



HHS Public Access

Author manuscript

Lancet Oncol. Author manuscript; available in PMC 2024 December 01.

Published in final edited form as:

Lancet Oncol. 2023 December ; 24(12): 1434–1442. doi:10.1016/S1470-2045(23)00510-7.

Development and Validation of Age-Specific Primary Ovarian Insufficiency Risk Prediction Models in Long-Term Survivors of Childhood Cancer: A Report from the Childhood Cancer Survivor Study and St. Jude Lifetime Cohort

Cindy Im[†],
Zhe Lu[†],
Sogol Mostoufi-Moab,
Angela Delaney,
Lin Yu,
Jessica L. Baedke,
Yutong Han,
Yadav Sapkota,
Yutaka Yasui,
Eric J. Chow,
Rebecca M. Howell,
Smita Bhatia,
Melissa M. Hudson,
Kirsten K. Ness,
Gregory T. Armstrong,
Paul C. Nathan^{*},
Yan Yuan^{*,‡}

Department of Pediatrics, University of Minnesota, Minneapolis, USA (C Im PhD); School of Public Health, University of Alberta, Edmonton, Canada (Z Lu MSc, L Yu MSc, Y Han BMath, Prof. Y Yuan PhD); Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia,

[†]**Corresponding author:** Yan Yuan, PhD, School of Public Health, University of Alberta, Edmonton, AB, T6G 1C9, Canada, yyuan@ualberta.ca.

Author contributions

YYu, PCN, CI, and SM-M developed the concept and provided supervision for the study. YYa, CI, GTA, EJC, KKN, MMH, SB, and YYu designed the study, YYa, ZL, CI, LY, SM-M, RMH, JLB, AD, YS, YH, GTA, KKN, MMH and YYu collected the data, ZL, CI, LY, JLB, YS, YH and YYu prepared and analyzed the data, ZL, YH and CI created the figures/tables, ZL, CI, LY, SM-M, PCN, and YYu interpreted the results, CI, ZL, PCN, YYu drafted the manuscript. Funding for this project was acquired by YYu, PCN, CI, LY, GTA, KKN, and MMH. All authors had access to the CCSS data reported in the study, contributed revisions and approved the final manuscript for publication.

[†]These authors contributed equally.

^{*}These senior investigators contributed equally.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interests

The authors declare no competing interests.

USA (S Mostoufi-Moab MD); Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, USA (A Delaney MD, JL Baedke MSc, Y Sapkota PhD, Prof. Y Yasui PhD, MM Hudson MD, KK Ness PhD, GT Armstrong MD); Division of Endocrinology, Department of Pediatric Medicine, St. Jude Children's Research Hospital, Memphis, USA (A Delaney MD); Public Health Sciences and Clinical Research Divisions, Fred Hutchinson Cancer Center, Seattle, USA (Prof. EJ Chow MD); Department of Radiation Physics, University of Texas at MD Anderson Cancer Center, Houston, USA (Prof. RM Howell PhD); Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, USA (Prof. S Bhatia MD); Department of Oncology, St. Jude Children's Research Hospital, Memphis, USA (MM Hudson MD, GT Armstrong MD); Division of Hematology/Oncology, The Hospital for Sick Children, Toronto, Canada (Prof. PC Nathan MD, MSc)

SUMMARY

Background: Female survivors of childhood cancer are at risk for primary ovarian insufficiency (POI), defined as cessation of gonadal function before age 40 years. We developed and validated models to predict age-specific POI risks among survivors.

Methods: To develop models to predict age-specific POI risk between ages 21–40 years, we used data from 7,891 female 5-year survivors in the Childhood Cancer Survivor Study (CCSS) with survey-ascertained ovarian status (922 with POI). Models utilizing statistical-/machine-learning-based algorithms considering cancer treatments were evaluated. Cross-validated prediction performance metrics (e.g., area under the receiver operating characteristic curve [AUROC]) were compared to select best-performing models. For external validation, we used data from 1,349 5-year survivors in the St. Jude Lifetime Cohort (SJLIFE) with clinically ascertained ovarian status (101 with POI). We evaluated a SJLIFE-based POI polygenic risk score among 1,985 CCSS survivors with genotype data.

Findings: Median follow-up from cancer diagnosis was 24 years (interquartile range [IQR]: 18–30) in CCSS and 15 years (IQR: 10–23) in SJLIFE. Between ages 21 to 40 years, POI prevalence increased from 7.9% to 18.6% in CCSS and 7.3% to 14.9% in SJLIFE. Age-specific logistic regression models considering ovarian radiation dosimetry or prescribed pelvic/abdominal radiation dose, along with individual chemotherapy predictors, performed well in CCSS. In the SJLIFE validation, the prescribed radiation dose model showed excellent performance (AUROC: 0.88–0.95), as did a simpler model that considered any exposures to pelvic/abdominal radiotherapy or alkylators (AUROC: 0.82–0.90). Addition of the polygenic risk predictor significantly improved the average positive predictive value (from 0.76 to 0.87, $P=0.029$) among CCSS survivors treated with ovarian radiation and chemotherapy.

Interpretation: POI risk prediction models utilizing treatment information showed robust prediction performance in adult-aged childhood cancer survivors. An associated online risk calculator is available to inform fertility preservation counseling.

Funding: Canadian Institutes of Health Research, US National Cancer Institute.

INTRODUCTION

Female survivors of childhood cancer are at risk of primary ovarian insufficiency (POI), a clinical syndrome defined by cessation of ovarian function before age 40 years.^{1–3} Ovarian tissue is vulnerable to gonadotoxic cancer treatments such as pelvic radiotherapy and alkylating agents, which can accelerate the age-related decline of the ovarian follicle pool.^{1,4} POI that occurs within five years of cancer diagnosis is defined as acute ovarian failure (AOF). An estimated 6% of survivors develop AOF, while an additional 9% develop POI years after completion of therapy.⁵ In the general population, the prevalence of non-surgical POI is considerably lower at 1–2%.^{3,6} Since POI limits fertility, several multidisciplinary groups have published recommendations surrounding the counseling of pediatric and young adult cancer patients and their families about risks for developing POI and options for fertility preservation.^{7,8}

We previously developed and validated an AOF risk calculator for female childhood cancer survivors.⁹ However, this tool only estimates POI risk within five years of cancer diagnosis. Currently, there are no tools that comprehensively consider cancer treatment regimens to predict age-specific patient risks for developing POI, especially during young adulthood when many survivors contemplate family planning. Available fertility preservation technologies such as ovarian tissue cryopreservation and oocyte harvesting are invasive, expensive, and there is limited data on safety and long-term success rates, especially in pediatric populations.^{10,11} The aim of this study was to develop and validate POI risk prediction models that provide patient-specific risks across an extended age spectrum, informing discussions surrounding the choice and timing of therapies to preserve fertility.

