\text{CIF}_{12}} - \lambda_{13} + 2\right)$$

with $\lambda_{12}^{-}$ being the correct solution. To extract the transition rate for non-reference groups (i.e. Black women), the SHRs are used as follows [4]:

$$\lambda_{\text{black}}(t) = \text{SHR}_{\text{black}} \cdot \lambda_{\text{white}}(t)$$

## Mortality Component

Two types of transitions lead to death in the MUSIC model:

1. Transitions between all non-death states and the natural/other-causes death state.
2. Transitions between all clinical cancer states and cancer-specific death states as well as transitions between all survivor states and cancer-specific death states.

Natural death (= transition to natural/other-causes death state) can happen from any non-death state. It does include all causes of death, except death from uterine cancer. Subjects in certain states (clinical cancer, survivor states) are subject to the competing risk of dying from uterine cancer and dying from other causes. Subjects dying because of uterine cancer in preclinical states will still transition to the other-cause death state, as that cancer was undetected until death.

Background mortality data is taken from the uterine obesity history generator, made by William D Hazelton from Fred Hutchinson Cancer Research Center. It covers mortality from 1968 to 2016. Other covariates are age (10-year bins) and race (Hispanic, NH white, NH black, AAPI, AI/AN). Linear interpolation is being applied to the age variable to account for the precise age each cycle.

```
import pandas as pd
from scipy.interpolate import interp1d

# Create new DataFrame from input excel file
df = pd.read_excel('Mortality-Calculations.xlsx')

# Create a piece-wise linear interpolation function for each column
interp_func = lambda group: pd.DataFrame({
    'interpolated_rate': interp1d(group['Age'], group['Rate'],
kind='linear', fill_value='extrapolate')(range(0, 86)),
    'interpolated_lower': interp1d(group['Age'], group['Lower'],
kind='linear', fill_value='extrapolate')(range(0, 86)),
    'interpolated_upper': interp1d(group['Age'], group['Upper'],
kind='linear', fill_value='extrapolate')(range(0, 86))
})

# Perform the interpolation for each group and store the results in a
new DataFrame
interpolated_data = df.groupby(['Year', 'Race']).apply(interp_func)

# Write the interpolated data into a new Excel/.csv file
interpolated_data.to_excel('Mortality-Split.xlsx')
interpolated_data.to_csv('Mortality-Split.csv', index=True)
```

The relative number of deaths due to uterine cancer has been estimated using the Global Burden of Disease tool. Subjects in clinical cancer states and survivor states have the additional risk of dying due to uterine cancer.

For non-recurred subjects, the underlying database interrogated is SEER Research Data, Nov 2023 Sub (2000-2021), looking at cause-specific survival after being diagnosed with cancer in Corpus Uteri and Uterus, NOS. Subjects were excluded if the diagnosis was listed only on the death certificate or the autopsy, or if age was unknown.
For recurred subjects, data from Pranav Gwalani is used (who in turn inquired SEER-Medicare databases). Survival chances are evaluated every 12 months after diagnosis (0-month, 12-month, …, 120-month). The covariates for this data are cancer subtype (endometrioid, non-endometrioid, sarcoma, other), race (NH white, NH black), AJCC stage (I - IV) and age at diagnosis in 5 years blocks. For cycle times not matching 12 months and age at diagnosis, linear interpolation is applied.

# References

1. Peter Sasieni, Rebecca Smittenaar, Earl Hubbell, John Broggio, Richard D Neal, Charles Swanton. Modelled mortality benefits of multi-cancer early detection screening in England. British Journal of Cancer. Nature Publishing Group UK London; 2023;129(1):72–80.
2. James F Epperson. An introduction to numerical methods and analysis. John Wiley & Sons; 2013.
3. Peter C Austin, Jason P Fine. Practical recommendations for reporting F ine-G ray model analyses for competing risk data. Statistics in medicine. Wiley Online Library; 2017;36(27):4391–4400.
4. Yizeng He, Kwang Woo Ahn, Ruta Brazauskas. Review of Competing Risks Data Analysis. Medical College of Wisconsin; 2021;



Icahn School of Medicine at **Mount Sinai**



# Output Overview

## Summary

Definitions and methodologies for the basic model outputs.

## Overview

## Calibration Targets

- age-adjusted SEER / Columbia uterine cancer incidence
  - endometrioid
  - non-endometrioid
  - sarcoma
  - other
- SEER AJCC stage distribution and survival by histology and race
  - AJCC I
  - AJCC II
  - AJCC III
  - AJCC IV
- age-adjusted SEER uterine cancer mortality
  - total mortality

## Output Listing

The model creates a variety of outputs. The 'main' output is the state of each subject at each point in time, including other parameters such as BMI. Due to the large size of this output, it is not saved, but a certain subset of the data is retained.

