

## ORIGINAL RESEARCH

# Prediction of Lifetime and 10-Year Risk of Cancer in Individual Patients With Established Cardiovascular Disease



Cilie C. van 't Klooster, MD,<sup>a</sup> Paul M. Ridker, MD, PhD,<sup>b,c</sup> Nancy R. Cook, PhD,<sup>b</sup> Joachim G.J.V. Aerts, MD, PhD,<sup>d</sup> Jan Westerink, MD, PhD,<sup>a</sup> Folkert W. Asselbergs, MD, PhD,<sup>e,f,g</sup> Yolanda van der Graaf, MD, PhD,<sup>h</sup> Frank L.J. Visseren, MD, PhD,<sup>a</sup> on behalf of the UCC-SMART Study Group

## ABSTRACT

**BACKGROUND** Cardiovascular disease (CVD) and cancer share many common risk factors; patients with CVD also may be at risk of developing cancer.

**OBJECTIVES** The aim of this study was to derive and externally validate prediction models for the estimation of lifetime and 10-year risk for total, colorectal, and lung cancer in patients with established CVD.

**METHODS** Data from patients with established CVD from the UCC-SMART cohort (N = 7,280) were used for model development, and from the CANTOS trial (N = 9,322) for model validation. Predictors were selected based on previously published cancer risk scores, clinical availability, and presence in the derivation dataset. Fine and Gray competing risk-adjusted lifetime models were developed for the outcomes total, colorectal, and lung cancer.

**RESULTS** Selected predictors were age, sex, smoking, weight, height, alcohol use, antiplatelet use, diabetes, and C-reactive protein. External calibration for the 4-year risk of lung, colorectal, and total cancer was reasonable in our models, as was discrimination with C-statistics of 0.74, 0.64, and 0.63, respectively. Median predicted lifetime and 10-year risks in CANTOS were 26% (range 1% to 52%) and 13% (range 1% to 31%) for total cancer; 4% (range 0% to 13%) and 2% (range 0% to 6%) for colorectal cancer; and 5% (range 0% to 37%) and 2% (range 0% to 24%) for lung cancer.

**CONCLUSIONS** Lifetime and 10-year risk of total, colorectal, and lung cancer can be estimated reasonably well in patients with established CVD with readily available clinical predictors. With additional study, these tools could be used in clinical practice to further aid in the emphasis of healthy lifestyle changes and to guide thresholds for targeted diagnostics and screening. (J Am Coll Cardiol CardioOnc 2020;2:400-10) © 2020 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

From the <sup>a</sup>Department of Vascular Medicine, University Medical Center Utrecht (UMCU), Utrecht University, Utrecht, the Netherlands; <sup>b</sup>Divisions of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; <sup>c</sup>Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston Massachusetts; <sup>d</sup>Department of Respiratory Medicine, Erasmus Medical Centre, Rotterdam, the Netherlands; <sup>e</sup>Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht (UMCU), Utrecht University, Utrecht, the Netherlands; <sup>f</sup>Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom; <sup>g</sup>Health Data Research UK and Institute of Health Informatics, University College London, London, United Kingdom; and the <sup>h</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht (UMCU), Utrecht University, Utrecht, the Netherlands. Dr. Ridker has served as the principal investigator of CANTOS and CIRT, funded by Novartis and the National Heart, Lung, and Blood Institute (NHLBI), respectively; has received investigator-initiated research grants from Kowa, Novartis, Pfizer, AstraZeneca, NHLBI, and National Cancer Institute; has served as a consultant to Corvidia, Inflammazome, Novartis, Amgen, Merck, and Civi

**T**reatment for cardiovascular disease (CVD) has improved substantially over the past decades, with more patients surviving CVD and living long enough to develop other diseases such as cancer. Besides an increased risk of new cardiovascular events, patients with established CVD have a higher risk of cancer compared with the general population (standardized incidence ratio of 1.19; 95% confidence interval [CI]: 1.10 to 1.29 adjusted for age, sex, and calendar year) (1), most likely due to shared risk factors including obesity, smoking, and low-grade inflammation (2,3). Furthermore, even though CVD is still the leading cause of mortality worldwide among adults, in some higher- and middle-income countries, cancer has become the predominant cause of death, partly due to improved prevention and treatment of CVD (4).