METHODS

Study design and participants

For model development, we used data from the Childhood Cancer Survivor Study (CCSS), an ongoing multi-institutional North American cohort study of 5-year survivors of childhood cancer treated between 1970–1999. Details regarding study methodology and recruitment have been documented previously.¹² Female survivors aged ≥18 years at their latest follow-up with self-reported menstrual history information and free of subsequent malignant neoplasms (SMNs) within 5 years of diagnosis were included. In CCSS, follow-up concluded on 11/25/2016.

For independent model validation, we used data from the ongoing St. Jude Lifetime Cohort (SJLIFE) study. Female 5-year survivors treated for pediatric cancer between 1962–2012 at St. Jude Children’s Research Hospital who completed at least one longitudinal multiple-day study visit with comprehensive clinical/laboratory health assessments and were not enrolled in CCSS were included. The SJLIFE study design and methodology have been described previously.¹³ In SJLIFE, follow-up concluded on 8/27/2021.

In this study, an established definition for POI in the childhood cancer survivor literature^{1,2,14} was used (details provided in Appendix). Ovarian status was classified as: (1)

POI, including survivors who never experienced menarche, had non-surgical self-reported permanent cessation of menses before age 40 (i.e., amenorrhea ≥ 12 months), or persistent amenorrhea and hormone measurements consistent with menopause before age 40; (2) surgical premature menopause (SPM), including survivors with hysterectomies or bilateral oophorectomy; and (3) normal, including survivors without POI or SPM before age 40 or at last survey.

In CCSS, baseline and longitudinal follow-up questionnaires querying age at menarche and last menstrual period, current menstrual status, cause of menopause if currently menopausal, pregnancy/childbirth, and use of hormonal contraception were used to ascertain POI and rule out unrelated causes of absent menses consistent with previous analyses in survivors^{1,2,14,15}. Five-year survivors without SPM aged ≥ 18 years at the time of questionnaire were assigned POI status if they: (a) never menstruated; (b) experienced their last menses within 5 years of cancer diagnosis; or (c) menstruated >5 years after cancer diagnosis but experienced their last menses before age 40. For POI assignment, last menses had to be reported >12 months before the questionnaire date. Ambiguous cases were reviewed by an endocrinologist (SM-M). Self-reported age at POI/SPM was used. For survivors who never menstruated, POI age was assigned at 16 years. SPM was treated as a competing risk. Survivors without events were censored at the last date of follow-up or the earliest SMN diagnosis date.

In SJLIFE, ovarian status was clinically ascertained using hormone measurements (menopause defined by follicle stimulating hormone >30 mIU/mL and estradiol <17 pg/mL) and medical chart/questionnaire review by an endocrinologist (AD). Age at POI clinical diagnosis was used. Survivors with unknown ovarian status, who experienced SPM or a SMN within 5 years of primary cancer diagnosis, or had missing treatment information were excluded. Gonadotropin deficiency was additionally treated as a censoring event.

Human subjects research approval was granted by the institutional review boards of all participating institutions prior to recruitment. All participants provided written/online/verbal informed consent.

Predictors

For model development, we considered host factors and predictors measured within five years of cancer diagnosis including self-reported race/ethnicity, primary cancer diagnosis and diagnosis age, hematopoietic stem-cell transplant (HSCT), and doses of abdominal/pelvic, ovarian radiation, total body irradiation (TBI), and 20 types of chemotherapy agents (including specific alkylators) among female survivors (medical record-abstracted sex) (Appendix p 15). Separate sensitivity analyses considered alkylators quantified as cyclophosphamide equivalent dose [CED]¹⁶ (Appendix). Interactions between age at cancer diagnosis and radiation dose variables and HSCT were assessed *a priori*. All primary cancer-related predictors were abstracted from medical records. Cumulative radiation therapy doses received within 5 years of primary cancer diagnosis were evaluated.⁵ Maximum target radiation doses to pelvic and abdominal regions were obtained by summing the prescribed (i.e., cumulative delivered) dose based on all overlapping fields to the respective body regions/organs.¹⁷ The average radiation doses delivered to right and left ovaries were estimated separately and the lesser of these two values was used in modeling (i.e., minimum

ovarian radiation dose [MORD]). Chemotherapy doses given within 5 years of diagnosis were used; however, if uncommon (i.e., <2% exposed), these exposures were dichotomized for some algorithms (see Appendix p 15).

We evaluated models that included a polygenic risk score (PRS) from a previous premature menopause genome-wide association study (GWAS) among SJLIFE survivors of predominantly European ancestry¹⁸ in European ancestry CCSS survivors with genotype data (Appendix). Methods used to generate genotype data in CCSS have been described in detail elsewhere.¹⁹ Six independent variants with suggestive associations ($P < 1 \times 10^{-5}$) with treatment-related POI risk¹⁸ were included in the PRS (shown in Appendix p 16).

Statistical analysis

Due to the variability in clinical predictor data availability, we considered three separate models (Appendix p 15): (1) an ovarian radiation dose model (referred to as the ovarian model hereafter), where dosimetry-based MORD was used, along with individual chemotherapies; (2) a prescribed radiation dose model which considered abdominal/pelvic radiation target doses (referred to as the prescribed model hereafter), along with individual chemotherapies; and (3) a simple model, in which any exposures to abdominal/pelvic radiation, TBI, and alkylating agents were considered. Race/ethnicity, cancer diagnosis and diagnosis age, and HSCT were included in all models.

