Combined Model Profile
Version: 1.0.00
Released: 2025-09-30

Uterine Cancer Combined Model Profile

Individual Model Profiles

The Columbia University Uterine Cancer Model (CU-UTMO): Model Profile
Columbia University
Version: 1.0.00
Released: 2025-09-30

Duke University Uterine Cancer Model (DU-CAM): Model Profile
Duke University
Version: 1.0.00
Released: 2025-09-30

Mount Sinai Uterine Cancer Model (MUSIC): Model Profile
Icahn School of Medicine at Mount Sinai
Version: 1.0.00
Released: 2025-09-30

Suggested citation

CISNET Uterine Working Group. Uterine Cancer Combined 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-combined-model-profile-1.0.00-2025-09-30.pdf

Combined Model Profile Version Table

Version	Date	Notes
1.0.00	2025-09-30	Initial release

Columbia University
Version: 1.0.00
Released: 2025-09-30

The Columbia University Uterine Cancer Model (CU-UTMO): Model Profile

Columbia University

Contact

Dr. Jason D. Wright (jw2459@columbia.edu)
Kevin J. Rouse (kr3164@columbia.edu)

Funding

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

Suggested Citation

Wright JD, Rouse K, Prest M. The Columbia University Uterine Cancer Model (CU-UTMO): 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-cu-utmo-model-profile-1.0.00-2025-09-30.pdf

Version Table

Version	Date	Notes
1.0.00	2025-09-30	Initial release

Columbia University
Readers Guide

COLUMBIA UNIVERSITY

Reader's Guide

Core Profile Documentation

These topics will provide an overview of the model without the burden of 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.

Columbia University
Model Purpose

Model Purpose

Summary

This document describes the primary purpose of the Columbia University uterine cancer model (UTMO).

Purpose

The aim of this model is to be used alongside other CISNET models for comparative exercises in projecting trends in uterine cancer in the US. The guiding premise of this model has been to build on a baseline biological understanding of uterine cancer carcinogenesis, incorporating a range of data sources prioritized by quality of the data. The model estimates incidence and mortality of cancer, by AJCC stage, for both endometrioid and non-endometrioid uterine cancers in Non-Hispanic Black and White women in the US. We are in the process of incorporating uterine sarcomas into the model as well. The model can be used to evaluate cancer control and prevention strategies, changes in epidemiologic and clinical factors over time and new treatment strategies.

UTMO is also designed to project future trends, providing critical insights into how the disease might evolve over time in different demographic groups. Model outputs include sojourn time, which helps to estimate the duration of the preclinical detectable phase, along with detailed projections of incidence and mortality by AJCC stage. It also evaluates key outcomes like the 5-year survival rate, allowing for a nuanced analysis of disease progression and survival trends. These outputs are essential for comparing the long-term impact of different screening and treatment strategies on patient outcomes, further guiding policy decisions.

Columbia University
Model Overview

Model Overview

Summary

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

Purpose

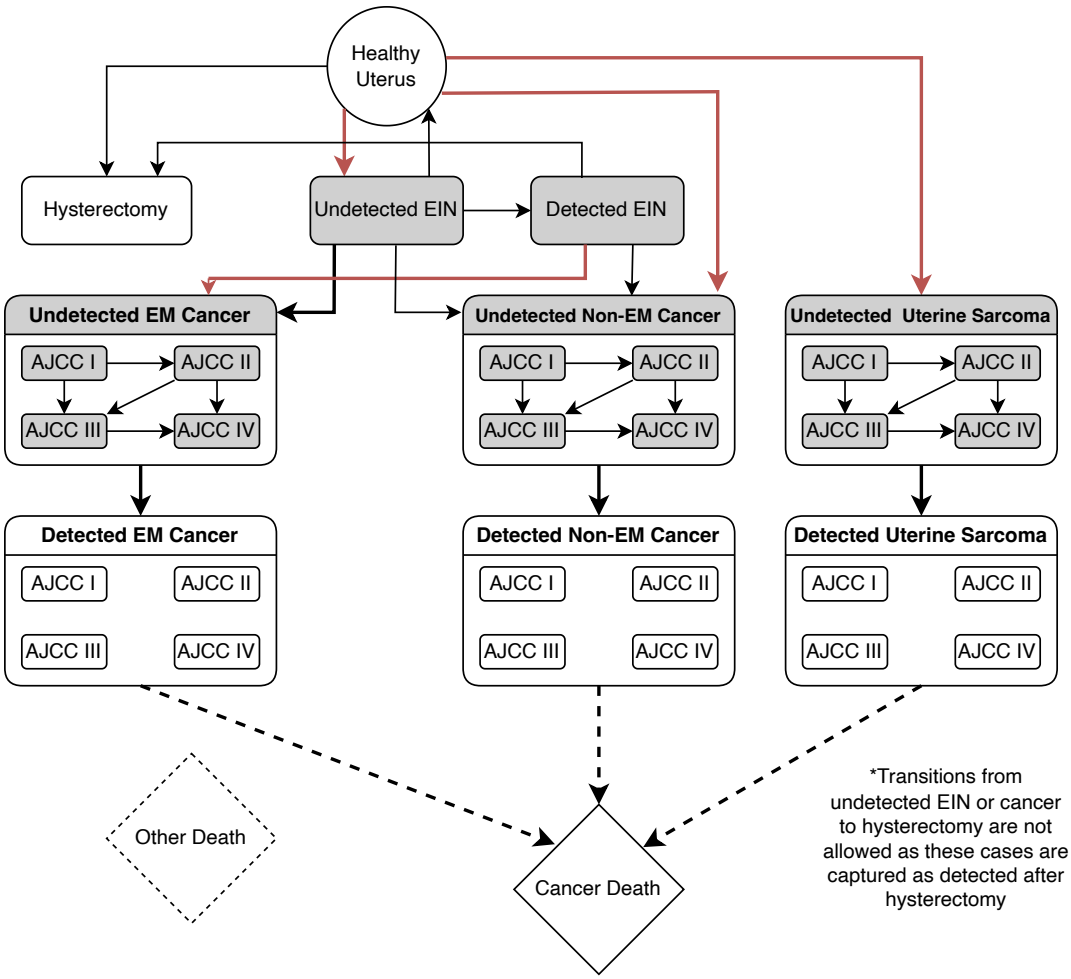
Model Purpose

Background

Uterine cancer is the most common gynecologic malignancy in the United States, with an estimated 67,880 cases and 13,250 deaths projected in 2024[1]. The last thirty years have seen a dramatic rise in both the incidence and death rate from uterine cancer[2]. The increase in the incidence and mortality of uterine cancer is likely driven by a number of factors. Uterine cancer is predominately a disease of older women and the aging of the population in the U.S. has undoubtedly contributed to these trends[3]. Likewise, the rising rate of overweight and obesity has influenced the changing trends in uterine cancer[4]. Adipocytes produce estrogen which can stimulate the endometrium making obesity one of the strongest risk factors for uterine cancer[5]. However, the aging of the population and rising prevalence of obesity likely only explain a part of the changing trends in uterine cancer incidence and mortality. Planning for cancer control and prevention activities requires an understanding of the projected burden of a given cancer. The changing risk profile for uterine cancer in the U.S. has challenged the ability to accurately forecast the burden of the disease. UTMO is a natural history model for uterine cancer calibrated to the population-based incidence and mortality of the disease.

Model Description

The model schematic below shows the possible states and transitions in UTMO, where solid arrows represent possible transitions, dashed arrows indicate terminal transitions, and all states are connected to ‘Other Death’. The transition probabilities between states have a complex structure of inputs and calibrated parameters (to be elaborated on later). In general, the transition probabilities vary by single year age, and 10-year birth cohort categories, the earliest cohort being 1910-1919. The state structure of the transition pathways in this model is based on an assumption of carcinogenesis of endometrioid cancers being mediated through a precancerous state, endometrial intraepithelial neoplasia (EIN). An additional crossover pathway, allowing for limited transitions from EIN and undetected EM cancer to non-endometrioid cancer is included based on the results of natural history data and a dedicated EIN precursor model, which showed up to 7% of non-endometrioid uterine cancers could occur via the same intermediary state. Due to the differing precursor structure and significantly lower incidence, sarcoma cases were excluded from the initial model. However, we are currently in the process of incorporating sarcomas, and the sarcoma pathway is also shown in the shcematic. Regression was restricted for all states except undetected EIN to healthy, as well as cancer recurrence. This improved identifiability by limiting non-primary pathways, enabling more accurate estimation of first-time cancer progression across states.



Time was incorporated into the model in two ways: age-specific transition probabilities and changing probabilities by birth cohort in 10-year groups. This captures secular trends and improves future projections. For transitions like mortality after diagnosis, where age alone wasn't sufficient, and for future screening and intervention modeling, a simple Markov model was inadequate. A patient-level microsimulation with state duration tracking was required.

The model integrates a range of data sources and literature estimates, prioritized by sample size and representativeness. Studies that did not stratify by race/ethnicity were given lower priority. CDC Wonder data and non-uterine cancer annual mortality estimates by age, birth cohort, and race/ethnicity were converted into monthly transition probabilities and directly incorporated into the model.

AJCC staging information (where available) in conjunction with historic (localized/regional/distant) data were extracted for each case listing. Multiple Imputation through Chained Equations (MICE) was used to re-classify patients with not otherwise specified histology as either endometrioid (EM) or non-endometrioid (Non-EM) and impute missing stage information based on all other available variables using the mice R package. The case listings data were then converted into incidence data, stratified by age, stage, race/ethnicity and birth cohort, and used as the primary calibration target under a mean squared error objective function. Surveillance, Epidemiology, and End Results Program (SEER) survival data were converted into duration-dependent hazard functions up to 10 years after diagnosis and imported directly at the microsimulation level.

SEER 18 case listings data was extracted with the following histology groups:

Group	ICD Codes
Endometrioid (EM)	8380-8383, 8480, 8570
Non-Endometrioid (Non-EM)	8020, 8050, 8260, 8310, 8441, 8460-8461, 8950-8951, 8980-8982
Not otherwise specified (NOS)	8000, 8140, 8255, 8323, 8481, 8560

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024 Jan;74(1):12–49.
2. Henley SJ, Miller JW, Dowling NF, Benard VB, Richardson LC. Uterine Cancer Incidence and Mortality - United States, 1999-2016. MMWR Morb Mortal Wkly Rep. 2018 Dec;67(48):1333–1338.
3. Lu KH, Broaddus RR. Endometrial Cancer. N Engl J Med. 2020 Nov;383(21):2053–2064.
4. Friedenreich C, Cust A, Lahmann PH, et al. Anthropometric factors and risk of endometrial cancer: the European prospective investigation into cancer and nutrition. Cancer Causes Control. 2007 Apr;18(4):399–413.
5. Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. Nat Rev Cancer. 2011 Dec;11(12):886–95.

Assumption Overview

Summary

An overview of the basic assumptions inherent in this model.

Background

Due to the unavailability of data or to avoid model complexity, some assumptions about the model were made.

Assumption Listing

No Regression: Regression from a higher to lower disease state was not permitted except for the regression of Undetected EIN to Healthy. Cancer recurrence is also not a part of the model with the goal of improving identifiability by limiting non-primary pathways, enabling more accurate estimation of first-time cancer progression across states. Not including regression or recurrence is an example of prioritizing model simplicity and interpretability.

Markov Property: Our natural history model assumes that a patient's health state is only dependent on their previous state. For example, a patient who has had stage II uterine cancer for four years has the same transition probabilities as a patient who has had stage II uterine cancer for one year (given that they have the same age, gender, cohort, etc). To account for previous states, a patient-level microsimulation is run which can account for things such as survival by years of follow-up after cancer diagnosis

Natural History: Our model consists of various assumptions about possible transitions that are based on the natural history of the disease:

- All cancers arise in an undetected state.
- While undetected, cancers may either persist in the given undetected stage, progress to a higher stage, or progress to a detected cancer of the corresponding stage if the tumor is identified clinically.
- Some individuals will die from other causes prior to the detection of the underlying uterine cancer.
- Obesity, survival, and hysterectomy rates are assumed not to change from their most recent available data point through 2050.

Parameter Overview

Summary

Describes the basic parameter set used to inform the model; more detailed information is available for each specific parameter.

Background

The main parameters in the model are the transition probabilities from one state to another. These parameters are calibrated and are what primarily drive the model results.

Parameter Listing Overview

The parameters of the model are guided by the possible transitions. Refer to the schematic in Model Purpose. For each of these possible transitions, there is a certain probability of transitioning from one state to another. These parameters are the primary drivers of the model's functionality and results and are calibrated using simulated annealing.

Table 1. Input parameters and calibration targets for the natural history model

Category	Implementation	Data Source	Example
Model inputs			
Obesity prevalence	Simulated BMI category by age and birth cohort	NHANES (2000–2020) [1]	1960–1970 cohort, NH White female, age 50, 40.5% BMI ≥ 30
Impact of obesity	Fixed transition risk ratio	Zhao et al. (2021) [2]; Epplein et al. (2008) [3]	BMI ≥ 25: OR = 2.7 for EMC vs. BMI < 25
Hysterectomy rates	Monthly, race and age-specific	NHANES (2000–2020) [1]	NH Black women, 45–49, P = 0.002
Hysterectomy mortality	Competing hazard	Wingo et al. (1985) [4]	Mortality (non-pregnancy/cancer): 6.0 per 10,000
All-cause mortality	Monthly, race/cohort/age-specific	CDC WONDER (1968–2016) [5]	1950–1960 cohort, NH White, age 55–59, P = 0.0008
Cancer-specific survival	10-year survival hazards	SEER (2000–2018)	NH White, 40–44, EM stage IV: P = 0.0078 (1st year)
Cancer sojourn time	Upper/lower bounds	Broder et al. (2021) [6]	Stage III–IV: median = 1.5 years
Calibration targets			
EIN incidence	Age-specific	Reed et al. (2009) [7]	Age 60–65: 28.9 per 100,000
EIN progression	Undetected progression bounds	Lacey et al. (2008) [8]	3-year risk = 8.2%
EIN prevalence	Age-specific	Korhonen et al. (1997) [9]	Age 45–55: subclinical = 0–2 per 3,000
EM cancer prevalence	Autopsy and screening	Horwitz (1981) [10]; Göl (2001) [11]	Rate = 22–31 per 10,000
Cancer incidence	Race/histology/stage/age/cohort	SEER (2000–2018)	1940–1950 cohort, NH Black, age 70, EM I: 38.8 per 100,000

Abbreviations: EIN: endometrial intraepithelial neoplasia. EM: endometrioid. NH: non-Hispanic. Non-EM: non-endometrioid. OR: odds ratio. SEER: Surveillance, Epidemiology, and End Results.

The NHANES reproductive history [1] survey was used to extract age- and race/ethnicity-specific hysterectomy hazards which were imported directly into the model. The NHANES reproductive history survey also provided BMI category data (<25, 25–30, >30), incorporated into the microsimulation. Age-dependent odds ratios linking BMI to EIN and endometrioid cancer progression were applied based on [2]. Hysterectomy prevalence and its associated small mortality risk [4] were also included.