Given one's absolute individual cancer risk varies, several risk prediction models have been developed to estimate the absolute risk for incident cancer of a specific type, notably lung cancer and breast cancer (5-9). However, no prediction models are available for patients with established CVD specifically. Furthermore, from a patient's perspective, risk of any cancer might be a more relevant metric, and no risk prediction models estimate total cancer risk. Furthermore, classic risk prediction models estimate prognosis in terms of absolute 5- or 10-year risk of cancer, and may not identify those patients who have a relatively low 5- or 10-year absolute risk, but a high cumulative lifetime risk (10). Finally, traditional 10-year risk prediction scores often do not consider the competing risk of noncancer mortality, and are prone to several types of bias (11). Especially in a population of patients with established CVD, the competing risk of noncancer mortality including cardiovascular death should be taken into account to prevent overestimation of cancer risk.

Estimating individualized probabilities could help in patients' and clinicians' understanding of cancer risk. As several modifiable risk factors are related to cancer (2) as well as to CVD, discussing these cancer

risks with patients could potentially aid in emphasizing healthy lifestyle changes, such as smoking cessation or weight loss. The aim of the current study was to develop and externally validate prediction models to estimate the 10-year and lifetime risk for total, colorectal, and lung cancer in patients with established CVD.

## METHODS

**STUDY POPULATIONS.** Model development was conducted in the UCC-SMART (Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease) study, an ongoing prospective cohort study, including 18- to 79-year-old patients referred to the University Medical Center Utrecht with clinically manifest vascular disease or atherosclerotic risk factors. The cohort was initiated in 1996 and is still recruiting patients annually. For the current study 7,280 patients age 45 to 80 years with clinically manifest vascular disease and who gave permission for data requests to other medical authorities were included.

External model validation was performed in the CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study; NCT01327846), a double-blind, placebo-controlled, randomized clinical trial, that included 10,061 participants with a myocardial infarction at least 1 month before study entry and elevated C-reactive protein (CRP) concentration ( $\geq 2$  mg/l). Eligible patients were randomized to receive either placebo or canakinumab at a dose of 50 mg, 150 mg, or 300 mg (12). For the current study 9,322 patients were included, after exclusion of patients younger than 45 years or older than 80 years. Detailed descriptions of the UCC-SMART cohort and the CANTOS trial have been published elsewhere (12-14). The studies were approved by institutional review boards and all participants provided written informed consent. An overview of eligibility criteria is provided in [Supplemental Table 1](#).

## ABBREVIATIONS AND ACRONYMS

**AIC** = Akaike's Information Criterion  
**CI** = confidence interval  
**CVD** = cardiovascular disease  
**CRP** = C-reactive protein  
**SD** = standard deviation

Bio; and is listed as co-inventor on patents related to the use of inflammatory biomarkers in cardiovascular disease and diabetes that are no longer active. Dr. Aerts has served as a consultant for Eli-Lilly, Merck Sharp & Dohme, Bristol-Myers Squibb, Boehringer Ingelheim, Amphera, AstraZeneca, Takeda, and Roche; and own stock in Amphera. Dr. Asselbergs is supported by UCL Hospitals NIHR Biomedical Research Centre. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose. The CANTOS trial was funded by Novartis Pharmaceuticals. The UCC-SMART study was financially supported by a grant of the University Medical Center Utrecht. The funder had no role in study design and conduct of the study, in the collection, analysis, and interpretation of the data, and in the preparation, review, or approval of the manuscript. The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the *JACC: CardioOncology* [author instructions page](#).

Manuscript received May 19, 2020; revised manuscript received July 5, 2020, accepted July 9, 2020.

**OUTCOMES.** During follow-up, participants enrolled in the UCC-SMART cohort received biannual questionnaires, gathering information on occurrence of recurrent CVD, bleeding events, incident diabetes mellitus, and end-stage renal disease. Additional information was collected from the hospital or general practitioner's records. An endpoint committee of 3 physicians adjudicated all clinical events independently and conflicting classifications were resolved through consensus. For data on cancer incidence, the UCC-SMART database was linked to the Dutch National Cancer Registry (INKL), a national registry receiving notifications of all new cancer diagnoses in the Netherlands through the Nationwide Network and Registry of Histopathology and Cytopathology, and hospital discharge diagnoses.