Statistical analysis details are given in the Appendix, along with an analytic overview (Appendix p 8). Before model development, we used an established multiple imputation method based on the predictive mean matching approach²⁰ that assumes data is missing at random to impute missing predictor values in CCSS. Four candidate prediction algorithms were evaluated: (1) Cox proportional hazards (PH) regression retaining all variables (Appendix p 15); (2) age-specific logistic regression retaining all variables (Appendix p 15); (3) penalized age-specific logistic regression, with elastic-net penalty method-based²¹ variable selection; and (4) age-specific XGBoost²², a machine learning algorithm that automatically models non-linear effects and complex interactions for continuous predictors. Age-specific algorithms were developed to predict POI risk by ages 21 to 40 years in CCSS, indirectly modeling age-related ovarian decline. Age (in years) as the time scale was used in the Cox PH model with the time origin being birth. In order to use binary classification methods and account for censoring and competing risk while minimizing bias due to loss to follow-up, inverse probability censoring weights (IPCW) were used.²³

Prediction performance was evaluated internally in CCSS and externally in SJLIFE. To avoid overly optimistic estimates of performance, metrics from 5-fold cross-validation or nested cross-validation (for algorithms requiring hyperparameter tuning) in CCSS were assessed. Performance metrics included: scaled Brier score (SBrS) for overall model performance;²⁴ area under the receiver operating characteristic curve (AUROC) for model discrimination, i.e., average sensitivity across varying thresholds;²⁵ and the area under the precision-recall curve (AUPRC), i.e., average positive predictive value across varying thresholds.²³ These metrics have a maximum value of one, with larger values indicating better performance (relative to metric-specific minimum values: SBrS=0; AUROC=0.5; AUPRC=event prevalence). A global calibration statistic, the Spiegelhalter-z statistic,²⁶

was calculated, where absolute values closer to zero indicate better calibration. Additional calibration curve-type plots comparing observed and model-predicted risks of POI in SJLIFE were visually inspected for all age cut-offs (see Appendix). In CCSS, age-specific metrics were calculated based on model-predicted risks averaged over imputed datasets. In SJLIFE, complete case age-specific metrics were calculated. To investigate the potential impact of differences in ovarian status ascertainment in CCSS and SJLIFE on prediction performance, a simulation study was performed considering varying degrees of binary outcome misclassification in the training data on risk prediction performance in a validation dataset without such misclassification, assuming the underlying biological mechanism for POI was the same (see Appendix).

Based on clinician input, POI predicted risks from the ovarian/prescribed models were categorized as follows⁹: low (5%), medium low (>5% to <20%), medium high (20% to <50%) and high (50%). For the simple model, POI predicted risks were classified as either low (5%) or high (>5%). Observed POI prevalence estimates accounting for IPCW and stratified by age and model-predicted risk groups were computed in CCSS.

All statistical analyses were done with R v4.0.2.

Role of the funding source

Funding sources had no role in study design, data collection, data analysis, data interpretation, or writing of this report.

RESULTS

For POI risk prediction model development, 7,891 female CCSS survivors were included, while 1,349 female SJLIFE survivors were included in the validation study (Figure 1). CCSS survivors were older at cancer diagnosis than SJLIFE survivors (median=7.5 versus 5.6 years, Table 1). While more CCSS survivors were treated with abdominal or pelvic radiation, chemotherapy exposures between cohorts were relatively consistent (Table 1, Appendix p 17–18). The median follow-up time from cancer diagnosis was longer among CCSS survivors (CCSS: 23.7 years, IQR: 18.3–30.0; SJLIFE: 15.1 years, IQR: 10.4–22.9). The estimated POI prevalence increased from 7.9% to 18.6% in CCSS and 7.3% to 14.9% in SJLIFE as survivors aged from 21 to 40 years. By age 40, 80% of the SJLIFE cohort had been censored, compared with 61% in CCSS (Appendix p 9). Among the 922 POI cases in CCSS, 531 experienced AOF, while 50 of the 101 POI cases in SJLIFE experienced AOF.

In CCSS, models based on age-specific logistic regression or XGBoost algorithms had the best performance overall (Appendix p 10). Of these, the ovarian and prescribed models were largely comparable over all evaluated ages (Appendix p 11). Therefore, age-specific logistic regression- and XGBoost-based models were selected for external validation in SJLIFE (details in Appendix). In the SJLIFE validation cohort, we observed no appreciable differences in performance metrics between the ovarian and prescribed models or the two risk prediction algorithms (Appendix p 11). Therefore, the following age-specific logistic regression-based models (with lower computational burden) were selected for dissemination:

the prescribed model, which showed excellent validation performance without requiring dosimetry, and the clinically-pragmatic simple model.

Performance metrics for the POI risk prediction models selected for dissemination are shown in Figure 2 (CCSS, panels 2A-2C; SJLIFE, panels 2D-2F). Across ages 21–40, overall risk prediction performance was substantially better in the SJLIFE validation cohort for both the prescribed model (SBrS_{CCSS}=16.8%–22.0%; SBrS_{SJLIFE}=47.2%–57.5%) and simple model (SBrS_{CCSS}=10.5%–15.1%; SBrS_{SJLIFE}=35.4%–46.4%) (Figures 2A, 2D). Discrimination for the prescribed model was reasonable in CCSS (AUROC=0.76–0.78) and very good to excellent in SJLIFE (AUROC=0.88–0.95). In SJLIFE, the simple model also showed very good discrimination (AUROC=0.82–0.90; Figures 2B, 2E). Similarly, the average positive predictive values were improved for both the prescribed (AUPRC=0.72–0.82) and simple models (AUPRC=0.58–0.71) in the SJLIFE (Figures 2C, 2F). Age-specific calibration metrics indicated both models were well-calibrated in CCSS (Figure 3A) and were reasonably calibrated in SJLIFE (Figure 3B, Appendix p 12). Detailed model performance metrics considering different age thresholds are shown in the Appendix (pp 19–20). In sensitivity analyses, we found the prescribed model with individual chemotherapy predictors had superior performance compared to models reducing individual alkylators to CED¹⁶, e.g., AUROCs of 0.77 versus 0.73 for the prescribed model and CED-based models, respectively, by age 40 (Appendix p 13). In CCSS, the observed POI prevalence by age 30 was 3%, 8%, 35% and 67% in the low, medium low, medium high, and high predicted risk groups under the prescribed model, respectively, and 2%, 9%, 39% and 74% by age 40 (Appendix p 21). The averaged regression coefficients for each model by age is shown in the Appendix (p 22).

The distribution of estimated odds ratios (ORs) for influential predictors in the prescribed model by age 22 in CCSS, i.e., $P < 0.1$ across all imputed datasets, is shown in the Appendix (p 14). For comparison, we also provide estimated POI risk associations for each of these influential predictors in non-imputed complete case data (6,322 CCSS survivors). Increasing pelvic radiation (per 10 Gray OR=2.03, 95% CI: 1.67–2.46) and procarbazine dose (per g/m² OR=1.09, 95% CI: 1.01–1.17), and treatment with HSCT (OR=4.65, 95% CI: 1.95–11.08), busulfan (OR=2.39, 95% CI: 1.06–5.41), or CCNU (OR=2.68, 95% CI: 1.14–6.24) were influential predictors and associated with increased POI risk in non-imputed data. Other influential predictors included age at cancer diagnosis and its respective statistical interactions with HSCT and pelvic radiation dose (details in Appendix, p 4).