Cancer incidence and survival data were obtained from SEER and the literature [5,6]. SEER data were stratified by race/ethnicity, histology, stage, age, and cohort. Subclinical and clinically detected prevalence estimates came from autopsy and screening studies [7-11]. SEER data also informed survival probabilities, adjusted by race, stage, histology, and cohort.

References

1. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. 2020;
2. Zhao J, Hu Y, Zhao Y, Chen D, Fang T, Ding M. Risk factors of endometrial cancer in patients with endometrial hyperplasia: implication for clinical treatments. BMC Womens Health. 2021 Aug;21(1):312.
3. Epplein M, Reed SD, Voigt LF, Newton KM, Holt VL, Weiss NS. Risk of complex and atypical endometrial hyperplasia in relation to anthropometric measures and reproductive history. Am J Epidemiol. 2008 Sep;168(6):563–70; discussion 571-6.
4. Wingo PA, Huezo CM, Rubin GL, Ory HW, Peterson HB. The mortality risk associated with hysterectomy. Am J Obstet Gynecol. 1985 Aug;152(7 Pt 1):803–8.
5. Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System, Mortality 1999-2020 on CDC WONDER Online Database. 2020;
6. Broder MS, Ailawadhi S, Beltran H, et al. Estimates of stage-specific preclinical sojourn time across 21 cancer types. American Society of Clinical Oncology. 2021;
7. Reed SD, Newton KM, Clinton WL, et al. Incidence of endometrial hyperplasia. Am J Obstet Gynecol. 2009 Jun;200(6):678 e1–6.
8. Lacey JV, Jr., Mutter GL, Nucci MR, et al. Risk of subsequent endometrial carcinoma associated with endometrial intraepithelial neoplasia classification of endometrial biopsies. Cancer. 2008 Oct;113(8):2073–81.
9. Korhonen MO, Symons JP, Hyde BM, Rowan JP, Wilborn WH. Histologic classification and pathologic findings for endometrial biopsy specimens obtained from 2964 perimenopausal and postmenopausal women undergoing screening for continuous hormones as replacement therapy (CHART 2 Study). Am J Obstet Gynecol. 1997 Feb;176(2):377–80.
10. Horwitz RI, Feinstein AR, Horwitz SM, Robboy SJ. Necropsy diagnosis of endometrial cancer and detection-bias in case/control studies. Lancet. 1981 Jul;2(8237):66–8.
11. Gol K, Saracoglu F, Ekici A, Sahin I. Endometrial patterns and endocrinologic characteristics of asymptomatic menopausal women. Gynecol Endocrinol. 2001 Feb;15(1):63–7.

Component Overview

Summary

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

Overview

The model is a state-transition microsimulation model where patients can transition to and from certain states with certain probabilities (visual in Model Overview). There are various components to the model that attempt to model the natural history of the disease as well as certain risk factors such as obesity.

Component Listing

Natural History Component:

The natural history component is the main component of the model. The goal of this component is to simulate the natural history of uterine cancer. The model begins at 18 years of age and uses a one-month cycle length to dynamically account for short duration changes in risk factors and health states over time. The model simulates non-Hispanic White and non-Hispanic Black women separately.

The model has four general states: healthy (unaffected), precursor lesion (endometrial intraepithelial neoplasia), cancer, and death. Individuals with uterine cancer may have two disease states. Undetected uterine cancer is based on the true presence or absence of cancer for a given tumor stage that has not yet been detected clinically or pathologically. Detected uterine cancer is also based on the true presence or absence of a cancer for a given tumor stage, but one that has been clinically identified based on a diagnostic test.

Calibration Component:

The calibration component is primarily responsible for adjusting probabilities (besides for death and hysterectomy rates which can be derived from available data). Transition rates change based on age, race, and birth cohort, and there is not enough data to accurately capture these changes without calibration. For some of the transition probabilities, the ranges were constrained based on the literature and domain knowledge. The Surveillance, Epidemiology, and End Results (SEER) program was used to obtain cancer incidence and survival in the United States.

Calibration of UTMO was performed in two phases, a multicohort phase and a cohort-specific phase. The motivation for this methodology was to incorporate as much of the available data as possible, while being able to forecast future trends. As some target data had limited sample sizes, high variance or insufficient information to stratify by birth cohort, this information was incorporated into the multicohort phase. The parameters from the multicohort phase were then used as starting values for the cohort-specific phase, where birth cohorts were grouped into 10-year intervals beginning with 1910-1920.

Detection Component:

The detection component controls how individuals move from undetected to detected cancer states. Detection probabilities are governed by constraints that maintain biological plausibility and consistency with published evidence. For pre-invasive disease, detection of endometrial intraepithelial neoplasia (EIN) is constrained to always exceed the probability of regression, ensuring that progression to cancer is realistically captured. Detection of endometrioid cancer must exceed EIN detection.

For malignant states, detection rates increase across stages to reflect higher clinical detectability at more advanced stages. Detection rates for non-endometrioid cancers are constrained to be higher than for endometrioid cancers, aligning with their more aggressive natural history and shorter sojourn times. These relationships are enforced programmatically in the model's constrain logic, which adjusts transition probabilities each cycle. Broder et al. (2021)[1] provides uterine cancer sojourn time estimates, which are used to inform baseline detection rates.

Survival Component

The survival component applies stage- and histology-specific survival probabilities to detected cases. Monthly hazards are derived from SEER's cause-specific death variable and reflect survival patterns by birth cohort, age, histology, and tumor stage for up to 120 months after diagnosis. These monthly hazards were calculated using the rstpm2 package in R. These survival hazards are directly applied in the microsimulation to ensure that modeled survival is consistent with population-based data and captures differences by race, histology, stage, age, and cohort.

References

1. Broder MS, Ailawadhi S, Beltran H, et al. Estimates of stage-specific preclinical sojourn time across 21 cancer types. American Society of Clinical Oncology. 2021;

Output Overview

Summary

Definitions and methodologies for the basic model outputs.

Overview

The uterine cancer natural history model produces key outputs that capture the burden and progression of the disease, allowing for insights into future trends and disparities across different populations. The model tracks incidence, prevalence, and mortality for uterine cancer, focusing on how these metrics evolve over time. By stratifying these results by race, age, birth cohort, and tumor histology (endometrioid vs. non-endometrioid; sarcoma to be added upon completion), the model provides an overview of how uterine cancer impacts different groups.

Output Listing

1. Stage Distribution: The model outputs AJCC stage distribution, which is informative of cancer detection patterns by histology and can indicate screening disparities between certain groups.
2. Sojourn time: The model estimates sojourn time, which is the duration that women remain in undetected stages of uterine cancer. This is done using two different intervals: a) pre-invasive sojourn time (time between the presentation of endometrial intraepithelial neoplasia to detection) and b) malignancy sojourn time (time between first malignant cell and cancer detection). By calculating the sojourn times separately for NH White and NH Black women, we can infer how quickly uterine cancer progresses in each group.
3. Incidence: Capturing incidence in the model is important for understanding the rate at which new cases of uterine cancer are occurring. Incidence is a foundational measure for tracking the disease's spread and forecasting future trends. By including stratifications such as race, histology, and birth cohort, the model can reveal important disparities in who is getting the disease and how the incidence varies across populations.
4. 5-year survival rate: The model calculated 5-year survival rates which simulate patient outcomes after cancer diagnosis. Their survival is dependent on various factors such as tumor type, stage, race, and age.
5. Incidence-Based Mortality: The model captures incidence-based mortality within a 5-year period. This can provide a more accurate measure of cancer mortality as it links deaths to original cancer diagnosis and has a pre-specified time interval allowed between diagnosis and death, making it more interpretable.

Results Overview

Summary

A guide to the results obtained from the model.

Overview

This section presents findings from the Columbia University uterine cancer model (UTMO) comparing stage distributions, incidence, and survival with SEER data to assess model alignment and validity. We also explore uterine cancer incidence by histology and BMI, stratified by race and age, and examine median ages of diagnosis. Further, sojourn times are analyzed across histological subtypes and racial groups. Lastly, results from model stress testing using the Maximum Clinical Incidence Reduction (MCLIR) methodology evaluate intervention effectiveness at different ages and cancer stages.

Results List

Table 1. Incidence and Stage Distribution of Uterine Cancer by year of diagnosis (UTMO vs SEER)

Incidence by Stage	Calibration Incidence (%)			Projection Incidence (%)		
	2000	2010	2018	2030	2040	2050
AJCC I						
SEER	37.3 (75.0)	40.8 (75.0)	42.9 (74.5)	NA	NA	NA
UTMO	33.1 (71.5)	38.0 (71.2)	41.8 (70.5)	44.7 (70.0)	48.5 (69.7)	52.4 (68.9)
AJCC II						
SEER	3.8 (7.6)	3.0 (5.5)	2.6 (4.5)	NA	NA	NA
UTMO	4.1 (9.0)	4.6 (8.7)	4.8 (8.2)	4.9 (7.7)	5.3 (7.6)	5.6 (7.4)
AJCC III						
SEER	4.5 (9.1)	6.7 (12.2)	6.8 (11.9)	NA	NA	NA
UTMO	4.7 (10.2)	6.1 (11.3)	7.2 (12.1)	8.4 (13.1)	9.4 (13.4)	10.8 (14.2)
AJCC IV						
SEER	4.1 (8.3)	3.9 (7.2)	5.2 (9.1)	NA	NA	NA
UTMO	4.3 (9.3)	4.7 (8.8)	5.4 (9.2)	5.8 (9.2)	6.5 (9.3)	7.2 (9.5)

Incidence per 100,000 in subjects 40 years of age and older (percentage of incidence per 100,000 diagnosed at each stage). Results are based on estimated incidence for non-Hispanic White and non-Hispanic Black women combined.
AJCC: American Joint Commission on Cancer. SEER: Surveillance, Epidemiology, and End Results database. UTMO: Columbia University uterine cancer model.

Table 1 shows encouraging alignment between Columbia's uterine cancer stage distribution and SEER, with UTMO closely matching SEER's data in most categories. For Stage I, UTMO reports 3-4% fewer cases than SEER, demonstrating a strong similarity. In Stage II, UTMO slightly overrepresents cases. For Stage III and Stage IV, UTMO's data aligns closely with SEER. In Stage IV, UTMO's results are more similar to SEER's, with just a slight 0.3% difference in 2018. Overall, UTMO's data is highly comparable to SEER's, indicating a strong alignment in stage distribution.

Table 2. Sojourn Times (Months) Stratified by Race and Histology

Type of Sojourn Time	Histology	NH Black Mean	NH Black SD	NH White Mean	NH White SD
Pre-invasive	Endometrioid	127.6	105.1	97.8	72.9
	Non-endometrioid	27.1	54.6	52.3	73.0
Malignancy	Endometrioid	22.4	19.7	21.9	19.8
	Non-endometrioid	8.2	6.6	8.7	6.6

Two sojourn times were estimated from model outputs by race and histology: 1) pre-invasive sojourn time, which reflects the time from first pre-invasive lesion (undetected EIN state) to detected cancer, and 2) malignancy sojourn time, which reflects the time from first malignant cell (undetected cancer) to clinically detected cancer (Table 3). The sojourn time was shorter for non-endometrioid compared to endometrioid cancers. For example, the malignancy sojourn time for endometrioid uterine cancer was 22.4 months in Black women compared to 21.9 months in White women. The corresponding sojourn times for non-endometrioid tumors were 8.2 and 8.7 months, respectively.

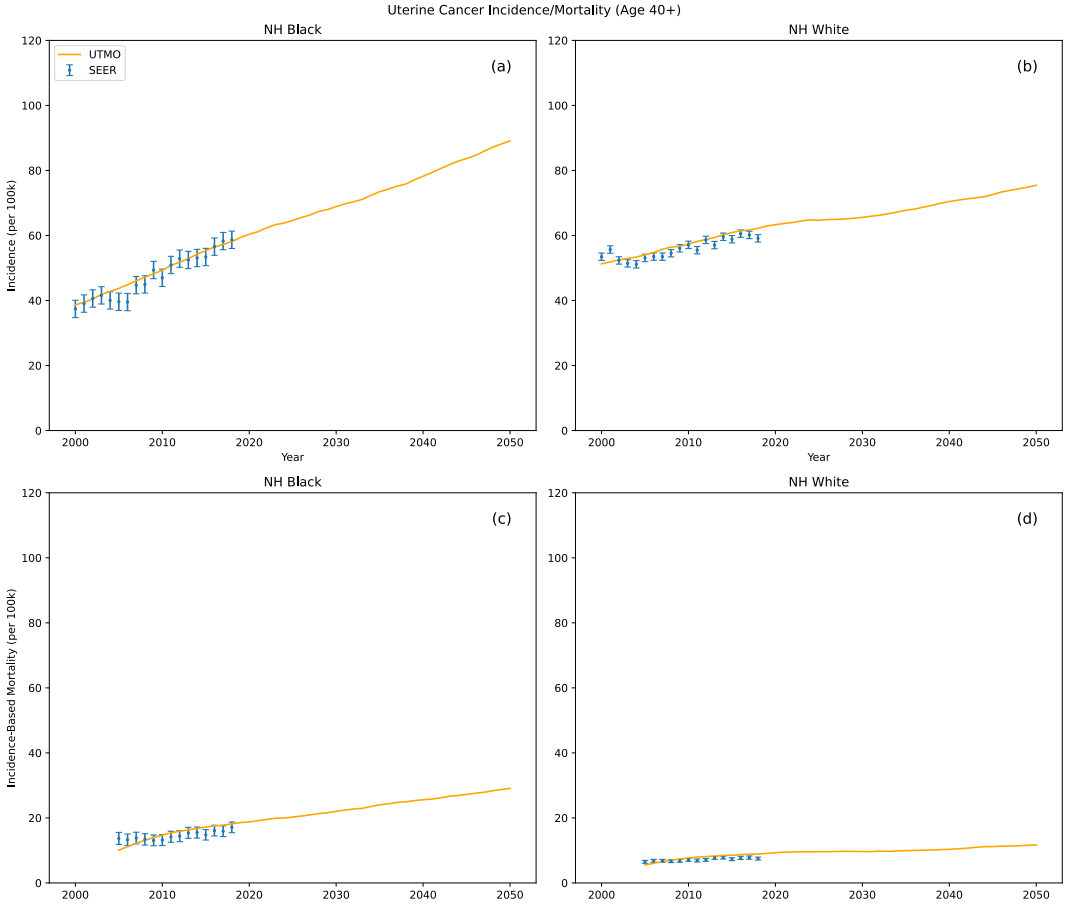


Figure 1. Projected age adjusted uterine cancer incidence and incidence-based mortality among women aged 40+ stratified by race and ethnicity to 2050. A. Incidence in Black women. B. Incidence in White women. C. Incidence-Based Mortality in Black women. D. Incidence-Based Mortality in White women

Table 3. Observed and projected incidence and mortality 5-year survival of uterine cancer overall and stratified by histology.