Participants in the CANTOS trial were followed for incident CVD as well as cancer diagnoses. Even though the primary endpoint of the trial was CVD incidence, patients' records were investigated for cancers reported during follow-up, as prespecified in the trial safety monitoring plan. Incident cancer reports were adjudicated by an endpoint committee of oncologists, blinded to treatment allocation (15). An overview of cancer diagnoses during follow-up for both study populations is provided in [Supplemental Tables 2 and 3](#). For the current study, total cancer was defined as any invasive neoplasm, excluding nonmelanoma skin cancer. As lung and colorectal cancer are the most common (not sex specific) cancers worldwide (16), these were chosen as separate outcomes. For the endpoint of total cancer, only first diagnoses of cancer were counted. For lung and colorectal cancer, the first diagnosis of that particular cancer type was included, possibly being the second or third primary diagnosis of cancer for a certain patient during follow-up.

**DATA PREPARATION AND PREDICTOR SELECTION.** Missing data (per variable  $\leq 1.1\%$  for UCC-SMART and  $\leq 0.2\%$  for CANTOS) were singly imputed by weighted probability matching using multivariable regression for the baseline and outcome data. Complete case analysis yielded similar model coefficients. Continuous variables were truncated at the 1st and 99th percentile to limit the effect of outliers on the model coefficients (i.e., leverage) (17). To prevent overfitting, predictors were preselected based on presence in previously published risk prediction models of multiple cancer types. Antiplatelet use (aspirin, P2Y<sub>12</sub>-ADP receptor antagonist, or other, such as dipyridamole) was added as a predictor, due to its inclusion in multiple previously published prediction models for colorectal cancer and due to

the common use of antiplatelet therapy in patients with CVD. Furthermore, it was required that the variables were readily clinically available, as well as present in the derivation dataset. This led to the following predictors: age, sex, smoking status, weight, height, alcohol use, use of antiplatelet medication, and diabetes mellitus ([Supplemental Table 4](#) details an overview of predictor selection). In addition, CRP was added as a predictor after a literature search for predictors of cancer was performed (3,15,18,19). Definitions of the predictors in the UCC-SMART cohort and CANTOS trial are provided in [Supplemental Table 5](#).

#### **DEVELOPMENT OF A PREDICTION MODEL FOR TOTAL CANCER, COLORECTAL CANCER, AND LUNG CANCER.**

Methods have been described in detail previously (10,11). Three separate complementary Fine and Gray competing risk-adjusted sub-distribution hazard functions (20,21) were developed in the UCC-SMART cohort for 10-year and lifetime risk predictions of: 1) total cancer; 2) colorectal cancer; and 3) lung cancer, with consideration for the competing risks of: 1) noncancer death; 2) non-colorectal cancer death; and 3) non-lung cancer death, respectively. As the endpoints colorectal and lung cancer included potential second or third primary diagnoses of cancer for a particular patient, the competing risks for these outcomes did not include other cancer types. The models were developed with left truncation: age rather than follow-up time was used as the underlying time scale. This way, patients contributed person-years between age at study entry and age at study exit, resulting in overlapping observations that allow for lifetime predictions across the range of baseline ages. Because a limited number of patients and events in certain age groups led to instability of predictions, the age range at baseline was restricted to 45 to 80 years.

The proportional hazards assumption was assessed visually by plotting scaled Schoenfeld residuals against time, and interactions with age (underlying time scale) were added to the model when a violation was observed. Log and quadratic associations between continuous predictors and the outcome variable were assessed by comparing model fit based on Akaike's Information Criterion (AIC) (17), and transformations were applied when appropriate to improve robustness of the model. AICs of models with and without addition of CRP as a predictor were compared to assess differences in model fit. Coefficients of the predictors were adjusted to account for optimism using a shrinkage factor acquired by bootstrapping with 1,000 bootstrap samples.

**INDIVIDUAL CANCER RISK PREDICTIONS.** Individual 10-year and lifetime risk of total, colorectal, and lung cancer, as well as life expectancy without cancer were estimated using the respective models. These predictions were derived from an individual lifetable with 1-year time intervals (22). First, starting at the baseline age for each patient, the risk of the event of interest ( $a_t$ ) and the risk of the competing event ( $b_t$ ) was calculated for each following life-year. Next, for each subsequent age year the probability of being healthy and alive at the start of that time interval (age year) ( $e_{t+1}$ ) was calculated by multiplying the survival probability ( $e_t$ ) by the event-free survival probability during that year ( $1 - a_t - b_t$ ). These steps were repeated from the age at baseline of an individual patient to the maximum age of 90 years, and together these predictions form an individual lifetable (10,23). The cancer-free life expectancy was determined as the age at which the median estimated cancer-free survival curve is 50%. For 10-year and lifetime risk of cancer, the cumulative cause-specific risks were truncated at 10 years after the age at baseline, and at the age of 90 years, respectively.