Because longitudinal self-reported menstrual history was used for POI ascertainment in CCSS, we simulated training datasets reflecting non-differential outcome misclassification and evaluated the risk prediction performance of developed models in a simulated validation dataset without outcome misclassification (i.e., similar to SJLIFE, with clinically/biochemically ascertained POI). We found that as the outcome misclassification rate increased in the simulated training data, model performance remained consistent and was always superior in the simulated validation data than that of the simulated training data (training AUROC: 0.72–0.87, validation AUROC: 0.93; Appendix p 23).

We compared predicted risks from our models with a risk prediction model for POI onset age²⁷ (Appendix p 24), where the latter was based on mathematically modeling the age-related decline of ovarian reserve in an untreated population and the estimated radiation dose required to reduce the human non-growing follicle pool by 50%. Results from the previous radiation-focused POI onset age model²⁷ suggest survivors diagnosed with cancer before age 20 treated with MORD between 4–9 Gy would all develop POI by age 39. Our ovarian model classified survivors with the same set of risk factors as having medium low (>5% to <20%) predicted POI risk by age 40. We found this risk classification was consistent with the observed POI prevalence of 15.5% by age 40 (n=71) among a comparable subgroup of no-chemotherapy/HSCT CCSS survivors with MORD ranging from 4–9 Gy.

Among CCSS survivors with genotype data, we found that the addition of a SJLIFE-based PRS for treatment-related POI to the prescribed model improved POI prediction by age 40 among survivors exposed to any ovarian radiation and chemotherapy (Table 2), and in particular, contributed to a significant gain in the average positive predictive value (with PRS: AUPRC=0.87, 95% CI: 0.80–0.94; without PRS: AUPRC=0.76, 95% CI: 0.63–0.89; P=0.029). This gain in AUPRC with the inclusion of the PRS in at-risk survivors appeared to be driven by the reduction in the false positive rate (with PRS: 21.7%, without PRS: 62.0%; using the risk threshold 50% for classification).

DISCUSSION

To our knowledge, we have developed and validated the first models to estimate patient-specific POI risks by specific ages throughout young adulthood (from 21 to 40 years) in female survivors of childhood cancer. These models showed strong discriminatory and positive predictive performance in independent data (AUROCs=0.88–0.95, AUPRCs=0.72–0.82), indicating their appropriateness for use in a clinical setting. Notably, we demonstrated that a model featuring prescribed pelvic/abdominal radiotherapy doses performed almost as well as a model requiring dosimetry-based ovarian radiation dose. Interestingly, while CED is generally used to evaluate POI risk in survivors¹⁶, models that included individual alkylators instead of CED had superior performance (4-point increase in AUROC values), and a subset of alkylators (procarbazine, busulfan, CCNU) were influential POI predictors. In addition, a simple model that required no dose information showed very good validation performance, with AUROCs exceeding 0.85 between ages 21–33. These POI risk prediction models are freely available as an easy-to-use online risk calculator (<https://ccss.stjude.org/tools-documents/calculators-other-tools/primary-ovarian-insufficiency-risk-calculator.html>, see Appendix).

We previously developed risk prediction models for AOF,⁹ an early toxicity event. These AOF risk prediction models⁹ limit POI risk prediction to the time period within five years of childhood cancer diagnosis: they are not, for example, useful for survivors whose menses resumed after treatment but have higher residual premature menopause risk by the time they are ready to have children (e.g., age >30). The current study addressed this gap. While these models are not a substitute for direct clinical assessments of fertility, the prescribed model's strong performance among older survivors in SJLIFE (AUROCs=0.88–0.94, survivors aged >30 years) highlights the potential clinical utility of these models across the reproductive

age spectrum. For reference, we provide applications of these models (Appendix) describing how survivors may be clinically stratified at different ages.

The Pediatric Initiative Network (PIN) of the Oncofertility Consortium recently developed a standardized system to stratify female survivors into three categories of infertility risk, where risk categories can be assigned based on specific treatment exposures and the timing of treatment.²⁸ In contrast, our POI risk prediction models provide personalized risk estimates that are informed by multiple demographic and treatment characteristics. We also found our models produced predicted risks that better reflected the observed prevalence of POI in a cohort of aging female survivors, with results that differed from a previous radiation-focused mathematical model predicting POI onset age²⁷. There are considerable methodological differences between these models: in the current study, statistical modeling of a treated (survivor) population was conducted followed by validation in an independent survivor cohort. Host factors (e.g., race/ethnicity) and additional clinical predictors besides radiation (HSCT; chemotherapy doses, including alkylators) were included.

The main strength of this study is the strong evidence of these POI risk prediction models' generalizability, as demonstrated by their excellent performance in validation data. We also investigated the potential utility of a SJLIFE-based POI PRS¹⁸ in CCSS. In survivors treated with ovarian radiation and chemotherapy, we found that considering additional germline variants beyond the previously identified *NPY2R* locus led to a significant improvement in the predictive ability to identify survivors at high POI risk by age 40 compared with the prescribed model. These results suggest a model with this PRS could more accurately detect/classify POI cases in the subgroup of survivors with higher POI risks than a model with clinical predictors only. However, additional external validation is needed, along with more comprehensive POI genetic association studies including survivors of non-European genetic ancestry.

A limitation of this study is the limited racial/ethnic diversity in CCSS. While race/ethnicity is included as a predictor, more precise risk estimates could be obtained by using training data with greater diversity. High rates of censoring in long-term follow-up (i.e., ages >35 in CCSS) is a further issue which was mitigated with use of IPCW. Other important limitations of this study include the use of self-reported cessation of menstruation as a surrogate marker for diminished ovarian function in CCSS, and the differences in POI ascertainment in CCSS and SJLIFE. However, POI ascertainment among CCSS survivors considered longitudinal menstrual history information, to maximize consistency with POI ascertainment with hormone measurements as in SJLIFE. Encouragingly, the strong validation of these POI risk prediction models in SJLIFE with clinically ascertained POI suggests POI ascertainment in CCSS did not diminish their robustness. With emerging literature calling for prospective measures of ovarian function and infertility risk in survivors,^{29,30} models developed in data with POI ascertainment supported by such measures may yield POI risk prediction models with stronger performance; such models may be developed in the future (e.g., in SJLIFE as longitudinal follow-up increases). Notably, we found risk prediction models with robust predictors developed in simulated training data with minor outcome misclassification can have consistently superior performance in simulated validation data with no outcome misclassification, indicating that our models should perform better among

survivors with POI that is ascertained directly with hormone measurements. Lastly, this study could not assess predictors related to the treatment of SMNs (such participants were censored at diagnosis) or newer treatments, e.g., immunotherapy or targeted therapies. These models will need to be updated as such data become available.