NH Black Women

Outcome	2000	2010	2018	2030	2040	2050
Incidence per 100,000 (40+)						
Overall (SEER)	36.0 (33.4, 38.6)	45.5 (42.9, 48.1)	56.8 (54.2, 59.4)	-	-	-
Overall (UT-MO)	35.5 (32.1, 38.9)	46.9 (43.5, 50.3)	56.6 (53.2, 60.0)	65.5 (62.1, 68.9)	75.6 (72.2, 79.0)	86.9 (83.5, 90.3)
Endometrioid (SEER)	23.5 (22.1, 24.9)	30.7 (29.3, 32.2)	34.2 (32.8, 35.7)	-	-	-
Endometrioid (UT-MO)	23.0 (21.2, 24.8)	29.6 (27.9, 31.4)	35.2 (33.4, 37.0)	39.3 (37.5, 41.1)	44.7 (42.9, 46.5)	50.5 (48.7, 52.3)
Non-Endometrioid (SEER)	12.5 (11.2, 13.8)	14.8 (13.5, 16.0)	22.5 (21.2, 23.8)	-	-	-
Non-Endometrioid (UT-MO)	12.5 (10.9, 14.1)	17.2 (15.6, 18.8)	21.4 (19.8, 23.0)	26.2 (24.6, 27.8)	30.9 (29.3, 32.5)	36.3 (34.7, 37.9)
5-Year Survival (%)						
Overall (SEER)	66.5 (62.1, 70.5)	68.5 (65.3, 71.6)	69.1 (66.1, 71.8)	-	-	-
Overall (UT-MO)	71.4 (69.5, 73.3)	67.2 (65.4, 69.0)	67.1 (65.3, 68.9)	66.7 (65.2, 68.3)	64.3 (62.8, 65.7)	68.2 (67.0, 69.5)
Endometrioid (SEER)	81.8 (75.4, 86.2)	84.5 (80.9, 87.5)	86.5 (83.3, 89.1)	-	-	-
Endometrioid (UT-MO)	88.8 (87.2, 90.5)	86.5 (84.8, 88.2)	86.4 (84.7, 88.2)	-	-	-
Non-Endometrioid (SEER)	38.4 (30.0, 46.7)	39.0 (32.4, 45.5)	42.9 (37.3, 48.4)	-	-	-
Non-Endometrioid (UT-MO)	44.5 (41.2, 47.8)	40.2 (37.3, 43.2)	37.5 (34.5, 40.5)	41.7 (39.3, 44.0)	40.5 (38.3, 42.6)	43.2 (41.2, 45.1)

NH White Women

Outcome	2000	2010	2018	2030	2040	2050
Incidence per 100,000 (40+)						
Overall (SEER)	52.1 (51.0, 53.2)	55.9 (54.8, 57.0)	57.7 (56.6, 58.9)	-	-	-
Overall (UT-MO)	48.0 (46.3, 49.7)	54.4 (52.7, 56.2)	59.7 (57.9, 61.4)	63.6 (61.9, 65.3)	68.7 (67.0, 70.4)	74.2 (72.5, 75.9)
Endometrioid (SEER)	43.6 (42.7, 44.6)	48.2 (47.3, 49.2)	49.2 (48.3, 50.2)	-	-	-
Endometrioid (UT-MO)	41.0 (39.5, 42.5)	46.7 (45.2, 48.2)	51.2 (49.8, 52.7)	54.5 (53.0, 56.0)	58.8 (57.3, 60.3)	63.4 (61.9, 64.9)
Non-Endometrioid (SEER)	8.5 (8.3, 8.7)	7.7 (7.4, 7.9)	8.5 (8.3, 8.8)	-	-	-
Non-Endometrioid (UT-MO)	7.0 (6.8, 7.3)	7.8 (7.5, 8.0)	8.4 (8.2, 8.7)	9.1 (8.8, 9.3)	9.9 (9.6, 10.1)	10.8 (10.6, 11.1)
5-Year Survival (%)						
Overall (SEER)	86.6 (85.7, 87.5)	86.6 (85.8, 87.4)	89.0 (85.1, 86.8)	-	-	-
Overall (UT-MO)	86.3 (85.0, 87.5)	85.7 (84.4, 87.0)	85.7 (84.4, 87.0)	85.1 (83.9, 86.3)	85.1 (83.9, 86.2)	83.8 (82.6, 84.9)
Endometrioid (SEER)	91.5 (90.5, 92.5)	92.5 (91.7, 93.2)	92.2 (91.4, 92.9)	-	-	-
Endometrioid (UT-MO)	92.3 (91.2, 93.3)	92.4 (91.3, 93.4)	92.6 (91.5, 93.6)	-	-	-
Non-Endometrioid (SEER)	50.0 (45.6, 54.3)	50.0 (46.0, 53.8)	52.4 (48.4, 55.9)	-	-	-
Non-Endometrioid (UT-MO)	50.5 (45.6, 55.4)	52.8 (48.3, 57.3)	45.9 (41.2, 50.7)	48.4 (44.1, 52.7)	51.6 (47.6, 55.6)	45.5 (41.7, 49.3)

SEER: Surveillance, Epidemiology, and End Results database.
UT-MO: Columbia University uterine cancer model.
* 2013 was the last year with 5-year survival data from SEER 18.

The incidence and mortality for uterine cancer have gradually increased over time for both White and Black women (Figure 2, Table 3). Our model closely fit SEER incidence and mortality data available through 2018. For uterine cancer incidence, our model has a square root normalized SSE (NSEE), a measure of average variance from the target, of 0.064 in White women and 0.084 in Black women. For uterine cancer incidence-based mortality, our model has an NSSE of 0.124 in White women and 0.333 in Black women. These results demonstrate excellent model validity.

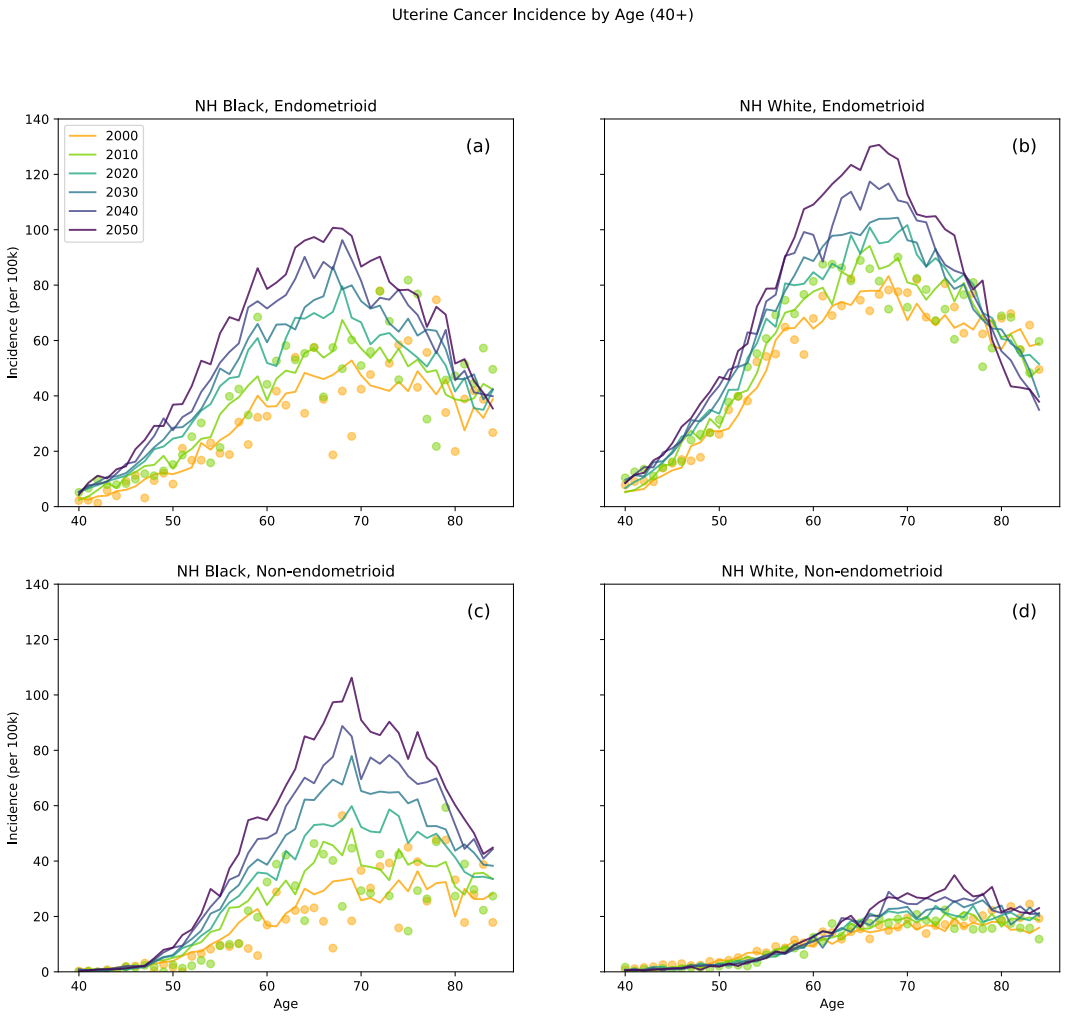


Figure 2. Histology-specific uterine cancer incidence stratified by race and ethnicity and birth cohort among women aged 40+. A. Endometrioid cancer in Black women. B. Endometrioid cancer in White women. C. Non-endometrioid cancer in Black women. D. Non-endometrioid cancer in White women.

The model performed well when additionally stratified by histology and age (Figure 2), while recognizing the additional noise present in single-age observations. The model-predicted median age of diagnosis across all cohorts for EM cancer was 65.4 and 65.7 for White and Black women respectively (compared to actual median ages of 65.7 and 66.9). The predicted median age of diagnosis for non-EM cancer was 70.7 and 68.9 for White and Black women respectively (compared to actual median ages of 70.7 and 69.7).

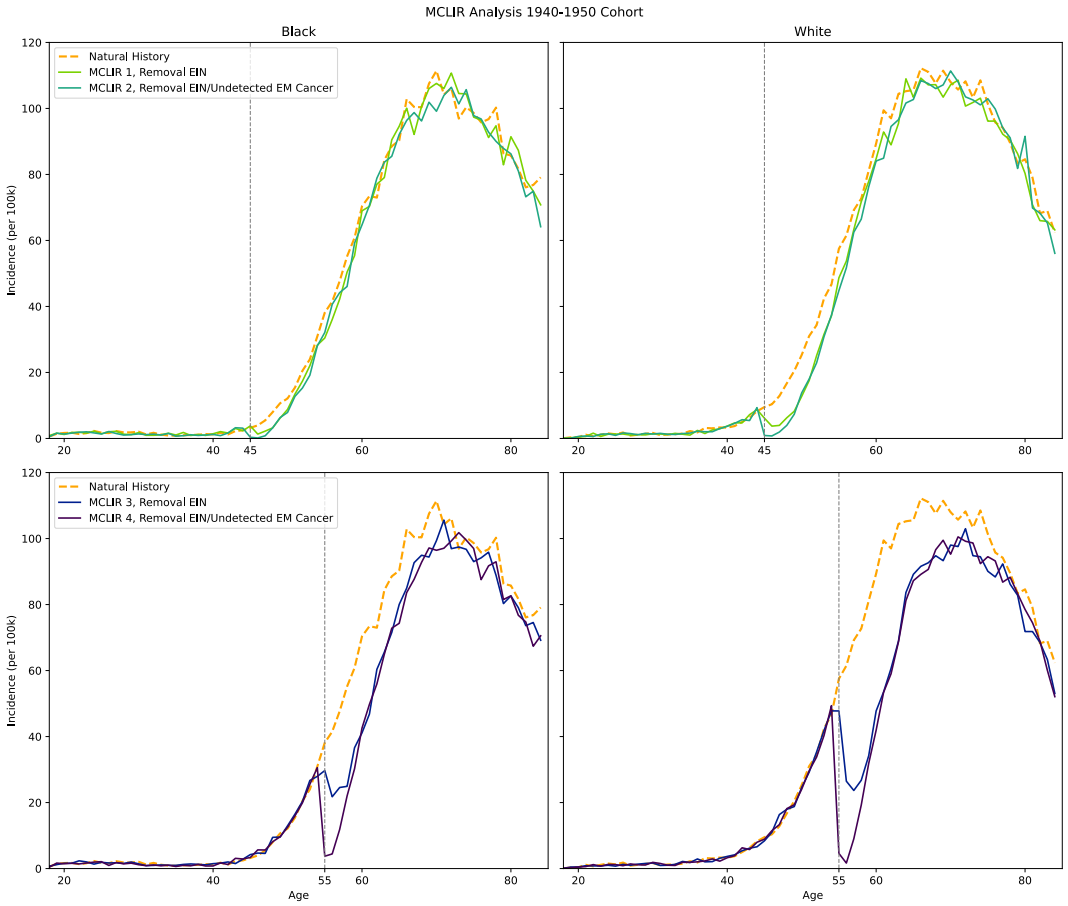


Figure 3. Sensitivity analysis using maximum clinical incidence reduction (MCLIR)[1] scenario analysis in the birth cohort of 1940-1950.
A. Removal of undetected endometrial intraepithelial neoplasia (scenario 1) or undetected endometrioid cancer (scenario 2) in 45-year-old Black patients.
B. Removal of undetected endometrial intraepithelial neoplasia (scenario 1) or undetected endometrioid cancer (scenario 2) in 45-year-old White patients.
C. Removal of undetected endometrial intraepithelial neoplasia (scenario 3) or undetected endometrioid cancer (scenario 4) in 55-year-old Black patients.
B. Removal of undetected endometrial intraepithelial neoplasia (scenario 3) or undetected endometrioid cancer (scenario 4) in 55-year-old White patients.

Results from stress testing of the model using the MCLIR methodology are shown in Figure 3. Scenarios 1 and 2 were applied for women at age 45 and eliminated further risk for patients with undetected EIN alone (scenario 1) or undetected EIN and endometrioid tumors (scenario 2). For both White and Black women there was a decline in cancer incidence starting at age 45 and lasting up to 7 years and 8 years, respectively. Women 55 years of age for scenario 3 (removal of undetected EIN alone) and scenario 4 (removal of undetected EIN and endometrioid tumors) saw a decline in cancer incidence lasting up to 15 and 16 years for White and Black women respectively. The magnitude of decline and the absolute number of cases eliminated was larger when the intervention was applied at 55 compared to 45 years of age. This indicates that potential screening and intervention options are more effective when targeting women aged 55 than those aged 45.

References

1. Kok, Inge M. C. M. de, van Rosmalen, Joost, Castle, Philip E., Berkhof, Johannes. The Impact of Different Screening Model Structures on Cervical Cancer Incidence and Mortality Predictions: The Maximum Clinical Incidence Reduction (MCLIR) Methodology. Medical Decision Making. 2020 Jun;40(4):474.

Key References

Broder MS, Ailawadhi S, Beltran H, et al. Estimates of stage-specific preclinical sojourn time across 21 cancer types. American Society of Clinical Oncology. 2021;

Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. 2020;

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

Epplein M, Reed SD, Voigt LF, Newton KM, Holt VL, Weiss NS. Risk of complex and atypical endometrial hyperplasia in relation to anthropometric measures and reproductive history. Am J Epidemiol. 2008 Sep;168(6):563–70; discussion 571-6.