Derived outputs include:

- the incidence of uterine cancer subtypes over time for one birth cohort
- the state distribution over the duration of the simulation (age 35 to 85)
- the difference of cancer incidence in different BMI groups.
- age-adjusted uterine cancer incidences by subtype
- age-adjusted uterine cancer mortality by subtype and combined, by race
- incidence of recurrence over time for one birth cohort
- age-adjusted incidence of recurrence by subtype and combined, by race

# Results Overview

## Summary

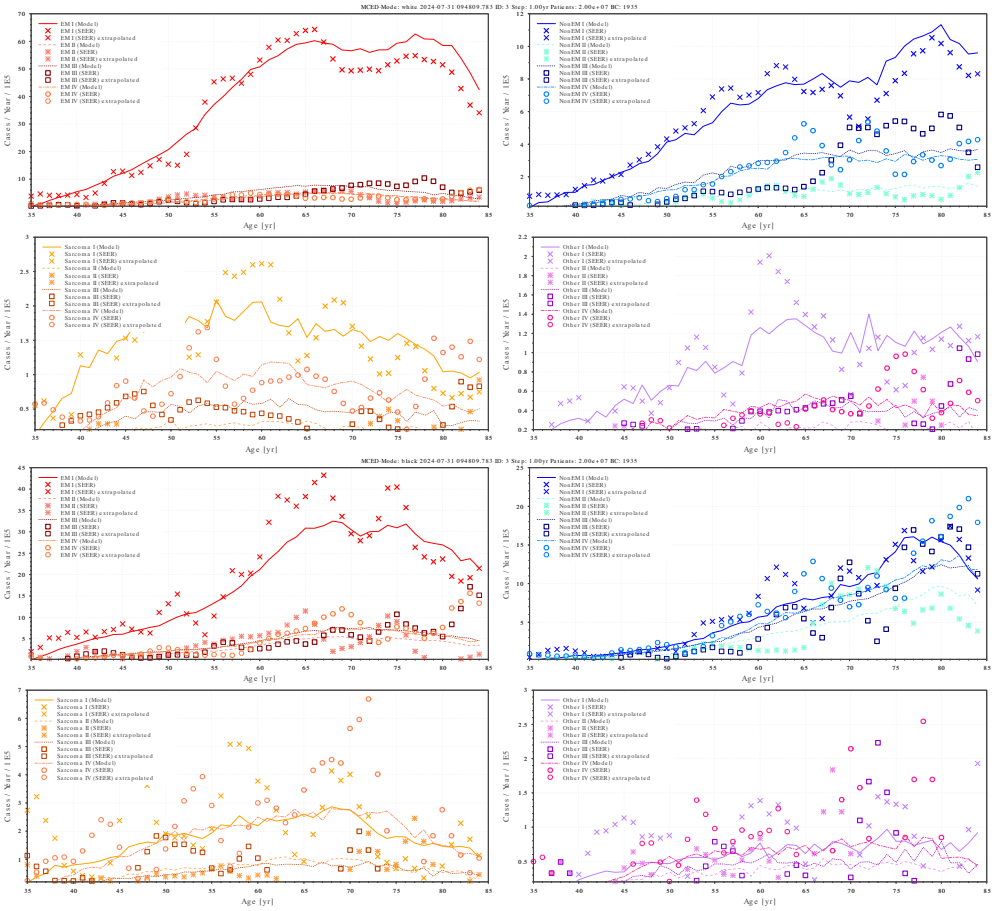
A guide to the results obtained from the model.