In summary, we developed and validated POI risk prediction models for female childhood cancer survivors as they age. Compared to our previous AOF risk calculator⁹ which is better suited to guide the counseling of female patients and their families before cancer treatment begins, these novel POI risk prediction models can critically inform continuing risks after treatment for childhood cancer concludes, supporting individualized needs for fertility preservation during adulthood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This work was funded by the Canadian Institutes of Health Research (FRN 148693, Y Yuan principal investigator; FRN 388395, P Nathan principal investigator); the US National Cancer Institute (R21 CA261833, C Im/Y Yuan principal investigators; R01 CA216354, Y Yasui/J Zhang, MPIs); Canadian Centre for Applied Research in Cancer Control (studentship, L Yu). The study data was from the Childhood Cancer Survivor Study (U24 CA55727, GT Armstrong principal investigator) and the St. Jude Lifetime Cohort Study (U01 CA195547, MM Hudson/KK Ness, principal investigators).

Data sharing

This study only involved secondary data analysis. All relevant CCSS data is currently available and may be requested through <https://ccss.stjude.org/>. SJLIFE data used for model validation is currently available and may be accessed from the St. Jude Cloud (<https://stjude.cloud>) under the accession number SJC-DS-1002. All CCSS genotype and associated clinical data used in polygenic risk score analyses are available through the database of Genotypes and Phenotypes (dbGaP accession number: phs001327.v2.p1)

REFERENCES

1. Chemaitilly W, Li Z, Krasin MJ, et al. Premature ovarian insufficiency in childhood cancer survivors: a report from the St. Jude Lifetime Cohort. *The Journal of Clinical Endocrinology & Metabolism* 2017; 102(7): 2242–50. [PubMed: 28368472]
2. Mostoufi-Moab S, Seidel K, Leisenring WM, et al. Endocrine abnormalities in aging survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *Journal of Clinical Oncology* 2016; 34(27): 3240. [PubMed: 27382091]
3. Webber L, Davies M, Anderson R, et al. ESHRE Guideline: management of women with premature ovarian insufficiency. *Human Reproduction* 2016; 31(5): 926–37. [PubMed: 27008889]
4. Johnston RJ, Wallace WHB. Normal ovarian function and assessment of ovarian reserve in the survivor of childhood cancer. *Pediatric blood & cancer* 2009; 53(2): 296–302. [PubMed: 19514070]
5. Levine JM, Whitton JA, Ginsberg JP, et al. Nonsurgical premature menopause and reproductive implications in survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *Cancer* 2018; 124(5): 1044–52. [PubMed: 29338081]
6. Mishra GD, Pandeya N, Dobson AJ, et al. Early menarche, nulliparity and the risk for premature and early natural menopause. *Human Reproduction* 2017; 32(3): 679–86. [PubMed: 28119483]

7. Mulder RL, Font-Gonzalez A, Hudson MM, et al. Fertility preservation for female patients with childhood, adolescent, and young adult cancer: recommendations from the PanCareLIFE Consortium and the International Late Effects of Childhood Cancer Guideline Harmonization Group. *The Lancet Oncology* 2021; 22(2): e45–e56. [PubMed: 33539753]
8. Lambertini M, Peccatori FA, Demeestere I, et al. Fertility preservation and post-treatment pregnancies in post-pubertal cancer patients: ESMO Clinical Practice Guidelines. *Annals of oncology* 2020; 31(12): 1664–78. [PubMed: 32976936]
9. Clark RA, Mostoufi-Moab S, Yasui Y, et al. Predicting acute ovarian failure in female survivors of childhood cancer: a cohort study in the Childhood Cancer Survivor Study (CCSS) and the St Jude Lifetime Cohort (SJLIFE). *The Lancet Oncology* 2020; 21(3): 436–45. [PubMed: 32066539]
10. Wallace WHB, Smith AG, Kelsey TW, Edgar AE, Anderson RA. Fertility preservation for girls and young women with cancer: population-based validation of criteria for ovarian tissue cryopreservation. *The Lancet Oncology* 2014; 15(10): 1129–36. [PubMed: 25130994]
11. Corkum KS, Rhee DS, Wafford QE, et al. Fertility and hormone preservation and restoration for female children and adolescents receiving gonadotoxic cancer treatments: a systematic review. *Journal of pediatric surgery* 2019; 54(11): 2200–9. [PubMed: 30773394]
12. Robison LL, Armstrong GT, Boice JD, et al. The Childhood Cancer Survivor Study: a National Cancer Institute–supported resource for outcome and intervention research. *Journal of clinical oncology* 2009; 27(14): 2308. [PubMed: 19364948]
13. Howell CR, Bjornard KL, Ness KK, et al. Cohort profile: the St. Jude Lifetime Cohort Study (SJLIFE) for paediatric cancer survivors. *International Journal of Epidemiology* 2021; 50(1): 39–49. [PubMed: 33374007]
14. Sklar CA, Mertens AC, Mitby P, et al. Premature menopause in survivors of childhood cancer: a report from the childhood cancer survivor study. *Journal of the National Cancer Institute* 2006; 98(13): 890–6. [PubMed: 16818852]
15. Chemaitilly W, Mertens AC, Mitby P, et al. Acute ovarian failure in the childhood cancer survivor study. *The Journal of Clinical Endocrinology & Metabolism* 2006; 91(5): 1723–8. [PubMed: 16492690]
16. Green DM, Nolan VG, Goodman PJ, et al. The cyclophosphamide equivalent dose as an approach for quantifying alkylating agent exposure: a report from the Childhood Cancer Survivor Study. *Pediatric blood & cancer* 2014; 61(1): 53–67. [PubMed: 23940101]
17. Howell RM, Smith SA, Weathers RE, Kry SF, Stovall M. Adaptations to a Generalized Radiation Dose Reconstruction Methodology for Use in Epidemiologic Studies: An Update from the MD Anderson Late Effect Group. *Radiation research* 2019; 192(2): 169–88. [PubMed: 31211642]
18. Brooke RJ, Im C, Wilson CL, et al. A high-risk haplotype for premature menopause in childhood cancer survivors exposed to gonadotoxic therapy. *Journal of the National Cancer Institute* 2018; 110(8): 895–904. [PubMed: 29432556]
19. Sapkota Y, Cheung YT, Moon W, et al. Whole–Genome Sequencing of Childhood Cancer Survivors Treated with Cranial Radiation Therapy Identifies 5p15. 33 Locus for Stroke: A Report from the St. Jude Lifetime Cohort Study. *Clinical Cancer Research* 2019; 25(22): 6700–8. [PubMed: 31462438]
20. Van Buuren S *Flexible imputation of missing data*: CRC press; 2018.
21. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* 2005; 67(2): 301–20.
22. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*; 2016; 2016. p. 785–94.
23. Yuan Y, Zhou QM, Li B, Cai H, Chow EJ, Armstrong GT. A threshold-free summary index of prediction accuracy for censored time to event data. *Statistics in medicine* 2018; 37(10): 1671–81. [PubMed: 29424000]
24. Hu B, Palta M, Shao J. Properties of R2 statistics for logistic regression. *Statistics in medicine* 2006; 25(8): 1383–95. [PubMed: 16059870]
25. Pepe MS. *The statistical evaluation of medical tests for classification and prediction*: Oxford University Press, USA; 2003.

26. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. *Statistics in medicine* 1986; 5(5): 421–33. [PubMed: 3786996]
27. Kelsey TW, Hua C-H, Wyatt A, Indelicato D, Wallace WH. A predictive model of the effect of therapeutic radiation on the human ovary. *Plos one* 2022; 17(11): e0277052. [PubMed: 36399448]
28. Meacham LR, Burns K, Orwig KE, Levine J. Standardizing risk assessment for treatment-related gonadal insufficiency and infertility in childhood adolescent and young adult cancer: the pediatric initiative network risk stratification system. *Journal of Adolescent and Young Adult Oncology* 2020; 9(6): 662–6. [PubMed: 32456570]
29. Anderson RA, Cameron D, Clatot F, et al. Anti-Müllerian hormone as a marker of ovarian reserve and premature ovarian insufficiency in children and women with cancer: a systematic review. *Human reproduction update* 2022; 28(3): 417–34. [PubMed: 35199161]
30. Drechsel KC, Pilon MC, Stoutjesdijk F, et al. Reproductive ability in survivors of childhood, adolescent, and young adult Hodgkin lymphoma: a review. *Human Reproduction Update* 2023: dmad002.

RESEARCH IN CONTEXT

Evidence before this study

Female survivors of childhood cancer treated with abdominal/pelvic radiotherapy and/or chemotherapy are at risk for cessation of gonadal function before the age of 40 years, i.e., primary ovarian insufficiency (POI). Previously, we developed treatment-informed prediction models to estimate survivors' risks for experiencing acute ovarian failure (AOF), defined as permanent loss of ovarian function within five years of cancer diagnosis. While these models showed excellent prediction performance in independent validation data, they cannot predict risks of developing POI later in life. Given that most childhood cancer patients will become long-term survivors, models that accurately estimate POI risks across the reproductive age spectrum would be valuable for informing survivors of their needs for fertility preservation interventions. A PubMed search until January 1st, 2022 with the terms “pediatric cancer OR childhood cancer” AND “primary ovarian insufficiency” AND “prediction” for research articles published in English showed no additional studies pertaining to this topic other than our previous work.

Added value of this study

To our knowledge, these are the first age-specific risk prediction models for POI that have been developed and validated in large, well-characterized cohorts of long-term childhood cancer survivors. We found a comprehensive model that incorporated cumulative prescribed/delivered doses of abdominal/pelvic radiation, hematopoietic stem cell transplantation information, and a range of chemotherapy agents, especially individual alkylators, performed well in internal and external data (validation area under the receiver operating characteristic curve [AUROC] values=0.88–0.95). Importantly, this model showed evidence of generalizability to older adult survivors (over 30 years of age: validation AUROCs=0.88–0.94). This study is among the first to evaluate genetic predictors, specifically a polygenic risk score consisting of genetic variants associated with treatment-related POI in survivors. These models are accessible as a user-friendly online calculator (<https://ccss.stjude.org/tools-documents/calculators-other-tools/primary-ovarian-insufficiency-risk-calculator.html>).

Implications of all the available evidence

Addressing fertility is an important aspect of oncology care and survivorship. Our models can inform reproductive health counseling by providing individualized POI risk estimates for childhood cancer patients at diagnosis and for survivors during the years following the completion of therapy. Our findings demonstrate the utility of developing accurate POI risk prediction tools for survivors, especially one that incorporates novel host and treatment risk predictors, to better inform the timing of discussions about fertility preservation.

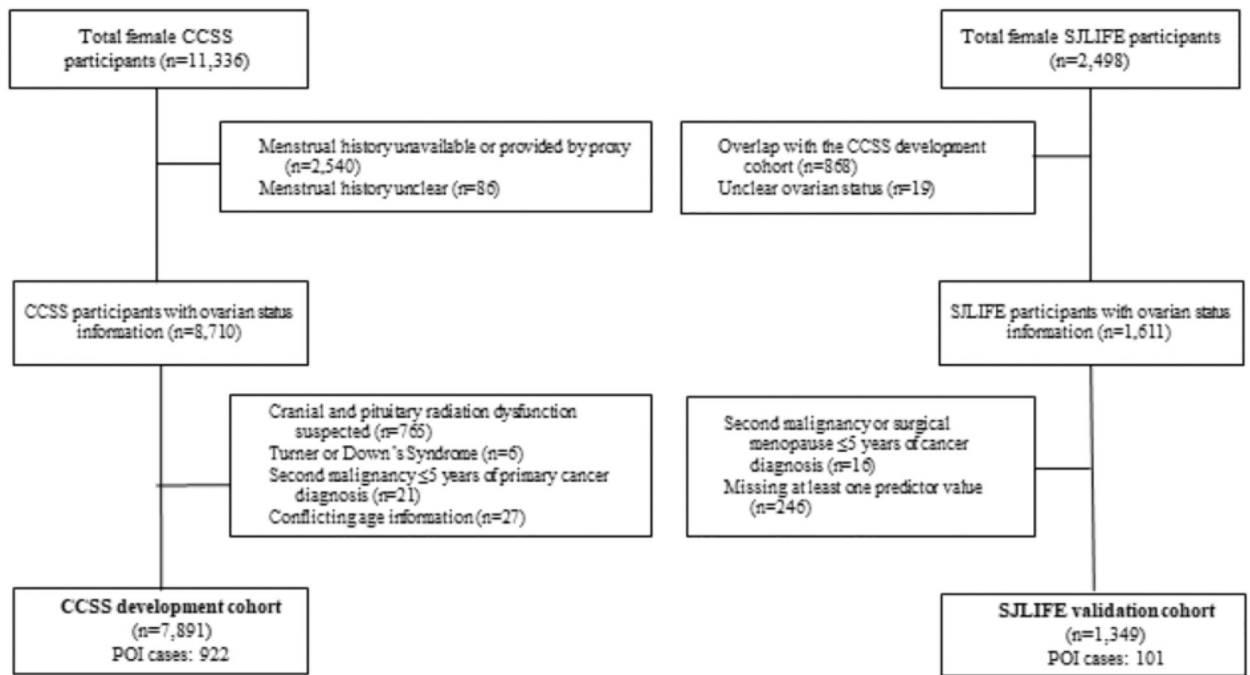


Figure 1: Study cohort flowchart for inclusion in POI risk prediction model development (CCSS) and validation (SJLIFE) datasets.