Friedenreich C, Cust A, Lahmann PH, et al. Anthropometric factors and risk of endometrial cancer: the European prospective investigation into cancer and nutrition. Cancer Causes Control. 2007 Apr;18(4):399–413.

Gol K, Saracoglu F, Ekici A, Sahin I. Endometrial patterns and endocrinologic characteristics of asymptomatic menopausal women. Gynecol Endocrinol. 2001 Feb;15(1):63–7.

Henley SJ, Miller JW, Dowling NF, Benard VB, Richardson LC. Uterine Cancer Incidence and Mortality - United States, 1999-2016. MMWR Morb Mortal Wkly Rep. 2018 Dec;67(48):1333–1338.

Horwitz RI, Feinstein AR, Horwitz SM, Robboy SJ. Necropsy diagnosis of endometrial cancer and detection-bias in case/control studies. Lancet. 1981 Jul;2(8237):66–8.

Khandekar MJ, Cohen P, Spiegelman BM. Molecular mechanisms of cancer development in obesity. Nat Rev Cancer. 2011 Dec;11(12):886–95.

Kok, Inge M. C. M. de, van Rosmalen, Joost, Castle, Philip E., Berkhof, Johannes. The Impact of Different Screening Model Structures on Cervical Cancer Incidence and Mortality Predictions: The Maximum Clinical Incidence Reduction (MCLIR) Methodology. Medical Decision Making. 2020 Jun;40(4):474.

Korhonen MO, Symons JP, Hyde BM, Rowan JP, Wilborn WH. Histologic classification and pathologic findings for endometrial biopsy specimens obtained from 2964 perimenopausal and postmenopausal women undergoing screening for continuous hormones as replacement therapy (CHART 2 Study). Am J Obstet Gynecol. 1997 Feb;176(2):377–80.

Lacey JV, Jr., Mutter GL, Nucci MR, et al. Risk of subsequent endometrial carcinoma associated with endometrial intraepithelial neoplasia classification of endometrial biopsies. Cancer. 2008 Oct;113(8):2073–81.

Lu KH, Broaddus RR. Endometrial Cancer. N Engl J Med. 2020 Nov;383(21):2053–2064.

Reed SD, Newton KM, Clinton WL, et al. Incidence of endometrial hyperplasia. Am J Obstet Gynecol. 2009 Jun;200(6):678 e1–6.

Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024 Jan;74(1):12–49.

Wingo PA, Huezo CM, Rubin GL, Ory HW, Peterson HB. The mortality risk associated with hysterectomy. Am J Obstet Gynecol. 1985 Aug;152(7 Pt 1):803–8.

Zhao J, Hu Y, Zhao Y, Chen D, Fang T, Ding M. Risk factors of endometrial cancer in patients with endometrial hyperplasia: implication for clinical treatments. BMC Womens Health. 2021 Aug;21(1):312.

Duke University
Version: 1.0.00
Released: 2025-09-30

Duke University Uterine Cancer Model (DU-CAM): Model Profile

Duke University

Contact

William Hazelton (hazelton@fredhutch.org)

Funding

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

Suggested Citation

Hazelton WD, Meyers ER, Havrilesky LJ. Duke University Uterine Cancer Model (DU-CAM): 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-du-cam-model-profile-1.0.00-2025-09-30.pdf

Version Table

Version	Date	Notes
1.0.00	2025-09-30	Initial release

Duke University
Readers Guide

DukeHealth

Reader's Guide

Core Profile Documentation

These topics will provide an overview of the model without the burden of 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.

Key References
A list of references used in the development of the model.

DukeHealth

Model Purpose

Summary

This document describes the primary purpose of the model.

Purpose

The overarching purpose of the DU-CAM model is to create a uterine cancer natural history model that allows examination of current and future trends in incidence and mortality and allows us to estimate the impact of various population-level strategies to reduce mortality and racial disparities in outcomes for uterine cancer.

DukeHealth

Model Overview

Summary

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

Purpose

DU-CAM is a multistage clonal expansion (MSCE) model.
The primary purpose of the DU-CAM model is to simulate the natural history of uterine cancer while accounting for age, period, cohort, race, reproductive history, obesity, and prior hysterectomy.

Background

The incidence and mortality of uterine cancer are increasing. Uterine cancer is one of the few tumor types in which both the incidence and death rates are rising.

A number of factors drive the increasing incidence of uterine cancer:

Obesity: At the population level, the rate of obesity, the strongest risk factor for uterine cancer, is increasing substantially. The obesity rate in the U.S. is rising rapidly and driving the increasing incidence of uterine cancer. At the population level, obesity is the most important risk factor for uterine cancer. There is a dose response relationship between body weight and uterine cancer, with a 50% increase in the risk of uterine cancer for every increase of five units in body mass index (BMI).

Reproductive factors: The hysterectomy rate has declined substantially, leaving a greater number of women at risk for uterine cancer as they age. Reproductive factors, mainly acting through alterations in unopposed estrogens which drive carcinogenesis, have a strong influence on uterine cancer risk. Higher parity reduces uterine cancer risk; a pooled analysis estimated a 27% lower risk in parous compared to nulliparous women, with a significant inverse association with increasing number of live births. Age at menarche is inversely associated with risk, with a meta-analysis of eight prospective studies demonstrating a 4% decrease in risk with every 2-year delay in menarche start. Similarly, older age at menopause increases the risk of uterine cancer due to increased number of ovulatory cycles.

Racial disparities: There is a profound racial disparity in outcomes for black and white women with uterine cancer and the gap is widening. While the incidence of uterine cancer is similar for black and white women (27 cases per 100,000 women), black women are much more likely to die from uterine cancer than white women. Numerous factors contribute to the observed racial disparities in poor outcomes for black women, including an increased prevalence of high-risk tumor histologies, more advanced stage tumors at diagnosis, molecular differences in tumors, disparities in treatment quality, decreased responsiveness to treatment, and socioeconomic factors.

Given the well-documented influence of reproductive risk factors and obesity trends on uterine cancer risk, they are modeled independently in DU-CAM using an obesity and reproductive history generator.

Model Description

The DU-CAM uterine cancer model consists of a biologically based MSCE natural history model of the effects of women's reproductive history events and history of body mass index (BMI) on uterine cancer incidence and mortality. Data on women's reproductive histories and BMI status are from the National Health and Nutrition Examination Surveys (NHANES) and National Health Examination Study (NHES) including over 100,000 women in 21 studies spanning years 1960-2020.

In this work, we comprehensively model the impact of current and future trends in obesity, reproductive history, race, age, period, and birth cohort on uterine cancer incidence and mortality. We utilize a MSCE model to fit to uterine cancer incidence by AJCC stage and mortality by type (Endometrioid, Non-Endometrioid, and Sarcomas) in the Surveillance, Epidemiology, and End Results (SEER) cancer registries, including 18 registries spanning years 2000-2018.

Assumption Overview

Assumption Overview

This section describes underlying assumptions of the DU-CAM model and their implications.

Summary

The primary assumption underlying the DU-CAM model is that we can utilize women's BMI and reproductive histories from NHES and NHANES to model trends in age-specific uterine cancer incidence, survival, and mortality.

Background

Numerous studies suggest that changing uterine cancer incidence trends are associated with historical changes in the distributions of body mass index (BMI) and women's reproductive histories (RH).[1-4]

To study this in more detail, we utilize nationally representative data from the National Health and Nutrition Examination Surveys (NHANES) and National Health Examination (NHES) spanning 1959 to 2020 along with uterine cancer incidence and mortality data from the Surveillance, Epidemiology, and End Results (SEER-18) that includes 18 cancer registries [5]. The aim is to gain a better etiological understanding of the potential roles of BMI and reproductive histories on US trends for age-specific uterine cancer incidence and mortality. [6]

The DU-CAM natural history model relates individual reproductive and BMI histories to cellular processes that contribute to the onset and progression of uterine cancer. We model cancer development as a multistage process of cell mutations and premalignant and malignant clonal expansions. We use a two-stage clonal expansion (TSCE) model to represent pre-malignant growth up to occurrent of the first malignant cell, followed by a lag time estimate of the mean time to cancer incidence. A single-stage clonal expansion (SSCE) model represents malignant clonal growth beginning with the first malignant cell. The SSCE model combines malignant clonal growth with possible metastatic transition while calibrating by stage at incidence, while matching to the lag time estimated from the TSCE model. [7-9]

Separate models are developed by race (non-Hispanic Blacks and non-Hispanic Whites) and by uterine cancer histology (Endometriod, non-Endometriod, and Sarcomas). Age-, period-, and cohort-specific trends for each histology are fit to SEER-18 data though likelihood-based calibration of BMI and RH dose-response effects on premalignant and malignant growth determined by the TSCE and SSCE parameters. The cohort and period trends reflect changes in the distributions of women's BMI and reproductive histories across successive NHES and NHANES surveys. The DU-CAM model does not contain any additional cohort or period trends.

Assumption Listing

We assume that we can utilize women's BMI and reproductive histories from NHES and NHANES to model age-specific uterine cancer incidence, survival, and mortality data from SEER. *This is necessary because NHANES does not include enough subjects to accurately estimate cancer outcomes, and SEER does not include reproductive or BMI histories.*

We assume that women's BMI histories can be simulated based on BMI measurements done by NHANES medical professionals at a single time point by using longitudinal data from NHANES to simulate reasonable BMI lifetime trajectories that pass through the measured value for each woman at their measurement age and date.

The TSCE natural history model includes two stochastic rate-limiting mutation events. Stochastic cell division and death rates define the clonal expansion of initiated cells, and a lag time represents the mean time from first

occurrence of a malignant cell to the time of uterine cancer incidence.

The SSCE model includes stochastic rates for malignant cell division and death and possible metastatic transition. The malignant stage undergoes clonal expansion from the time of first malignant cell to detection. Stochastic size thresholds and metastatic status define the fractions of individuals diagnosed in AJCC stages I - IV.

The background rates for initiation and malignant conversion are set equal to each other. This assures mathematical identifiability of the MSCE model parameters. Growth of the population of premalignant cells (promotion) is modeled stochastically through a cell birth and death process. This process cannot be observed directly but is consistently estimated between different cohorts as the most important mechanism that regulates cancer progression and the shape of the age-specific incidence curve.

References

1. Onstad MA. Addressing the Role of Obesity in Endometrial Cancer Risk, Prevention, and Treatment. J Clin Oncol. 2016;34(35):4225–4230.
2. Wu QJ. Parity and endometrial cancer risk: a meta-analysis of epidemiological studies. Sci Rep. 2015;5:14243.
3. Setiawan VW. Age at last birth in relation to risk of endometrial cancer: pooled analysis in the epidemiology of endometrial cancer consortium. Am J Epideiol. 2012;176(4):269–78.
4. Cote ML. Risk factors for endometrial cancer in black and white women: a pooled analysis from the Epidemiology of Endometrial Cancer Consortium (E2C2). Cancer Causes Control. 2015;26(2):287–296.
5. National Cancer Institute, DCCPS, Surveillance Research Program. SEER*Stat Database: Incidence - SEER Research Data, 18 Registries, Nov 2020 Sub (2020-2018). Released April 2024. 2025;
6. National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Centers for Disease Control and Prevention (CDC). 2025;
7. Moolgavkar SH. Mutation and cancer: a model for human carcinogenesis.. J Natl Cancer Inst. 1981;66(6):1037–52.
8. Moolgavkar SH. Two-MoolgavkarTwoEventModel1990. Risk Analysis. 1990;10(2):323–41.
9. Hazelton WD. Longitudinal multistage model for lung cancer incidence, mortality, and CT detected indolent and aggressive cancers. Mathematical Biosciences. 2012;240(1):20–34.

Parameter Overview

Summary

This section describes the primary data sources and DU-CAM model parameters that were calibrated using maximum likelihood estimation (MLE) methods.

Background

The DU-CAM model is based on the Two-Stage Clonal Expansion (TSCE) model of cancer development [1-3]. The TSCE model parameters represent stochastic rates for inition of normal cells to become premalignant, and for premalignant cells to undergo clonal expansion through cell division and death and further mutation that causes malignant transformation, followed by a lag time to cancer detection. The DU-CAM model allows these cellular rates to depend through dose-response relationships on women's history of BMI and reproductive history (RH) events (ages at menarche, first and last birth, number of births, age at menopause, and age at hysterectomy, if that occurs). Data from NHANES was used to sample women's complete BMI and RH events by race, age, and birth cohort whilel calibrating the DU-CAM TSCE model to SEER incidence data by race and uterine cancer histology. By using women's complete BMI and reproductive histories we maintain the correlations between these events that together may influence the risk for uterine cancer.

Parameter Listing Overview

Category	Implementation	Data source	Example
Model Inputs			
BMI at time of NHES, NHANES surveys	BMI (non-pregnant) by age, race, cohort	NHES and NHANES (1959 - 2020)	BMI(age 45) = 43 kg/m^2
Simulated age-specific BMI trajectories	BMI at ages 5, 15, ..., 85 by BMI percentile	Longitudinal & cross-sectional NHANES	BMI(age 5) = 18, BMI(age 15) = 22, ..., BMI(age 85)=34
Hysterectomy data from NHANES	Hysterectomy age by race, cohort	NHANES (1988 - 2020)	Hysterectomy at age 52
Reproductive histories from NHANES			
Age at menarche	Menarche age by race, cohort	NHANES (1988 - 2020)	Menarche at age 14
Age at first birth	First birth age by race, cohort	NHANES (1988 - 2020)	First birth at age 16
Age at last birth	Last birth age by race, cohort	NHANES (1988 - 2020)	Last birth at age 23
Number of live births	Number of live births by race, cohort	NHANES (1988 - 2020)	Number of live births = 3
Age at menopause	Menarche age by race, cohort	NHANES (1988 - 2020)	Menopause at age 46
Weight retention following childbirth(s)	Weight retention at 5, 10 years after last birth	NCHS Natality, CDC Wonder (2005 - 2023)	Increase of 1.5 BMI units by weight retention at age 33
All-cause mortality	Mortality by age, race, cohort	CDC Wonder (1968-2016)	Survival of 90.3% for NHW female at age 60 in 2020
Cancer-specific survival	10-year hazards by histology, stage, age, race	SEER (2000 - 2018)	NHW age 52 Endometrioid cancer AJCC stage 1
Calibration targets	Endometrioid (EM), Non-EM, Sarcomas		
Uterine cancer incidence	Age-specific incidence by histology, stage, race	SEER (2000 - 2018	38.8 per 100,000 at age 70 for NHB EM AJCC-I in 2015
Age at peak of EIN incidence (for EM histology)	Used to adjust lag time for for EIN in TSCE model	Semere, Obstet Gyneol. 2011, 118(1):21-28.	Median 53 years of age for EIN incidence
MLE calibration of TSCE models			
Premalignant TSCE model parameter MLEs:			
P1 background cell division rate	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P1 = 1.639372e+02 per year for NHW EM histology
P2 background premalignant promotion rate	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P2 = 1.642056e-01 per year for NHW EM histology
P3 background first and second mutation rates	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P3 = 1.211614e-07 per year for NHW EM histology
P4 increase in promotion rate after menarche	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P4 = 1.657073e+00 for NHW EM histology
P5 promotion rate during pregnancies	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P5 = 2.200011e+01 for NHW EM histology
P6 promotion rate after menopause	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P6 = 6.830476e-01 for NHW EM histology
P7 BMI promotion dose-response coefficient	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P7 = 4.360609e-01 for NHW EM histology
P8 BMI promotion dose-response power	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P8 = 1.399533e+00 for NHW EM histology
P9 lag-time from 1st malignant cell to incidence	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P9 = 1.224001e+00 year lagtime for NHW EM
Malignant TSCE model parameter MLEs:			
P10 background malignant cell division rate	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P10 = 2.190000e+02 per year for NHW EM cancer
P11 background malignant promotion rate	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P11 = 9.3566025e-01 per year for NHW EM cancer
P12 Stochastice size threshold AJCC I to AJCC II	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P12 = 8.663401e-08 threshold AJCC I to II
P13 Metastatic transition threshold	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P13 = 1.718970e-07 metastatic threshold
P14 Stochastice size threshold AJCC II to AJCC III	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P14 = 2.350917e-07 threshold AJCC II to III
P15 Stochastice size threshold AJCC III to AJCC IV	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P15 = 2.5853101e-07 threshold AJCC III to IV
P16 Estimated lag-time to EIN for EM histology	Maximim likelihood estimate (MLE)	SEER incidence, NHANES histories	MLE for P16 = 1.110000e+00 year lagtime for NHW EIN

Abbreviations: MLE: maximum likelihood estimate, NHW: non-Hispanic Whites, NHB: non-Hispanic Blacks, BMI: body mass index, RH: reproductive histories, NHES: National Health Examination Studies [4], NHANES: National Health And Nutrition Examination Studies [5], SEER: Surveillance, Epidemiology, and End Results cancer registries [6], NCHS: National Center for Health Statistics, CDC: Centers for Disease Control. EM: endometrioid, EIN: Endometrial Intraepithelial Neoplasia [7].