## Overview

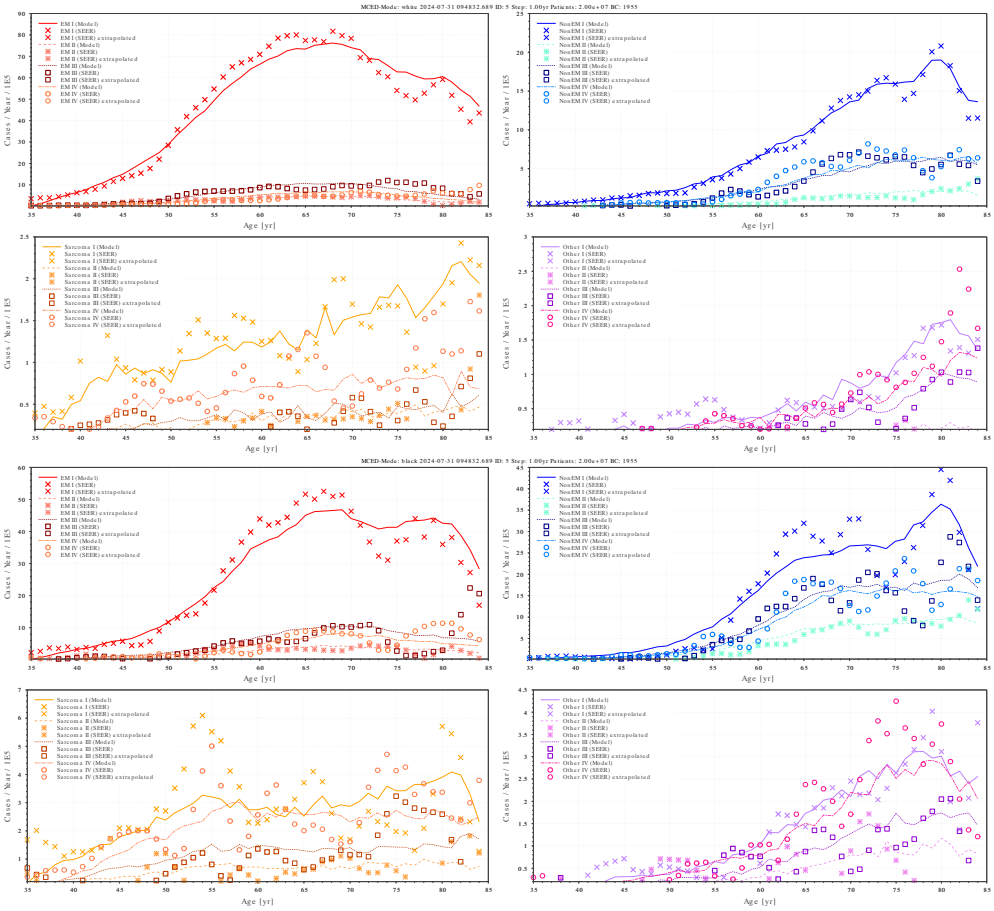
## Results List

### Calibration results

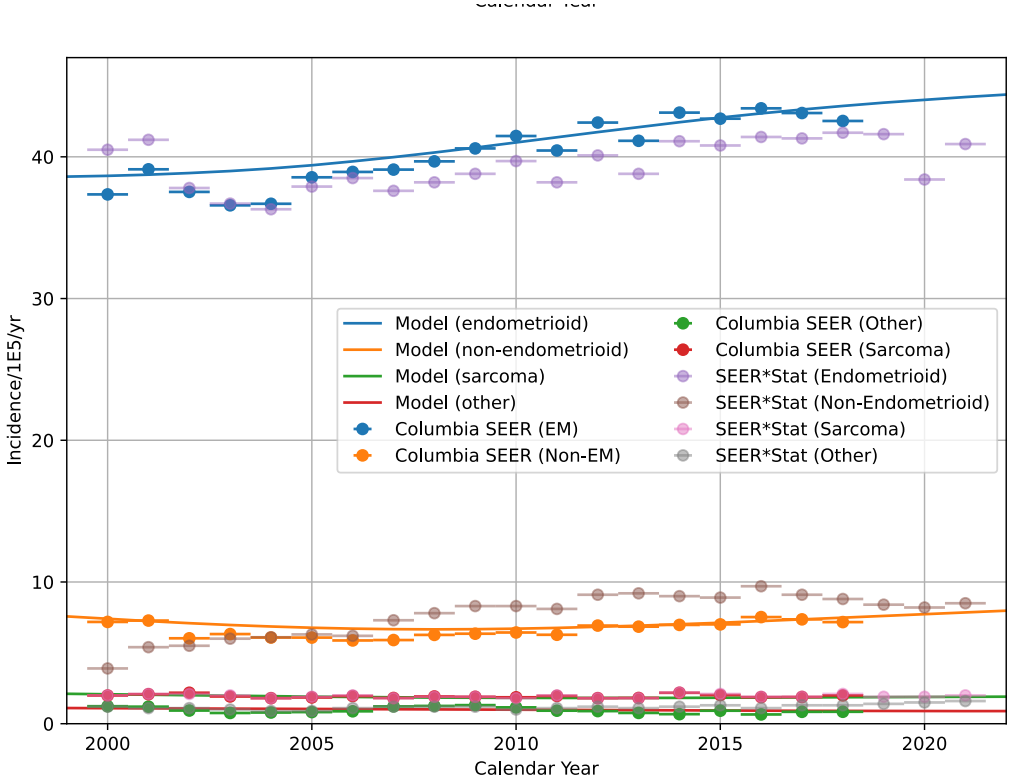
1. Calibration of the MUSIC model by birth cohort 1935, histologies, and race. Model projections are shown by lines, and SEER calibration targets are shown by symbols in matching colors for each stage. Bold symbols represent SEER data, and regular symbols represent projections of SEER data. Data $y(t)$ have been smoothed as such: $y(t) = \frac{y(t-2)+2\cdot y(t-1)+3\cdot y(t)+2\cdot y(t+1)+y(t+2)}{9}$, with $t$ being in units of years. For different cycle lengths, the formula is adapted to reflect the need to incorporate more/less bins.