**INTERNAL AND EXTERNAL VALIDATION OF VALIDATION OF THE MODELS.** Internal validation of the total cancer, colorectal cancer, and lung cancer models was performed at 10 years of follow-up in the UCC-SMART data. External validation of the total, colorectal, and lung cancer models was evaluated in outcome data from the CANTOS trial at 4 years of follow-up (approximation of the median follow-up time in the CANTOS trial) by implementing the 4-year baseline hazard from the derivation dataset (UCC-SMART). To adjust for treatment effects of canakinumab, hazard ratios of treatment effects of canakinumab on cancer outcomes and their competing mortality were determined and added to the respective models. Discrimination was assessed using Harrell's c-statistic for survival data, and goodness of fit was assessed by calibration plots of the predicted versus observed risks. For the calibration plot, patients were divided into equal groups of increasing predicted risk. Based on the number of events, patients were divided into 10 equal groups for the total cancer model, and patients were divided into 6 equal groups for the colorectal and lung cancer models. Observed risks were estimated in these groups by using a cumulative incidence function, accounting for competing risks. Recalibration was performed based on the expected to observed ratio. Predicted risks in the CANTOS trial were estimated after recalibration. The Brier score was calculated for 4-year predictions in CANTOS, with confidence intervals based on the percentile

method with 1,000 bootstrap samples with replacement.

For comparison, simple models for total, colorectal, and lung cancer with sex and smoking status as the only predictors and with age as the underlying time scale were developed in the UCC-SMART study and externally validated in the CANTOS study population by the same methodology.

All analyses were performed in R statistical software, version 3.5.1, for model development, and 3.6.0. for external validation analyses (packages Hmisc, rms, cmprsk, car). To facilitate the use of this model in clinical practice, an online calculator will be developed.

## RESULTS

Baseline characteristics of the UCC-SMART and CANTOS study populations are shown in [Table 1](#). During a median follow-up time of 8.1 years (interquartile range 4.5 to 12.1 years), a total number of 1,143 first cancers were diagnosed in patients enrolled in the UCC-SMART cohort. Lung cancer occurred in 258 patients and colorectal cancer in 180 patients. Incidence rates for total cancer and noncancer mortality as a competing event were 1.97 (95% CI: 1.85 to 2.08) and 1.91 (95% CI: 1.80 to 2.02) per 100 person-years, respectively. Median follow-up time of the CANTOS trial was 3.8 years (interquartile range: 3.2 to 4.5 years), during which a total number of 509 incident cancers were diagnosed, 123 lung cancers, and 72 colorectal cancers. Incidence rates of total cancer and noncancer mortality were 1.48 (95% CI: 1.35 to 1.61) and 2.21 (95% CI: 2.05 to 2.37), respectively. An overview of incidence rates is shown in [Supplemental Table 6](#).

**DEVELOPMENT OF LIFETIME RISK PREDICTION MODELS FOR COLORECTAL, LUNG, AND TOTAL CANCER IN UCC-SMART.** Results of model development are shown in [Supplemental Tables 7 to 10](#). Transformations of continuous predictors, and interactions with age for continuous as well as categorical predictors are shown in [Supplemental Table 7](#). Age-specific baseline survival is shown in [Supplemental Table 8](#). Subdistribution hazard ratios and shrinkage factors are shown in [Supplemental Table 9](#), and model formulas of the total cancer, colorectal cancer, and lung cancer models are provided in [Supplemental Table 10](#). The AIC was lower for total cancer, colorectal cancer, and lung models with CRP compared with the same model without CRP.

**INTERNAL AND EXTERNAL VALIDATION OF TOTAL, COLORECTAL, AND LUNG CANCER MODELS.** Internal validation showed good agreement between the predicted and observed 10-year risk for total,