Maximum likelihood estimation (MLE) methods were utilized to calibrate the DU-CAM model by age and calendar year to 2000-2018 SEER incidence data and NHANES BMI and RH data for NHW and NHB race/ethnicity groups and EM, non-EM, and Sarcoma histology groups. We utilized a Poisson likelihod to estimate the dose-response parameter relationships between BMI and RH events to the observed race and histology-specific SEER incidence cases. These methods provided estimates relating uterine cancer incidence trends by age and year to historical changes in BMI and reproductive history trends in the US between 2000 and 2018.

References

1. Moolgavkar SH. Mutation and cancer: a model for human carcinogenesis.. J Natl Cancer Inst. 1981;66(6):1037–52.
2. Moolgavkar SH. Two-MoolgavkarTwoEventModel1990. Risk Analysis. 1990;10(2):323–41.
3. Hazelton WD. Longitudinal multistage model for lung cancer incidence, mortality, and CT detected indolent and aggressive cancers. Mathematical Biosciences. 2012;240(1):20–34.
4. National Center for Health Statistics (NCHS). National Health Examination Survey Data. Centers for Disease Control and Prevention (CDC). 2025;
5. National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Centers for Disease Control and Prevention (CDC). 2025;
6. National Cancer Institute, DCCPS, Surveillance Research Program. SEER*Stat Database: Incidence - SEER Research Data, 18 Registries, Nov 2020 Sub (2020-2018). Released April 2024. 2025;
7. Semere LM, Ko E, Johnson NR, Vitonis AF, Phang LJ, Cramer DW, Mutter GL. Endometrial Intraepithelial Neoplasia Clinical Correlates and Outcomes. Obstet Gynecol. 2011 Jul;118(1):21–28.

Component Overview

Summary

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

Overview

Several components are involved in construction of the DU-CAM uterine cancer model. A Population Component uses individual simulated BMI, reproductive history (RH), and other cause mortality histories to generate a simulated US population. A Natural History Component utilizes the TSCE model, previously calibrated to BMI and RH data, to estimate uterine cancer deaths in the simulated US population based on the TSCE model. A Survival-Mortality Component includes effects of the lag time from first malignant cell to uterine cancer death in the TSCE model, and adjustments for additional age, period, and birth cohort to improve the fit to US uterine cancer mortality.

Component Listing

Population component

The National Health and Examination Survey (NHANES) and National Health Examination Study (NHES) comprise a series of population-based cross-sectional surveys of approximately 100,000 women in 21 studies with individual sampling weights to provide a representative sample of the US population that have been collected from 1960 -2020 and include periodic longitudinal follow up surveys of participants. Pertinent data include medical examinations, BMI, race, ethnicity, birth year, age at menarche, age at menopause, hysterectomy status and age, and use of hormonal medications.

The Surveillance, Epidemiology, and End Results (SEER) Program is an authoritative source of information on cancer incidence and survival in the United States. The SEER Program collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 48% of the U.S. population. SEER 17 (previously 18*) is available for cases diagnosed from 2000 through the current data year and includes expanded races.

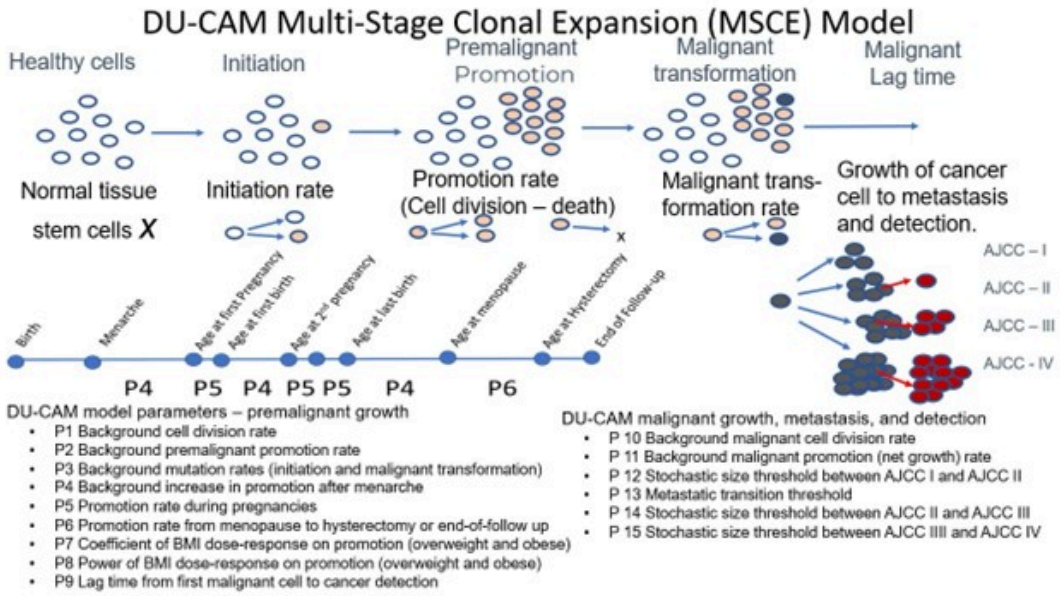
Hysterectomy incidence is currently incorporated into DU-CAM from NHANES (stratified by age, race, and BMI). Subsequent updates to DU-CAM will use data from the National Hospital Discharge Survey (NHDS), National Inpatient Sample (NIS), and National Ambulatory Surgery Sample (NASS), allowing stratification by age, indication for hysterectomy, race, and geographic region.

Reproductive and Obesity History Generator Component

NHANES surveys include a single BMI measurement. In order to generate individual BMI histories, patients are matched based on birth cohort, ages at live births, and age at menarche/menopause and NHANES BMI measurement at study age and date to generate individual BMI trajectories based on longitudinal data that all pass through the measured BMI value.

Natural History Component

DU-CAM is designed as a multistage clonal expansion (MSCE) model, with a two-stage component with lag time initially fit to SEER incidence, followed by a single stage model of malignant growth and metastasis to fit to the SEER AJCC stage I - IV distributions. The models are informed by a reproductive and obesity history “generator” and then calibrated to observed cancer registry incidence, and further modified to generate stage-dependent detection rates.



Modeling endometrial cancer incidence:

Because the reproductive histories and obesity profiles in NHANES are derived from a cross-sectional sampling with individual sampling weights, they are representative of the US population. Endometrial cancer incidence derived from SEER data is also cross-sectional and representative and was mapped/calibrated to obesity and RH profiles based on birth cohort, age, and race/ethnicity.

The nascent model is calibrated to age-specific incidence in SEER. Risk per birth cohort differs by age. Future projections of risk to 2050 are based on projecting reproductive history and BMI trends in NHANES observed between 2000 and 2020 to extend to the future date of 2050.

Race and histologic type were modeled independently.

Of note, DU-CAM models the natural history of uterine cancer using 3 parallel disease cohorts: (1) Endometrioid uterine cancers, (2) High-risk carcinomas: uterine papillary serous carcinomas (UPSC), clear cell carcinomas (CCC), and carcinosarcomas (MMT); and (3) leiomyosarcomas (LMS). Our model includes AJCC stages I – IV.

Survival-Mortality Component

Stage-dependent malignant growth and detection are modeled using a 1-stage clonal expansion model. DU-CAM is then recalibrated to BMI and reproductive histories.

We also use women's histories from NHANES to calibrate to uterine cancer mortality using a separate MSCE model, and are currently working on fitting to 1-, 5-, and 10-year survival by histological subtype age, and birth cohort based on SEER data.

Output Overview

Summary

Definitions and methodologies for the basic model outputs.

Overview

Calibration Data

- NHANES reproductive histories**
    - Age at menarche
    - Age at first birth
    - Age at last birth
    - Number of live births
    - Age at hysterectomy
    - Age at menopause
    - BMI
- SEER-18 uterine cancer incidence
    - Endometrioid
    - Non-Endometrioid
    - Sarcomas
- SEER-18 AJCC stage distributions and survival by histology
    - AJCC - I
    - AJCC - II
    - AJCC – III
    - AJCC – IV
- SEER-18 Uterine cancer mortality – all histologies

Output Listing

Outputs of the DU-CAM model include:

- EM, non-EM, and sarcoma uterine cancer incidence for NHB and NHW women that are calibrated to SEER-18 data between years 2000 and 2018, with projections to 2050.
- Endometrial intraepithelial neoplasia (EIN) that is a precursor for EM cancer incidence for NHB and NH for years 2000 - 2050.
- AJCC stage distributions at incidence for EM, non-EM, and sarcomas by race/ethnicity for years 2000 - 2050.
- Age-specific incidence-based mortality for EM, non-EM, and sarcomas for NHB and NHW for years 2000 - 2050.
- Age-specific attributable risks for BMI and reproductive histories as etiological risk factors in uterine cancer incidence by race/ethnicity and histology.
- Estimates of the age-specific impact on uterine cancer incidence of hysterectomies by race/ethnicity and histology.

Results Overview

Results Overview

This section describes results from fitting the DU-CAM model to uterine cancer incidence and mortality.

Summary

We used maximum likelihood estimation (MLE) methods to optimize the DU-CAM model fit age-specific uterine cancer incidence from SEER-18 across birth cohorts spanning years 2000-2018. Separate models were fit for non-Hispanic Whites (NHW) and non-Hispanic Blacks (NHB) and by histology: endometrioid (EM), non-EM, and sarcomas. The general age incidence trends depend on the TSCE model parameters. Differences by cohort and period arise through dose-response relationships for BMI and reproductive histories (RH) that modify the TSCE model parameters. Women undergoing hysterectomy were modeled as not being subsequently at risk for uterine cancer.

Overview

To identify the best fitting model for each race and histology, we optimized and compared over 20 different dose-response models influencing the TSCE model parameters. Each model was fit to histology-specific uterine cancer incidence from SEER-18. The different models assumed alternative dose-response relationships for the TSCE model parameters based on contemporaneous nationally-representative women's BMI and reproductive histories (RH) from NHANES. The RH events included ages at menarche, pregnancies defined by first and last birth, menopause, and hysterectomy, if it occurred.
Model comparisons were done using the Akaike Information Criterion (AIC) which adjusts the model likelihood scores based on the number of model parameters [1].

After identifying the best fitting TSCE model for each race and histology, we fit a single-stage clonal expansion (SSCE) model to represent the growth, metastatic transformation, and stage distribution at incidence observed in SEER-18, while matching to the estimated lag time estimated earlier from the TSCE model. Histology-specific cancer stage was classified according to the staging manual of the American Joint Committee on Cancer (AJCC), 7th edition, as AJCC stages I-IV.

Uterine cancer incidence-based mortality in the DU-CAM model was calculated for years 2000-2018 based on DU-CAM incidence and monthly histology-specific SEER-18 survival data.

Results List

In comparing over 20 different BMI and RH dose-response models for each race and histology, we identified the best fitting model based on AIC. We found a high degree of consistency for best fitting model by race and histology. For all histologies (EM, non-EM, and sarcoma), the best fitting models included a linear dose-response relationship for BMI influencing malignant conversion, and a piecewise-constant dose-response for RH intervals influencing premalignant promotion. However, for NHB EM, we found an additional significant linear dose-response relationship for BMI affecting initiation.

Figure 1 shows model fits by race for EM cancers. The difference in incidence by year are attributable to the effects of BMI and RH influencing the TSCE model rates, as described above.

Figure 2. shows predictions for endometrial intraepithelial neoplasia (EIN), a precursor to EM cancers. EIN was modeled as a secondary result of fitting the DU-CAM model to EM cancer incidence by adjusting lag time used in the TSCE model to approximate the observed age-specific peak in EIN incidence around ages 50-54.

Figure 3 shows model fits by race for non-EM cancers. Rates for non-EM cancers are increasing more rapidly among NHB than NHW women. Mortality rates tend to be higher for non-EM cancers than for EM cancers.

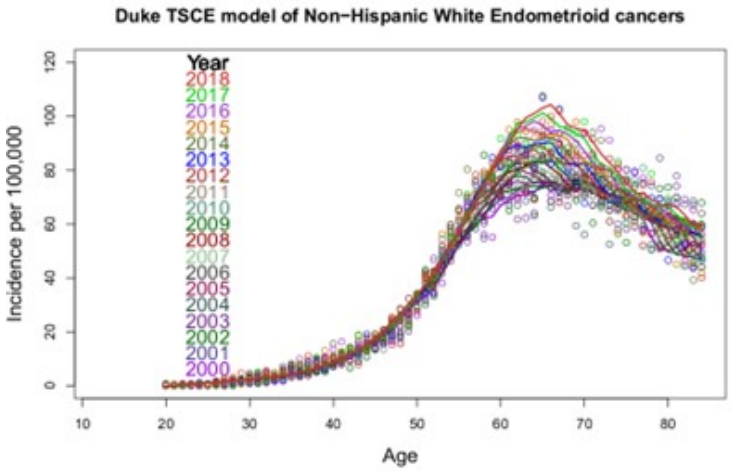
Figure 4 shows model fits by race for sarcomas. Sarcomas occur less frequently than other uterine cancers and tend to have poor survival.