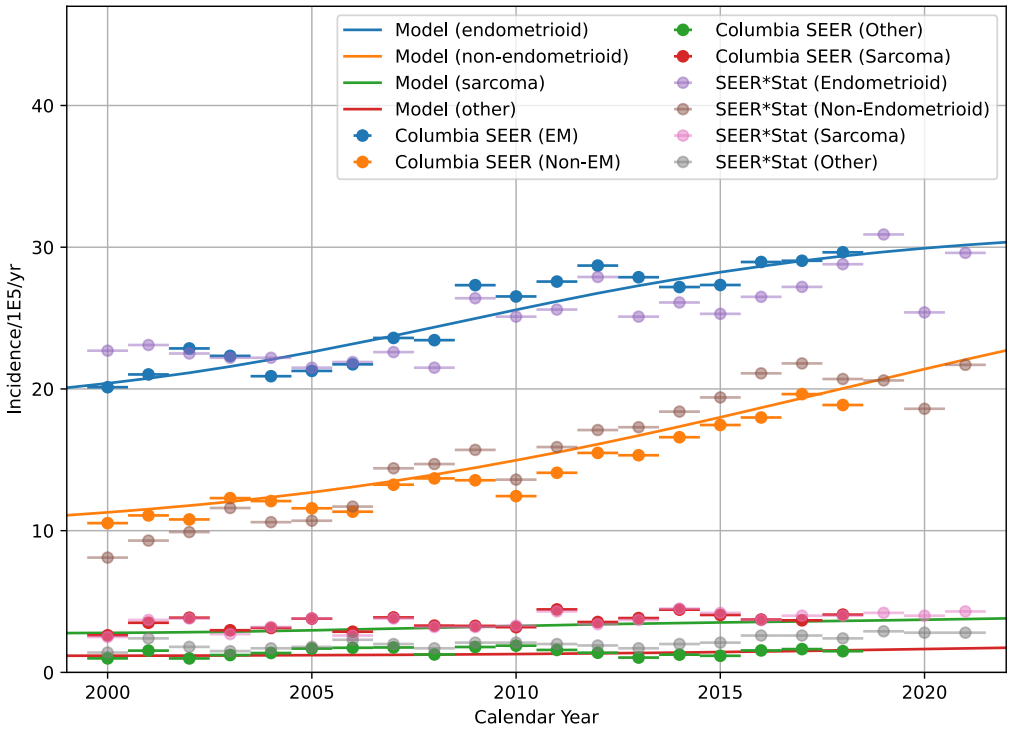


2. Calibration of the MUSIC model by birth cohort 1955, histologies, and race. Model projections are shown by lines, and SEER calibration targets are shown by symbols in matching colors for each stage. Bold symbols represent SEER data, regular symbols represent projections of SEER data. Data $y(t)$ have been smoothed as such: $y(t) = \frac{y(t-2)+2\cdot y(t-1)+3\cdot y(t)+2\cdot y(t+1)+y(t+2)}{9}$, with $t$ being in units of years. For different cycle lengths, the formula is adapted to reflect the need to incorporate more/less bins.