**TABLE 1** Baseline Characteristics of UCC-SMART and CANTOS Study Populations

	UCC-SMART (N = 7,280)	CANTOS (N = 9,322)
Male	5,470 (75)	6,869 (74)
Age, yrs	62 ± 9	62 ± 8
Former smoking	3,582 (49)	4,437 (48)
Current smoking	2,146 (29)	2,197 (24)
Alcohol consumption >0 and <10 U/weeks	3,850 (53)	1,654 (18)
Alcohol consumption >10 U/weeks	2,173 (30)	1,124 (12)
Medical history		
Cerebrovascular disease	2,128 (29)	712 (8)
Coronary heart disease	4,530 (62)	9,322 (100)
Peripheral vascular disease	1,300 (18)	844 (9)
Diabetes mellitus	1,321 (18)	3,829 (41)
Physical examination and laboratory measurements		
Body mass index, kg/m <sup>2</sup>	27 ± 4	31 ± 6
Systolic blood pressure, mm Hg	140 ± 20	130 ± 16
Diastolic blood pressure, mm Hg	81 ± 11	78 ± 9
LDL cholesterol, mmol/l	2.7 (2.1-3.5)	2.1 (1.7-2.8)
C-reactive protein, mg/l	2.0 (0.9-4.4)	4.2 (2.8-7.1)
Creatinine, μmol/l	91 ± 23	86 ± 29
Medication		
Lipid-lowering medication	5,038 (69)	8,711 (93)
Blood pressure-lowering medication	5,549 (76)	7,591 (81)
Antiplatelet therapy	5,652 (78)	8,488 (91)
Anticoagulants	816 (11)	718 (8)

Values are n (%), mean ± SD, or median (25th and 75th percentile).  
CANTOS = Canakinumab Anti-Inflammatory Thrombosis Outcomes Study; LDL = low density lipoprotein;  
UCC-SMART = Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease.

colorectal, and lung cancer (Supplemental Figure 1). C-statistics were 0.61 (95% CI: 0.59 to 0.63), 0.61 (95% CI: 0.57 to 0.66), and 0.74 (95% CI: 0.70 to 0.77), respectively, in the UCC-SMART study population.

External calibration plots in Figures 1A to 1C show reasonable agreement between the predicted and observed 4-year risk for total, colorectal, and lung cancer in the CANTOS study population. The expected to observed ratios of the event of interest and competing event were 1.06 and 0.99 for the total cancer model, 1.16 and 0.85 for the colorectal cancer model, and 0.60 and 0.99 for the lung cancer model, accounting for differences in baseline risk. Assessment of discrimination provided a C-statistic of 0.63 (95% CI: 0.61 to 0.66) for the total cancer model, 0.64 (95% CI: 0.58 to 0.70) for the colorectal cancer model, and 0.74 (95% CI: 0.70 to 0.78) for the lung cancer model in the CANTOS data. Supplemental Figures 2 to 4 show calibration plots and C-statistics for the competing risks and cancer-free survival of the total, colorectal, and lung cancer models. The Brier score for 4-year predictions of total, colorectal, and lung cancer was 0.052; 95% CI: 0.048 to 0.057, 0.008; 95% CI: 0.006 to 0.010, and 0.013; 95% CI: 0.010 to 0.016, respectively.