For all histologies, we found that BMI contributes more than RH to age-specific risk at earlier ages, while RH contributes more to uterince cancer risk at older ages.

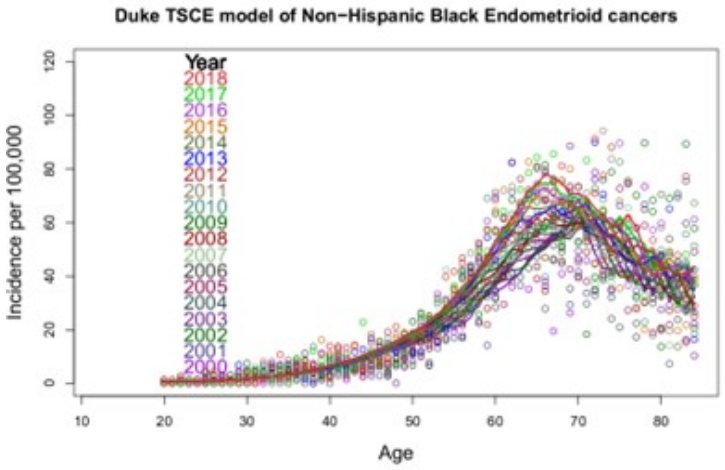
Fit/Figures

1. Calibration of DU-CAM model of endometrioid (EM) cancers to reproductive history and SEER incidence for endometrioid cancers. SEER EM incidence data by year are shown by colored circles and MSCE model fits are shown by corresponding colored lines.

   1a. NH White EM

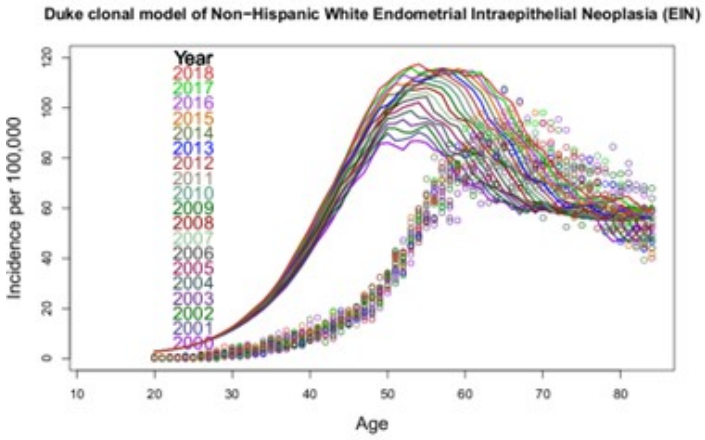


   1b. NH Black EM

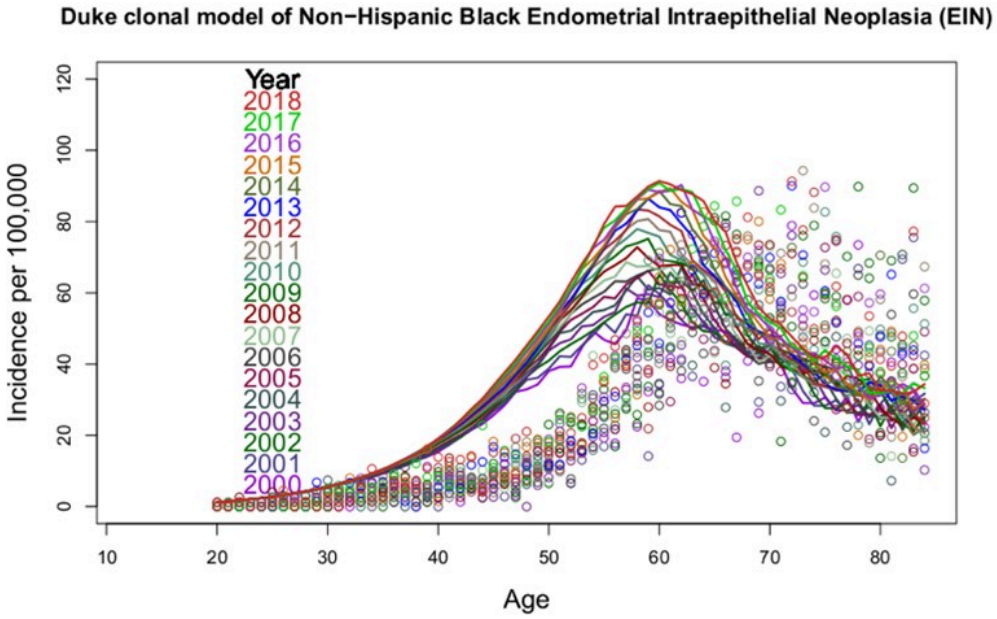


2. EIN prediction.
   SEER incidence data by year are shown by colored circles. Likelihood-based prediction of earlier EIN (a precursor to EM cancers) are shown by corresponding colored lines. EIN is potentially detectable prior to cancer diagnosis (as shown below).

2a. NH White EIN

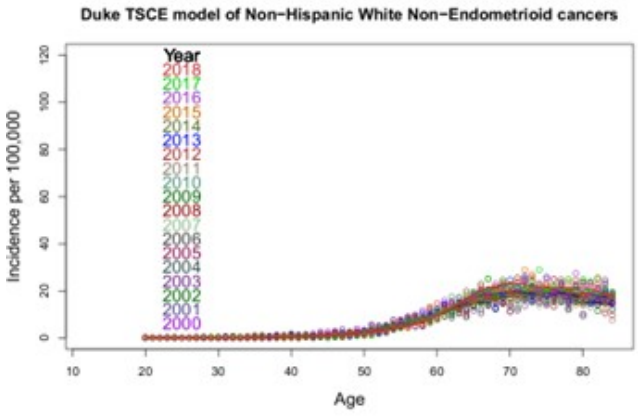


2b. NH Black EIN

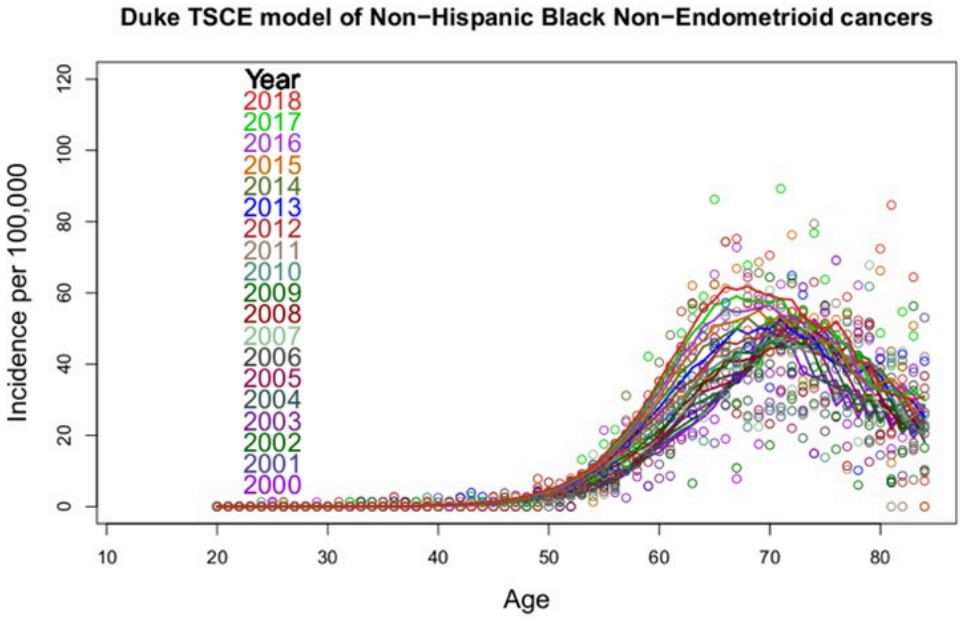


3. Calibration of DU-CAM TSCE of non-endometrioid (non-EM) cancers to reproductive history and SEER incidence. SEER non-EM incidence data by year are shown by colored circles and MSCE model fits are shown by corresponding colored lines.

3a. NH White Non-EM

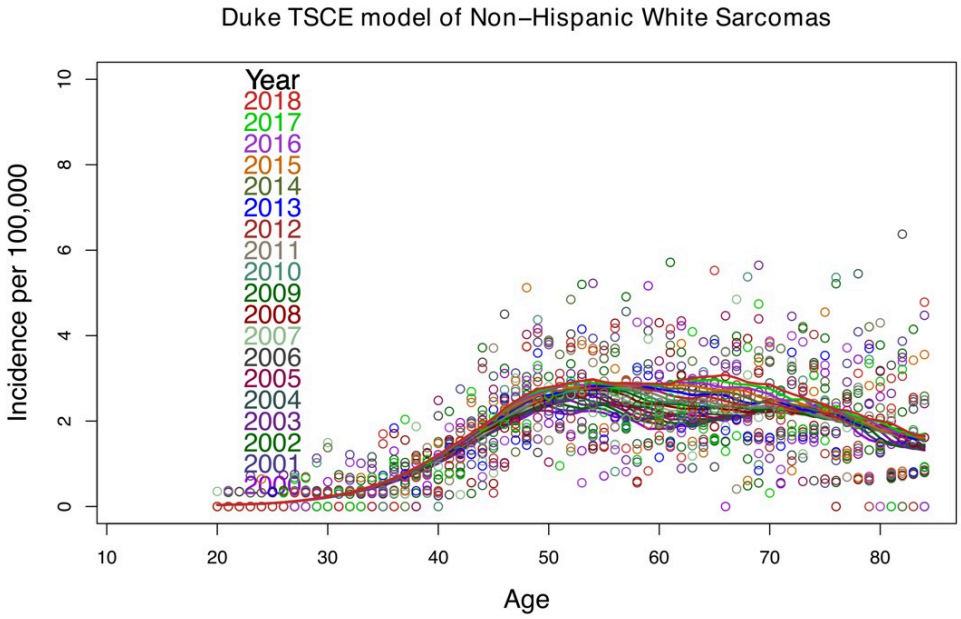


3b. NH Black Non-EM

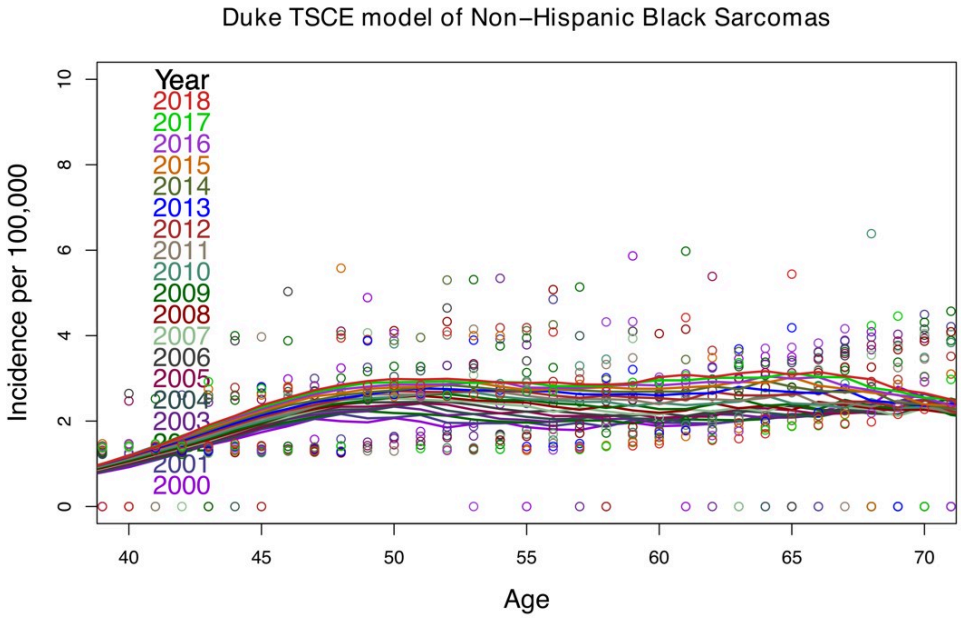


4. Calibration of DU-CAM TSCE of uterine sarcomas to reproductive history and SEER incidence. SEER sarcoma incidence data by year are shown by colored circles and MSCE model fits are shown by corresponding colored lines.

4a. NH White



4b. NH Black



References

1. Akaike HA. A New Look at the Statistical-Model Identification. IEEE Transactions on Automatic Control. 1974;19:716–723.

Duke University
Natural History
Component

DukeHealth

Natural History Component

Summary

Women's history of BMI and reproductive events are recognized as important risk factors in uterine cancer.

Overview

The DU-CAM model is informed by women's natural histories of BMI and reproductive history events as risk factors for uterine cancer.

Detail

We harmonized data on women's individual histories of BMI and reproductive events by race/ethnicity using US representative data from NHANES surveys between 1991 and 2020. These thousands of individual histories were utilized to inform the natural history of uterine cancer cellular progression from normal tissue by initiating mutations that generate premalignant cells. Premalignant cells undergo clonal expansion and malignant transformation to generate clonally expanding malignancies that may progress to cancer incidence and mortality.

By using full individual histories from women we are able to capture the complex correlations between BMI and and reproductive history events, (including ages at menarche, pregnancies associated with first and last birth, menopause, and hystorectomies) that contribute to uterine cancer progression. Changes by birth cohort and calendar year in the joint distributions of BMI and reproductive history events are captured by successive NHANES surveys, allowing modeling of current and future uterine cancer trends by histology and race/ethnicity.

Duke University
Population Component

DukeHealth

Population Component

Summary

We utilize population data on uterine cancer incidence from SEER, and contemporaneous data on women's BMI and reproductive histories from NHES and NHANES.

Overview

A large fraction of uterine cancer incidence in the US is available from the SEER-18 cancer registries, while US representative data on women's BMI and reproductive histories is available from NHES and NHANES.

Detail

To study the relationships between uterine cancer incidence and women's BMI and reproductive histories, we need extensive US population-representative data of both types. SEER data is the best available large data source for uterine cancer incidence, survival, and mortality, but it does not include information on women's BMI or reproductive histories. NHES and NHANES survey data spanning 1959 - 2020 is one of the best available sources of BMI and reproductive histories, but it is not large enough to provide sufficient uterine cancer incidence or mortality. The DU-CAM model links the two data sets to infer mechanistic associations between BMI and reproductive histories with uterine cancer incidence between 2000 and 2018.

Duke University
Reproductive and Obesity History Generator Component

DukeHealth

Reproductive and Obesity History Generator Component

Summary

The DU-CAM model utilizes inputs on women's obesity and reproductive histories to model the etiology of uterine cancer and predict future trends.

Overview

US representative data on reproductive histories and obesity were generated by extracting and harmonizing individual level records from NHES and NHANES survey data spanning 1959 - 2020.

Detail

There are complex correlations between BMI and different components of women's reproductive histories. It would be difficult to model these correlations in separate models of BMI and ages of reproductive history events. Instead we capture these correlations and inform the DU-CAM model by utilizing thousands of individual BMI and reproductive histories from NHANES.