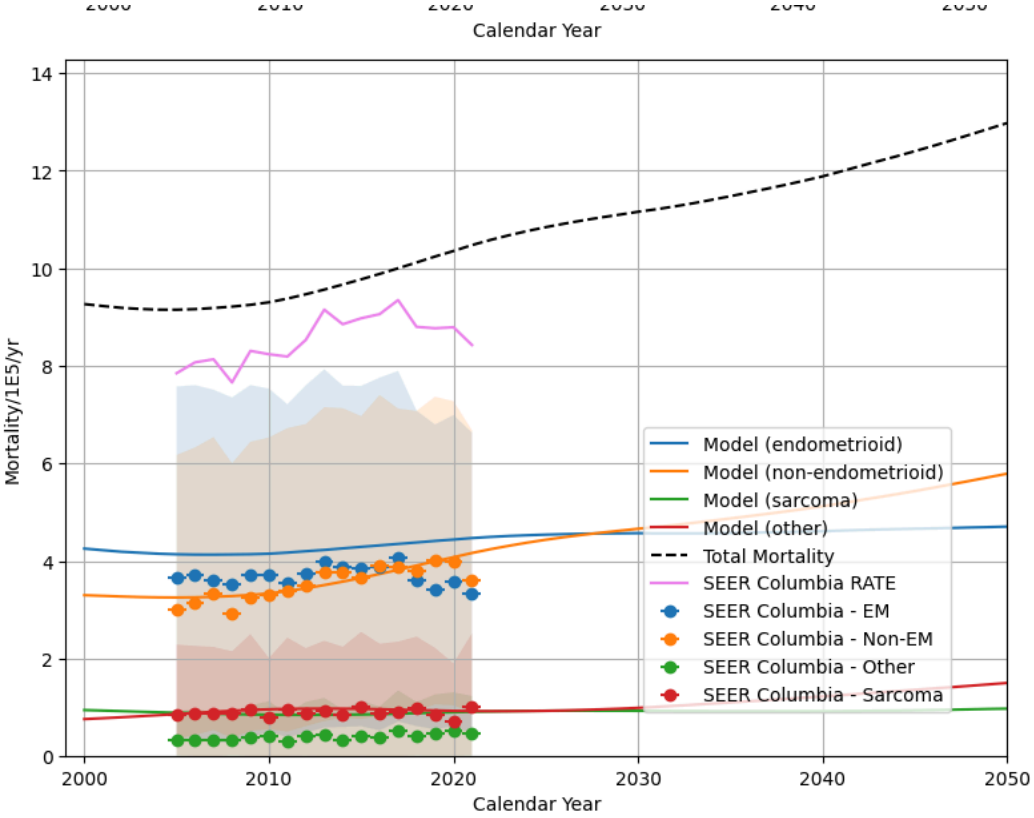


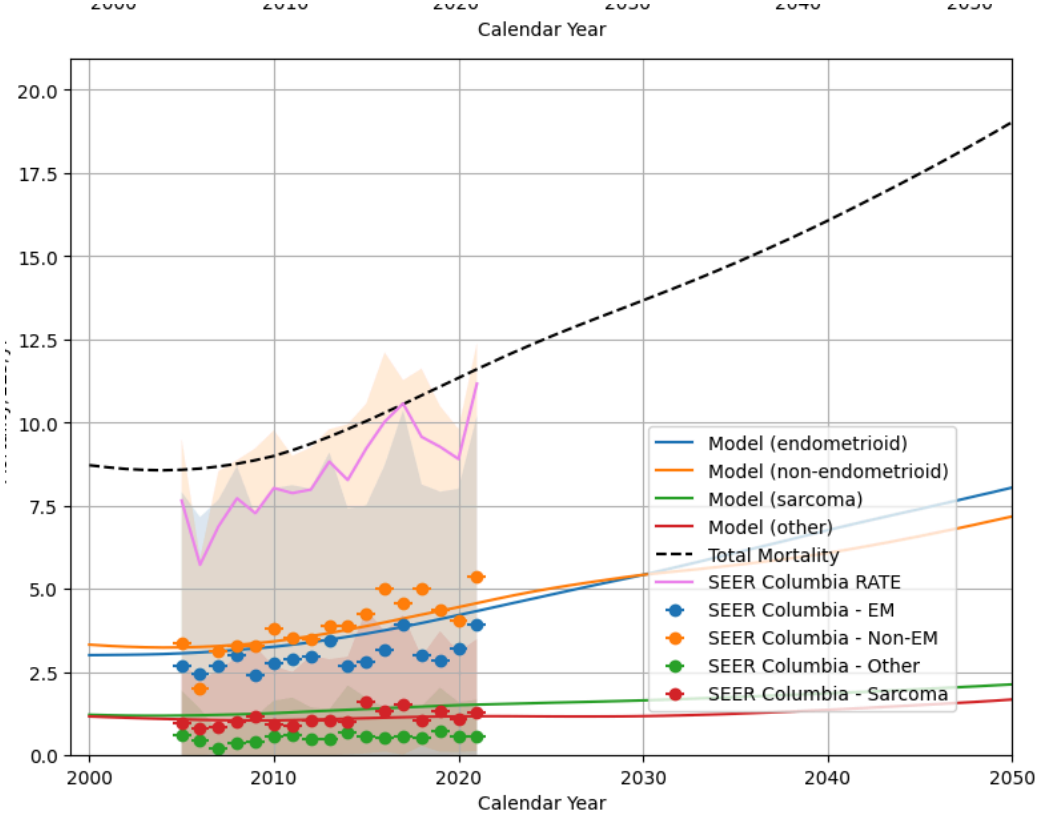
3. Calibration of the age-adjusted uterine cancer incidences by histology for NH White women (top) and NH Black women (bottom). The calibration targets are labeled 'Columbia SEER', which do contain multiple imputed data. The 'SEER*Stat' data is from SEER Nov 2023 Sub.