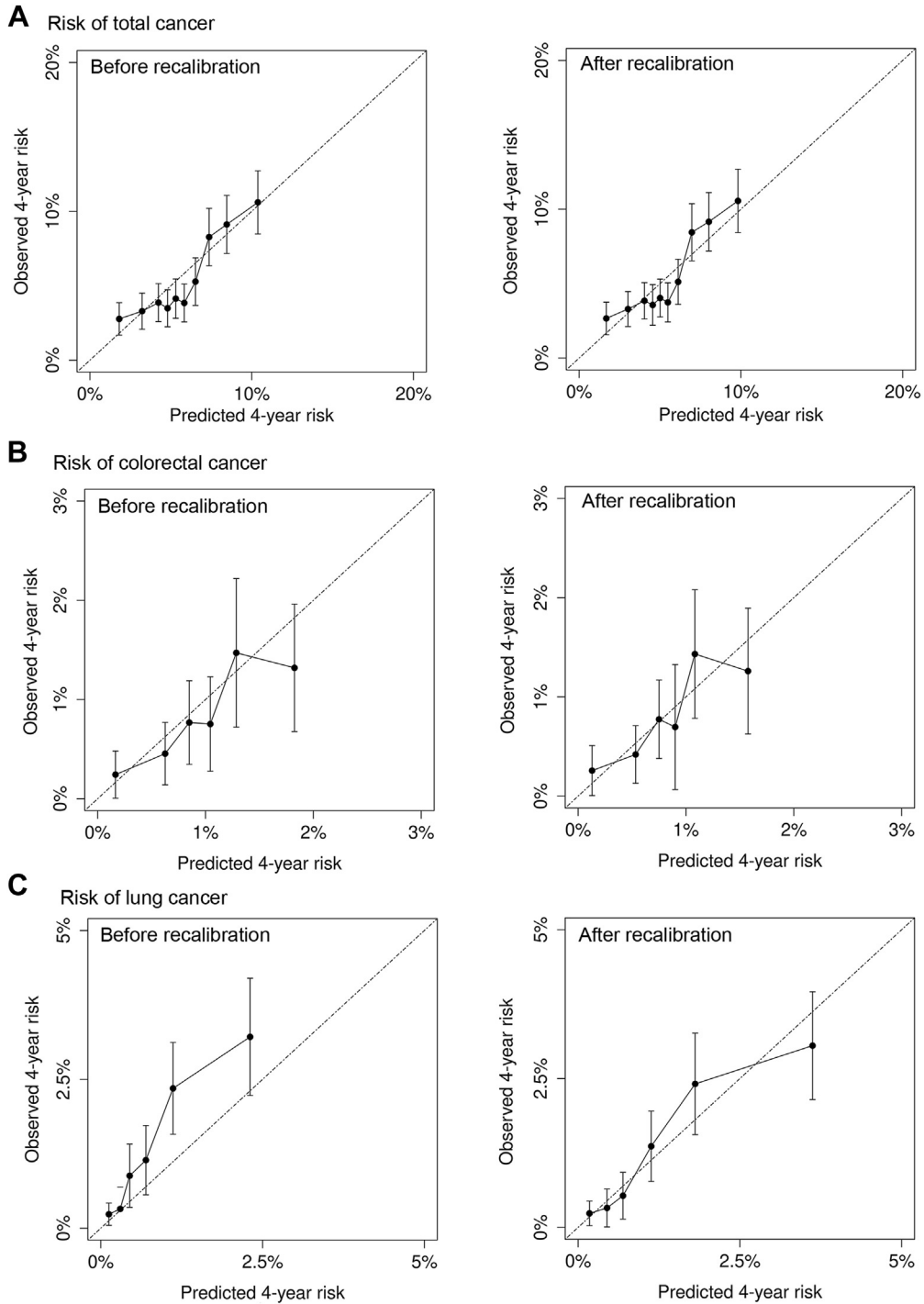
Compared with a simple model with sex and smoking status as the only predictors and with age as the underlying time scale, the full model had a better fit according to the likelihood ratio tests for total and lung cancer ( $p = 0.005$  and  $p < 0.001$ , respectively). For the colorectal cancer model, the full model did not improve model fit ( $p = 0.174$ ). Although the C-statistics of the simple models in CANTOS were similar or even slightly higher; 0.65; 95% CI: 0.62 to 0.67 for total cancer, 0.65; 95% CI: 0.62 to 0.66 for colorectal cancer, and 0.74; 95% CI: 0.70 to 0.79 for lung cancer, and although calibration was similar for colorectal and lung cancers, calibration was worse for total cancer and for the competing risks (Supplemental Figure 5). As calibration is a more clinically relevant performance measure for risk prediction accuracy than the C-statistic (24), the full model for total cancer was considered superior. As all predictors are needed for estimations of total cancer risk, the advantage of a simple model with a limited number of predictors was no longer relevant, and full models were used for risk predictions of total, colorectal, and lung cancer.

#### PREDICTED 10-YEAR AND LIFETIME RISK OF CANCER.

Median predicted absolute 10-year risks were 13% (range 1% to 31%) for total cancer, 2% (range 0% to 6%) for colorectal cancer, and 2% (range 0% to 24%) for lung cancer in the CANTOS study population. In the UCC-SMART study population, predicted 10-year risks were 16% (range 2% to 33%) for total cancer, 2% (range 0% to 5%) for colorectal cancer, and 2% (range 0% to 20%) for lung cancer. Median predicted absolute lifetime risks were 26% (range 1% to 52%) for total cancer, 4% (range 0% to 13%) for colorectal cancer, and 5% (range 0% to 37%) for lung cancer in the CANTOS study population. In the UCC-SMART study population, median predicted absolute lifetime risks were 35% (range 2% to 59%) for total cancer, 5% (range 0% to 11%) for colorectal cancer, and 7% (range 0% to 32%) for lung cancer. Median predicted 10-year and lifetime risks per age group with a 5-year interval for the UCC-SMART and CANTOS study populations are provided in Supplemental Table 11. The distribution of lifetime risks for total, colorectal and lung cancer for UCC-SMART and CANTOS study populations is shown in Figures 2A to 2C.