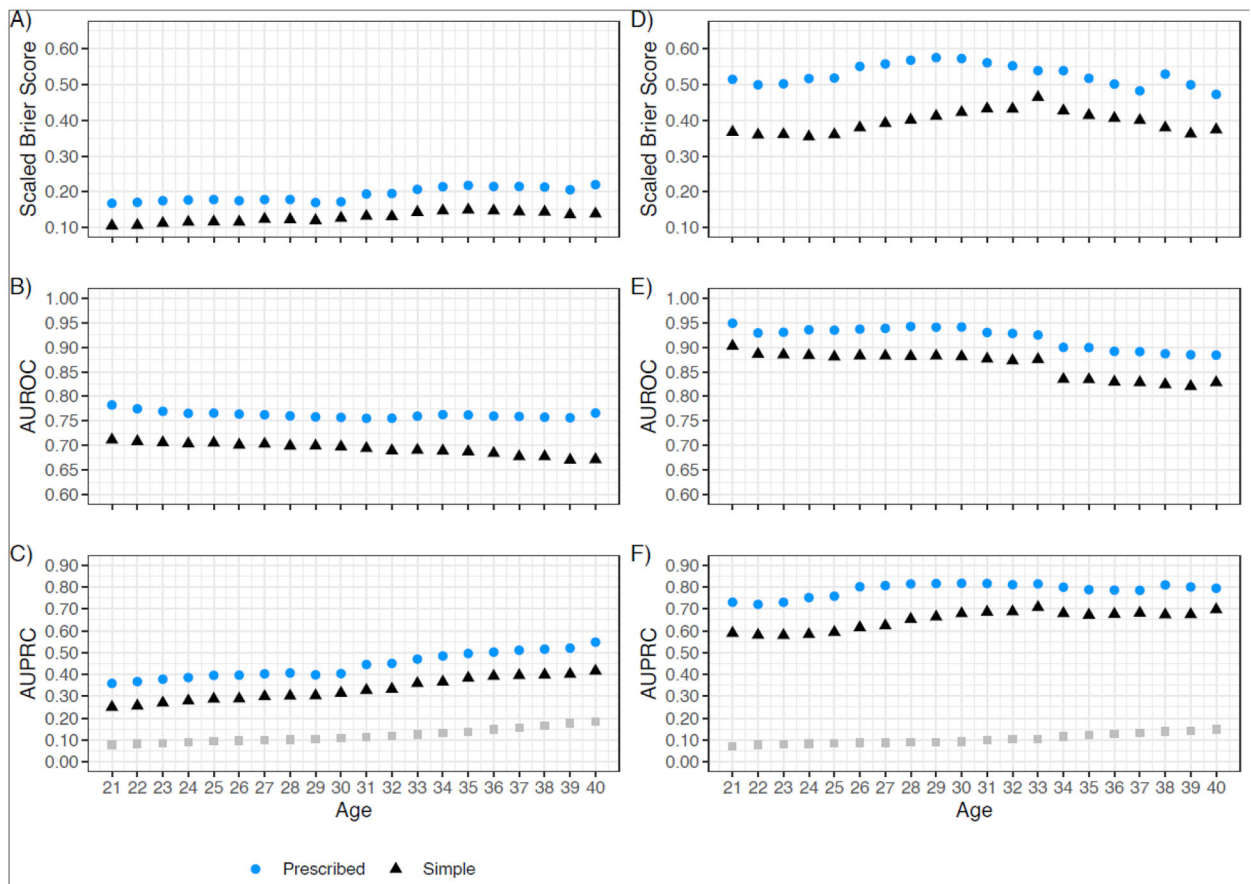


Figure 2:

POI risk prediction performance indices for the age-specific logistic regression algorithm for the prescribed dose model and simple model in the CCSS model development cohort and SJLIFE model validation cohort. CCSS results are shown on the left (panels A-C), SJLIFE results are shown on the right (panels D-F). Ages (in years) are shown on the x-axis. Age-specific logistic regression-based metrics are differentiated by color and shape for the prescribed dose model (blue circles) and the simple model (black triangles). Estimated age-specific rates of POI in each cohort are shown in panels C and F as grey squares. Scaled Brier scores are shown in the top row in CCSS (A) and SJLIFE (D); area under the ROC curve values (AUROC) are shown in the middle row for CCSS (B) and SJLIFE (E); area under the precision-recall curve (AUPRC) values are shown in the bottom row for CCSS (C) and SJLIFE (F).

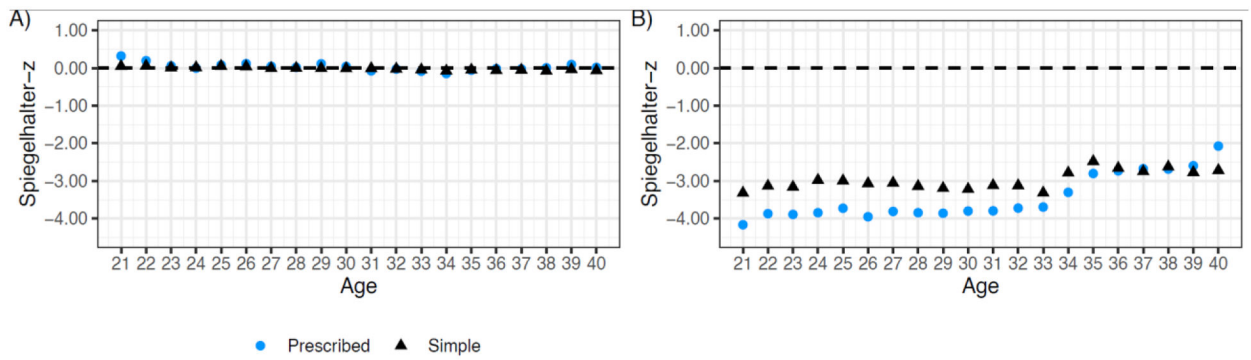


Figure 3: Calibration indices of the age-specific logistic regression algorithm from ages 21 to 40 years in the CCSS model development cohort and SJLIFE model validation cohort. CCSS results are shown on the left (panel A) and SJLIFE results are shown on the right (panel B). Ages (in years) are shown on the x-axis for each panel. Metrics are differentiated by color and shape for the prescribed dose model (blue circles) and the simple model (black triangles).