4. Preliminary benchmark of the MUSIC model to the age-adjusted uterine cancer mortality by histology for NH White women (top) and NH Black women (bottom). The benchmark targets are labeled 'SEER Columbia', which do contain multiple imputed data. 'SEER Columbia RATE' denotes the total mortality rate.









# Key References

Peter C Austin, Jason P Fine. Practical recommendations for reporting F ine-G ray model analyses for competing risk data. Statistics in medicine. Wiley Online Library; 2017;36(27):4391–4400.

Global Burden of Disease Collaborative Network. Global Burden of Disease Study. Seattle, United States: Institute for Health Metrics; 2021;

Centers for Disease Control, National Center for Health Statistics Prevention. National Vital Statistics System, Mortality 1999-2020 on CDC WONDER Online Database. 2021;

James F Epperson. An introduction to numerical methods and analysis. John Wiley & Sons; 2013.

Deborah Grady, Tebeb Gebretsadik, Karla Kerlikowske, Virginia Ernster, Diana Petitti. Hormone replacement therapy and endometrial cancer risk: a meta-analysis. Obstetrics & Gynecology. Elsevier; 1995;85(2):304–313.

Baoxia Gu, Xiaogai Shang, Mengqing Yan, Xiao Li, Wei Wang, Qi Wang, et al. Variations in incidence and mortality rates of endometrial cancer at the global, regional, and national levels, 1990–2019. Gynecologic Oncology [Internet]. 2021;161(2):573–580. Available from: https://www.sciencedirect.com/science/article/pii/S0090825821000962

Yizeng He, Kwang Woo Ahn, Ruta Brazauskas. Review of Competing Risks Data Analysis. Medical College of Wisconsin; 2021;

Michael L Hicks, Maya M Hicks, Roland P Mathews, Dineo Khabele, Camille A Clare, Onyinye Balogun, et al. Racial disparities in endometrial cancer: Where are we after 26 years? Gynecologic oncology. Elsevier; 2024;184:236–242.

Ruth M Pfeiffer, Yikyung Park, Aimée R Kreimer, James V Lacey Jr, David Pee, Robert T Greenlee, et al. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS medicine. Public Library of Science San Francisco, USA; 2013;10(7):e1001492.

Gillian K Reeves, Kirstin Pirie, Valerie Beral, Jane Green, Elizabeth Spencer, Diana Bull. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Bmj. British Medical Journal Publishing Group; 2007;335(7630):1134.

Andrew G Renehan, Margaret Tyson, Matthias Egger, Richard F Heller, Marcel Zwahlen. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. The lancet. Elsevier; 2008;371(9612):569–578.

Peter Sasieni, Rebecca Smittenaar, Earl Hubbell, John Broggio, Richard D Neal, Charles Swanton. Modelled mortality benefits of multi-cancer early detection screening in England. British Journal of Cancer. Nature Publishing Group UK London; 2023;129(1):72–80.

Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. Wiley Online Library; 2021;71(3):209–249.

Liu Yang, Yue Yuan, Rongyan Zhu, Xuehong Zhang. Time trend of global uterine cancer burden: an age-period-cohort analysis from 1990 to 2019 and predictions in a 25-year period. BMC women's health. Springer; 2023;23(1):384.

Harry K Ziel, William D Finkle, Sander Greenland. Decline in incidence of endometrial cancer following increase in prescriptions for opposed conjugated estrogens in a prepaid health plan. Gynecologic oncology. Elsevier; 1998;68(3):253–255.