Duke University
Survival-Mortality
Component

Survival-Mortality Component

Summary

The DU-CAM model utilizes incidence-based cancer mortality methodology.

Overview

The DU-CAM model provides a mechanistic model for uterine cancer incidence by race/ethnicity and histology in relation to women's BMI and reproductive histories. We utilize SEER-18 survival data by histology to calculate incidence-based cancer mortality.

Detail

We utilize 10-year monthly survival from SEER-18 by race/ethnicity and histology tp predict histology specific uterine cancer mortality between years 2000-2018, while predicting mortality out to year 2050.

However the DU-CAM incidence calibration began in year 2000, so we do not have complete incidence data to predict cancer mortality for the first ten years between 2000 and 2010. To fill in this interval, we assumed that incidence trends remained fixed between years 1990 and 2000.

Duke University
Key References

DukeHealth

Key References

Akaike HA. A New Look at the Statistical-Model Identification. IEEE Transactions on Automatic Control. 1974;19:716–723.

Cote ML. Risk factors for endometrial cancer in black and white women: a pooled analysis from the Epidemiology of Endometrial Cancer Consortium (E2C2). Cancer Causes Control. 2015;26(2):287–296.

Hazelton WD. Longitudinal multistage model for lung cancer incidence, mortality, and CT detected indolent and aggressive cancers. Mathematical Biosciences. 2012;240(1):20–34.

Moolgavkar SH. Mutation and cancer: a model for human carcinogenesis.. J Natl Cancer Inst. 1981;66(6):1037–52.

Moolgavkar SH. Two-MoolgavkarTwoEventModel1990. Risk Analysis. 1990;10(2):323–41.

National Cancer Institute, DCCPS, Surveillance Research Program. SEER*Stat Database: Incidence - SEER Research Data, 18 Registries, Nov 2020 Sub (2020-2018). Released April 2024. 2025;

National Center for Health Statistics (NCHS). National Health Examination Survey Data. Centers for Disease Control and Prevention (CDC). 2025;

National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Centers for Disease Control and Prevention (CDC). 2025;

Onstad MA. Addressing the Role of Obesity in Endometrial Cancer Risk, Prevention, and Treatment. J Clin Oncol. 2016;34(35):4225–4230.

Semere LM, Ko E, Johnson NR, Vitonis AF, Phang LJ, Cramer DW, Mutter GL. Endometrial Intraepithelial Neoplasia Clinical Correlates and Outcomes. Obstet Gynecol. 2011 Jul;118(1):21–28.

Setiawan VW. Age at last birth in relation to risk of endometrial cancer: pooled analysis in the epidemiology of endometrial cancer consortium. Am J Epideiol. 2012;176(4):269–78.

Wu QJ. Parity and endometrial cancer risk: a meta-analysis of epidemiological studies. Sci Rep. 2015;5:14243.

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

Icahn School of Medicine at **Mount Sinai**

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

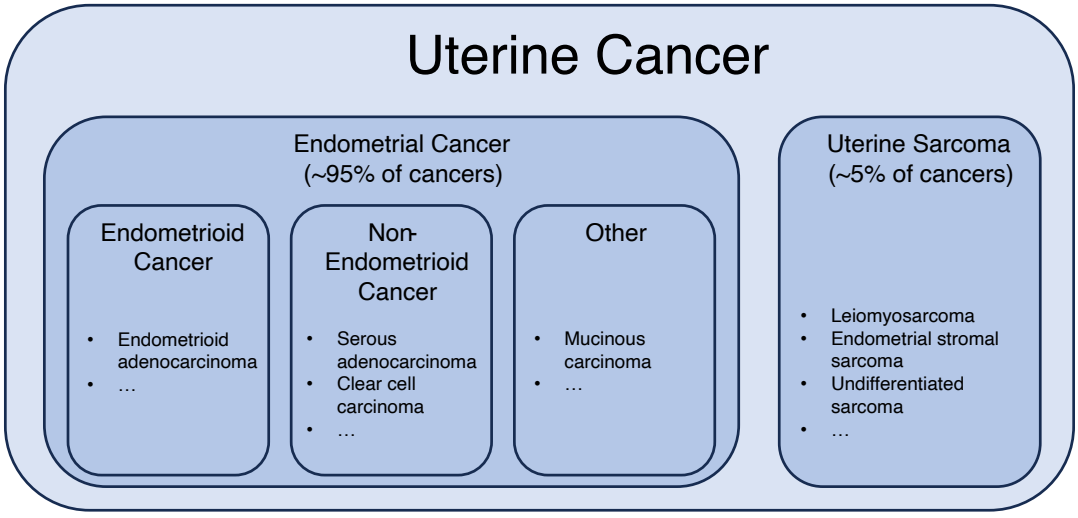
Purpose

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 $5\,kg/m^2$ increase in body mass index (BMI) and uterine cancer is 1.59, with this ratio increasing for BMIs above $28\,kg/m^2$ [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:

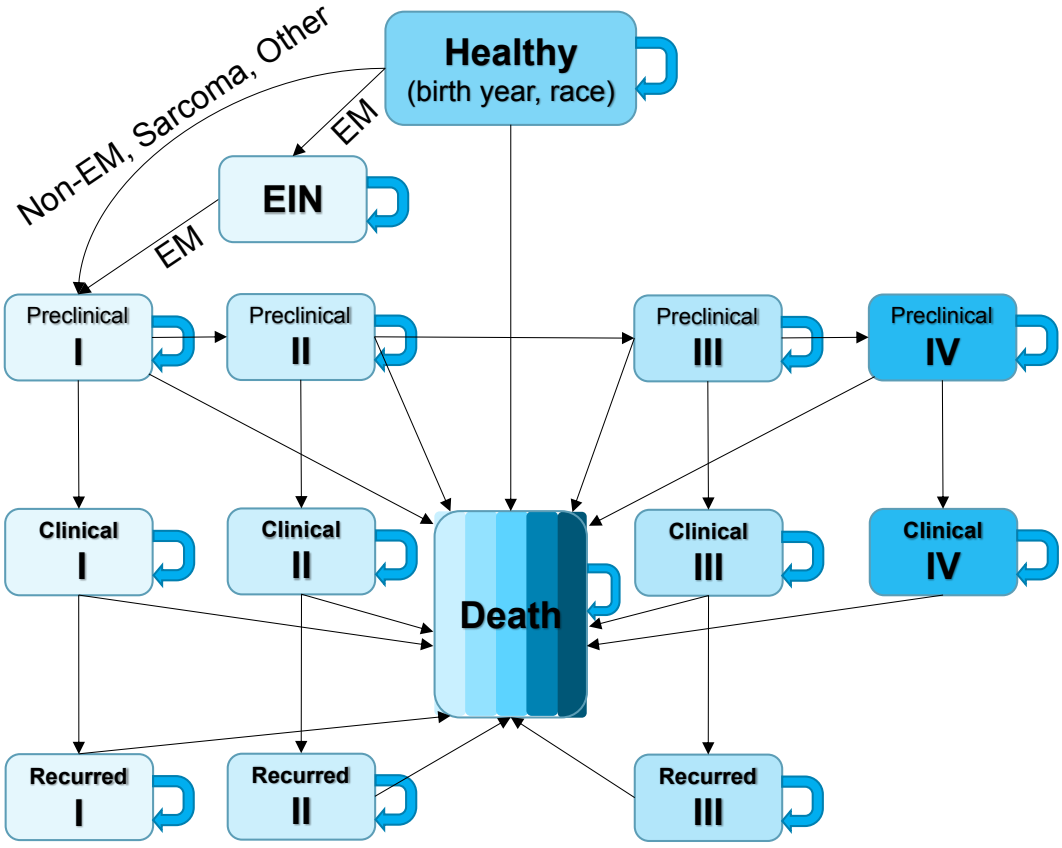
Category	ICD-O-3 code
Endometroid	8050, 8141, 8210, 8260-8263, 8380-8383, 8440, 8480-8481, 8560, 8570
Non-Endometrioid	8255, 8310, 8323, 8441, 8460-8461, 8950-8951, 8980-8981

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.

Mount Sinai
Assumption Overview

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

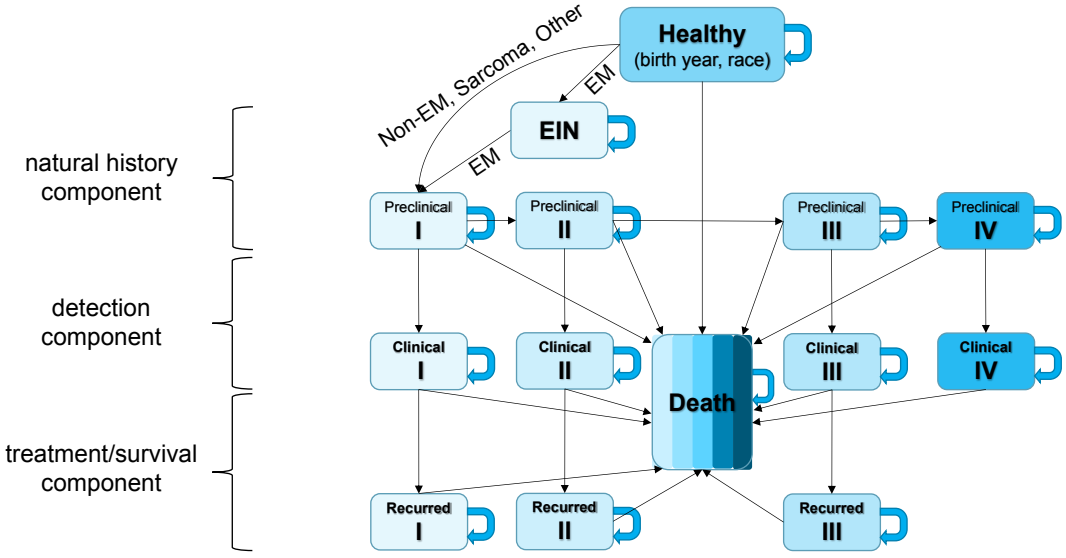
This page contains a description of the different computational elements which make up the Mount Sinai Uterine Cancer Model.

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_{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_{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.

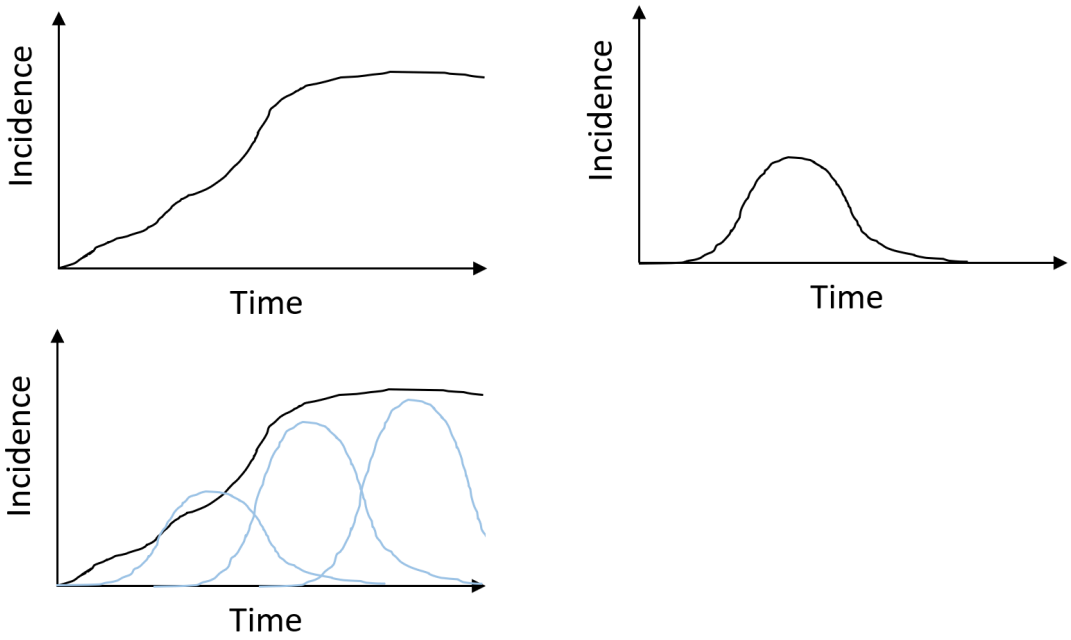
$$f_I(t) = P(\text{Preclinical I} \rightarrow \text{Clinical I} \mid [\text{Healthy} \rightarrow \text{Preclinical I}]_{t=\text{Age}}) \quad \text{Age} + 0\text{yr} \leq t \leq \text{Age} + 30\text{yr}$$

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_{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 λ being the target transition rate and t_{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_{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 α_i being the distribution of subjects within the BMI groups at that time and its sum being 1. The modified transition rates λ_i^{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 λ_i^{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* λ_{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 λ_{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 λ_{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;

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

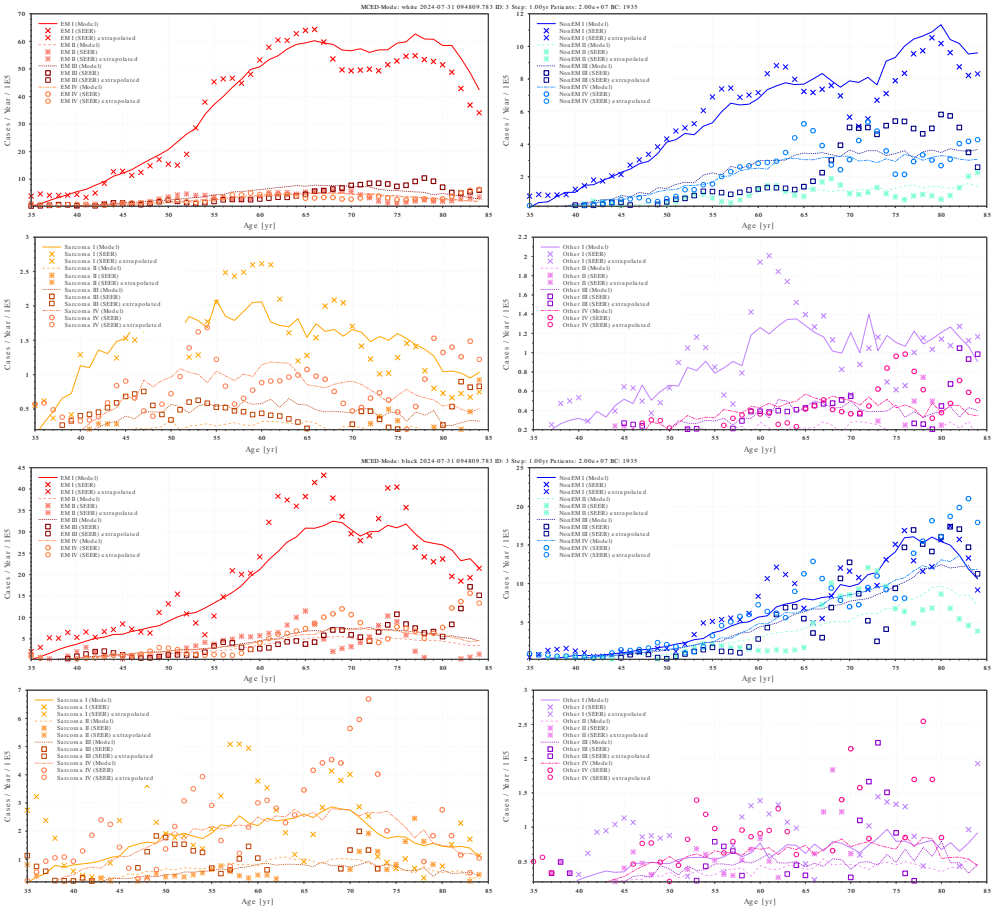
A guide to the results obtained from the model.