Table 1:

Clinical and treatment characteristics of the CCSS development cohort and SJLIFE validation cohort

	CCSS imputed (N=7891) N (%)	CCSS complete case (N=6322) N (%)	SJLIFE (N=1349) N (%)
Median age at cancer diagnosis, years (IQR)	7.5 (3.3 – 13.9)	7.3 (3.2 – 13.8)	5.6 (2.2 – 12.3)
Cancer Diagnosis			
Leukemia	2589 (32.8)	2061 (32.6)	408 (30.2)
Hodgkin disease	1182 (15.0)	922 (14.6)	89 (6.6)
Kidney tumors	948 (12.0)	791 (12.5)	84 (6.2)
Bone cancer	782 (9.9)	639 (10.1)	76 (5.6)
Central nervous system tumors	767 (9.7)	646 (10.2)	184 (13.6)
Neuroblastoma	643 (8.1)	516 (8.2)	52 (3.8)
Non-Hodgkin lymphoma	527 (6.7)	405 (6.4)	49 (3.6)
Soft tissue sarcoma	453 (5.7)	342 (5.4)	88 (6.5)
Other	NA	NA	319 (23.6)
Race			
White	6795 (86.1)	5538 (87.6)	1032 (76.5)
Black	483 (6.1)	370 (5.9)	267 (19.8)
Other	523 (6.6)	414 (6.5)	50 (3.7)
Missing	90 (1.1)	NA	NA
Minimum ovarian radiation dose, Gy			
None	3734 (47.3)	3411 (54.0)	846 (62.7)
<10	2993 (37.9)	2541 (40.2)	414 (30.7)
10 to <20	294 (3.7)	234 (3.7)	61 (4.5)
20	161 (2.0)	136 (2.2)	28 (2.1)
Missing	709 (9.0)	NA	NA
Abdominal radiation dose, Gy			
None	3734 (47.3)	3411 (54.0)	836 (62.0)
< 10	2037 (25.8)	1680 (26.6)	318 (23.6)
10 to < 20	502 (6.4)	406 (6.4)	93 (6.9)
20	969 (12.3)	825 (13.0)	102 (7.6)
Missing	649 (8.2)	NA	NA
Pelvic radiation dose, Gy			
None	3734 (47.3)	3411 (54.0)	836 (62.0)
< 10	2400 (30.4)	1995 (31.6)	331 (24.5)
10 to < 20	402 (5.1)	327 (5.2)	83 (6.2)
20	707 (9.0)	589 (9.3)	99 (7.3)
Missing	648 (8.2)	NA	NA
Hematopoietic stem cell transplant			
Yes	338 (4.3)	222 (3.5)	83 (6.2)
No	6934 (87.9)	6100 (96.5)	1266 (93.8)
Missing	619 (7.9)	NA	NA

	CCSS imputed (N=7891) N (%)	CCSS complete case (N=6322) N (%)	SJLIFE (N=1349) N (%)
Total body radiation dose, Gy			
None	7085 (89.8)	6208 (98.2)	1310 (97.1)
< 10	27 (0.3)	17 (0.3)	0 (0.0)
10 to < 20	135 (1.7)	97 (1.5)	39 (2.9)
Missing	644 (8.2)	NA	NA
Alkylating agent dose ^a (CED, mg/m ²)			
None	3620 (45.9)	3289 (52.0)	647 (48.0)
< 4000	1001 (12.7)	816 (12.9)	261 (19.3)
4000 to < 8000	1042 (13.2)	863 (13.7)	167 (12.4)
8000	1255 (15.9)	1354 (21.4)	274 (20.3)
Missing	973 (12.3)	NA	NA
POI prevalence ^b (95% CI)	18.6 (17.3 – 20.0)	16.4 (14.9 – 17.9)	14.9 (11.6 – 19.1)

Abbreviations: POI, primary ovarian insufficiency; IQR, interquartile range; CI, 95% confidence interval; Gy, Gray; mg, milligrams; m, meters; CED, cyclophosphamide equivalent dose.

^a. See Appendix p 17 for dose distributions for specific chemotherapy predictors.

^b. POI prevalence by age 40 was estimated in the whole cohort accounting for inverse probability censoring weights.

Table 2:

Comparison of POI risk prediction models with non-genetic clinical predictors only versus models including a POI polygenic risk predictor in CCSS

Risk prediction metrics	Clinical risk score ^b	POI PRS	P-value
All female survivors (n=1,985; POI prevalence ^a : 11%)			
SBrS (95% CI)	15.0% (10.0% to 21.0%)	16.0% (11.0% to 21.0%)	0.85
AUPRC (95% CI)	0.40 (0.34 to 0.48)	0.39 (0.33 to 0.47)	0.43
AUROC (95% CI)	0.79 (0.75 to 0.83)	0.77 (0.71 to 0.82)	0.13
Spiegelhalter-z (95% CI)	-2.42 (-4.38 to -0.46)	-0.34 (-2.30 to 1.62)	0.19
Any ovarian RT and chemotherapy (n=158; POI prevalence ^a : 58%)			
SBrS (95% CI)	-1.4% (-31.6% to 16.3%)	20.0% (1.0% to 35.0%)	0.018
AUPRC (95% CI)	0.76 (0.63 to 0.89)	0.87 (0.80 to 0.94)	0.029
AUROC (95% CI)	0.70 (0.57 to 0.82)	0.78 (0.70 to 0.87)	0.12
Spiegelhalter-z (95% CI)	2.43 (0.47 to 4.39)	1.63 (-0.33 to 3.59)	0.98
True positive rate ^c	38.0%	78.3%	--
Positive predictive value ^d	71.6%	70.2%	--

Abbreviations: CI: confidence interval; SBrS, scaled Brier score; AUROC, area under the ROC curve; AUPRC, area under the precision-recall curve; POI PRS, 6-variant POI polygenic risk score developed from a SJLIFE genome-wide association study.

^aPOI prevalence by age 40 years; estimated considering inverse probability censoring weights.

^bLogistic regression-based prescribed dose model at age 40 years.

^cDefining high risk as model predicted risks >50%, the true positive rate = (survivors with POI classified as high risk)/(survivors with POI).

^dDefining high risk as model predicted risks >50%, the positive predictive value = (survivors with POI classified as high risk)/(survivors classified as high risk).