Overview

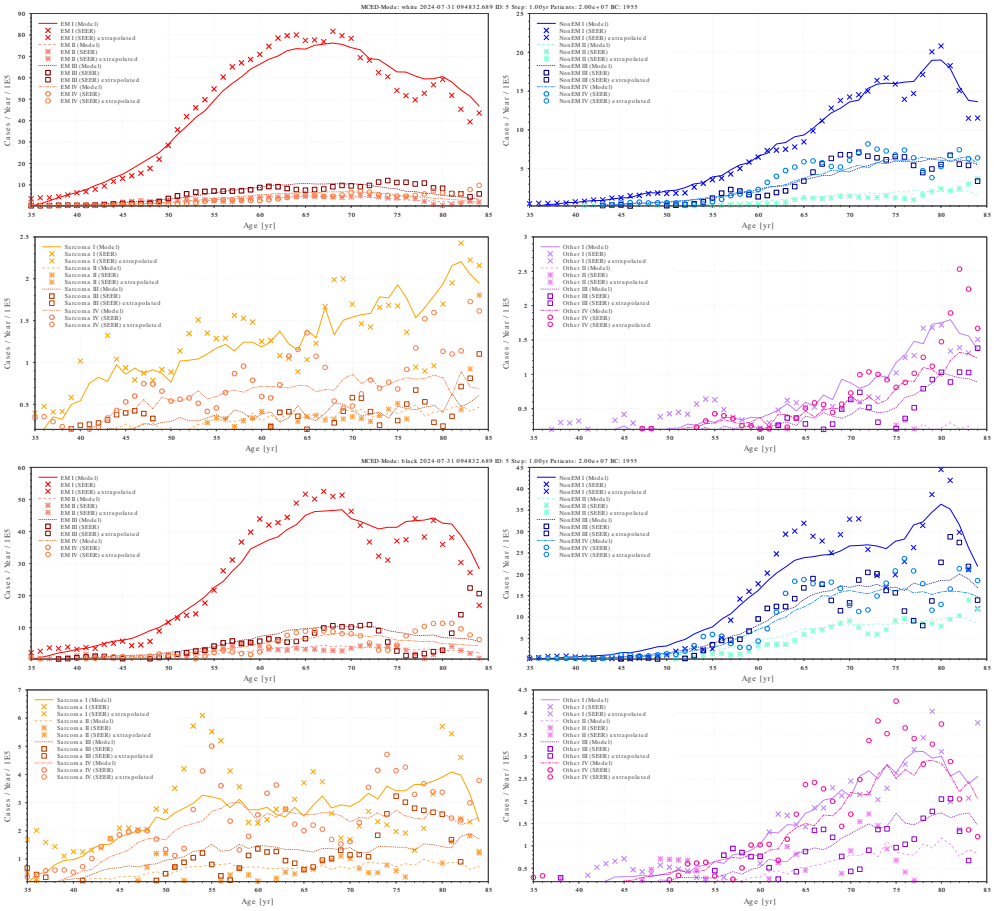
Results List

Calibration results

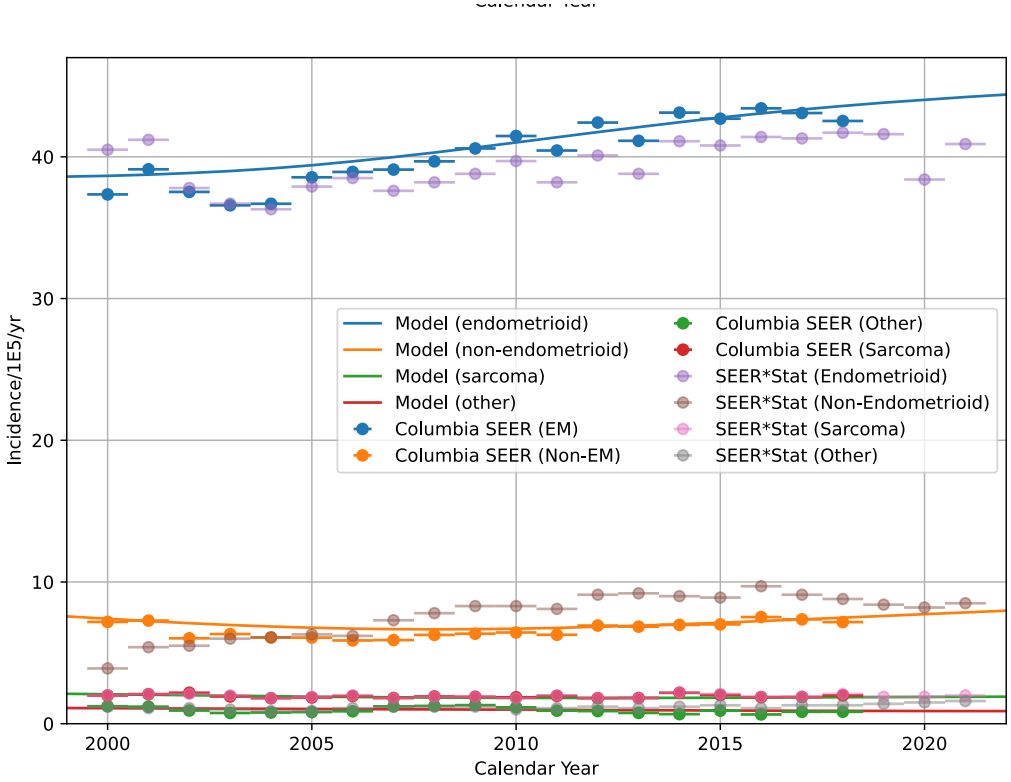
1. Calibration of the MUSIC model by birth cohort 1935, histologies, and race. Model projections are shown by lines, and SEER calibration targets are shown by symbols in matching colors for each stage. Bold symbols represent SEER data, and regular symbols represent projections of SEER data. Data $y(t)$ 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.

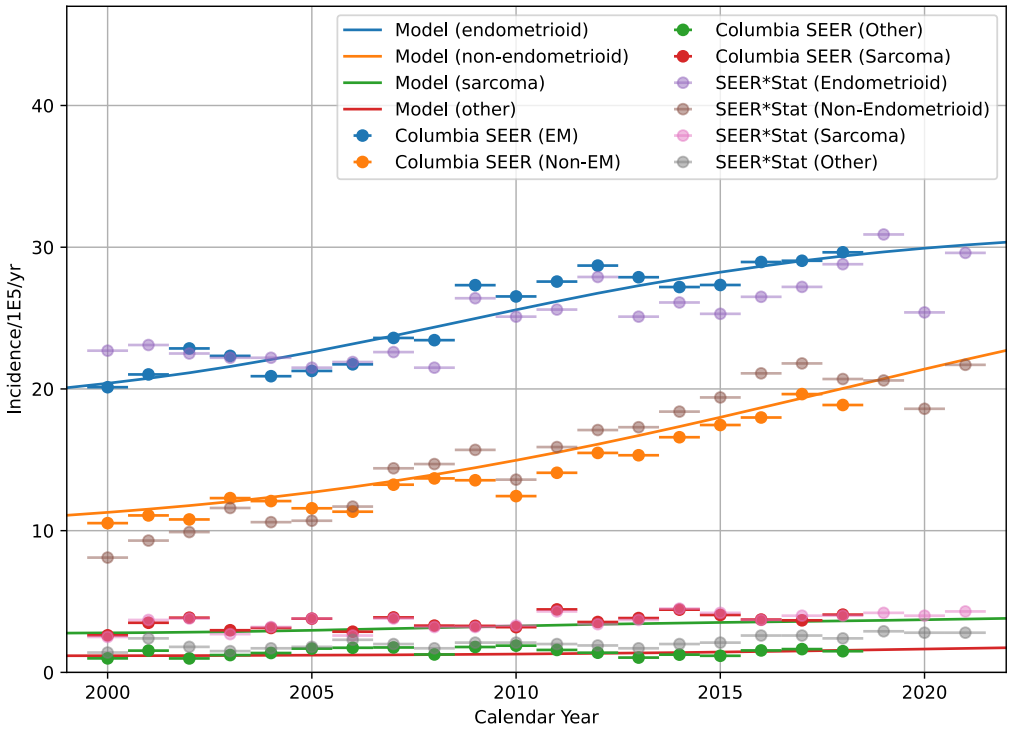


2. Calibration of the MUSIC model by birth cohort 1955, histologies, and race. Model projections are shown by lines, and SEER calibration targets are shown by symbols in matching colors for each stage. Bold symbols represent SEER data, regular symbols represent projections of SEER data. Data $y(t)$ have 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.

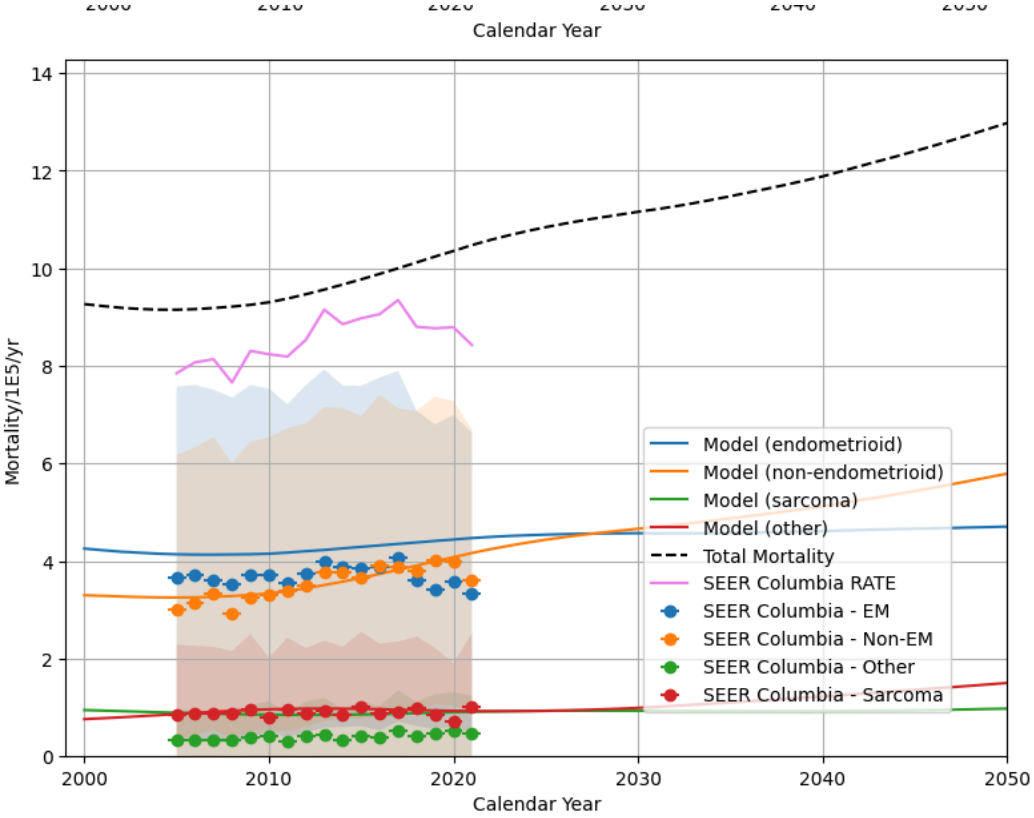


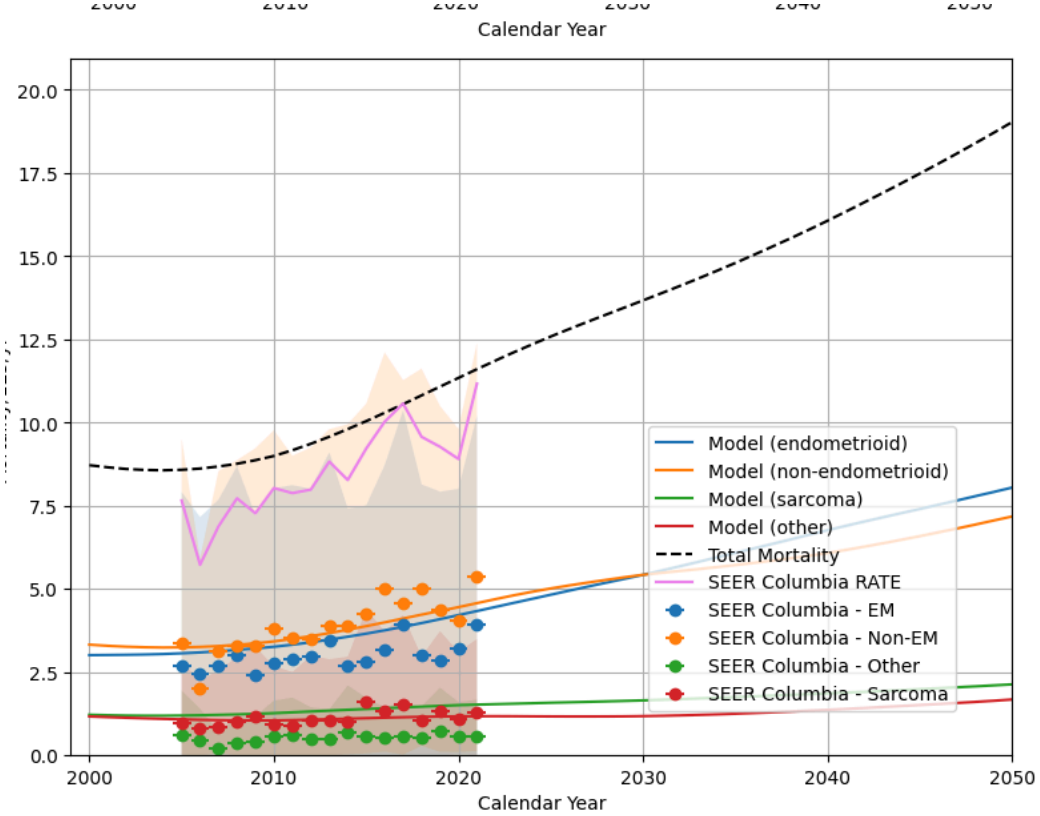
3. Calibration of the age-adjusted uterine cancer incidences by histology for NH White women (top) and NH Black women (bottom). The calibration targets are labeled 'Columbia SEER', which do contain multiple imputed data. The 'SEER*Stat' data is from SEER Nov 2023 Sub.





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.





Icahn School of Medicine at **Mount Sinai**

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.