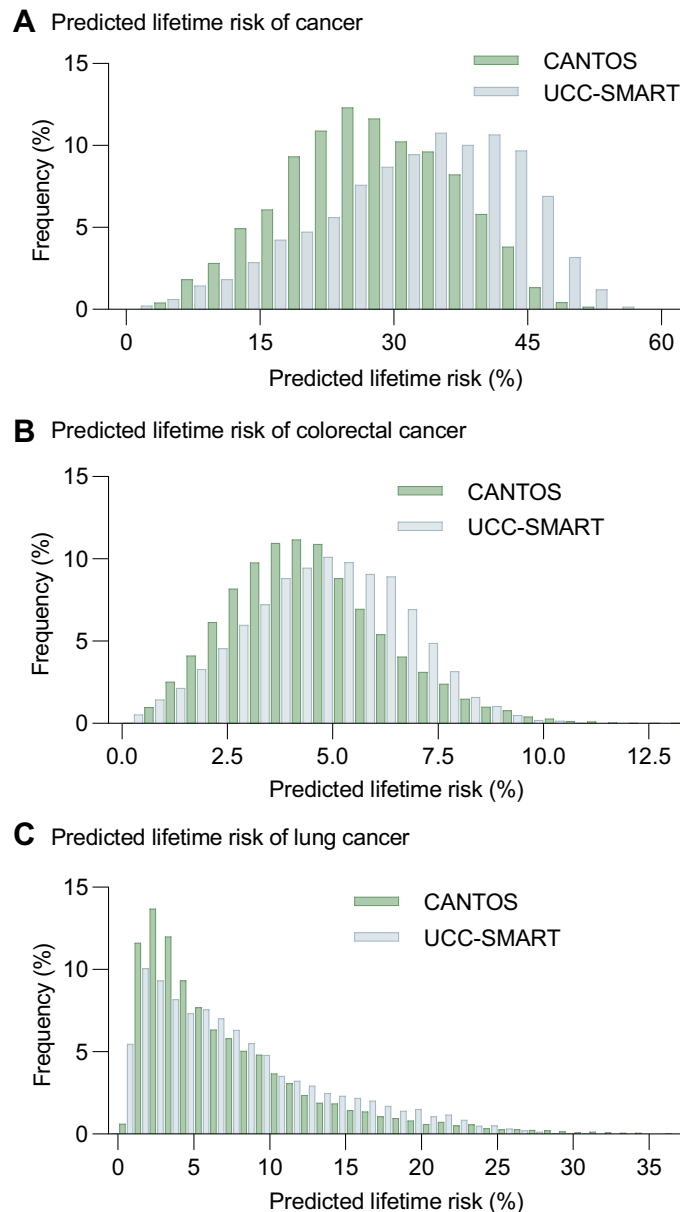
As an example, for a 50-year-old man with average values of UCC-SMART for all other predictors, his predicted lifetime risk of total cancer is 48% if he is a current smoker, 45% if he is a former smoker, and 35% if he has never smoked. The predicted lifetime risks of colorectal cancer for this 50-year old male are 6% (current smoker), 7% (former smoker), and 6% (never smoker). This 50-year old male has a predicted lifetime risk of lung cancer of 18% if he is a smoker,

**FIGURE 1** External Calibration in the CANTOS Trial Population of Cancer Models Before and After Recalibration



Calibration plots are shown of the predicted versus observed 4-year risk of total (A), colorectal (B), and lung cancer (C) in the CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study) study population, before and after recalibration. The study population is divided into quantiles based on the predicted risk, and ordered according to increasing predicted risk. The diagonal dotted line represents perfect calibration.



**FIGURE 2** Distribution of Predicted Lifetime Risk for Total Cancer, Colorectal Cancer, and Lung Cancer

Histograms show the distribution of the predicted lifetime risk for total (A), colorectal (B), and lung cancer (C) in the UCC-SMART (Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease) and CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study) populations. The predicted lifetime risk is the predicted risk of incident cancer up to the age of 90 years. Presented predicted risks are estimated by the full models for total, colorectal, and lung cancer.

10% if he is a former smoker, and 4% if he is a never smoker. In order to facilitate risk predictions in clinical practice, the prediction model is available in the [Supplemental Appendix](#).

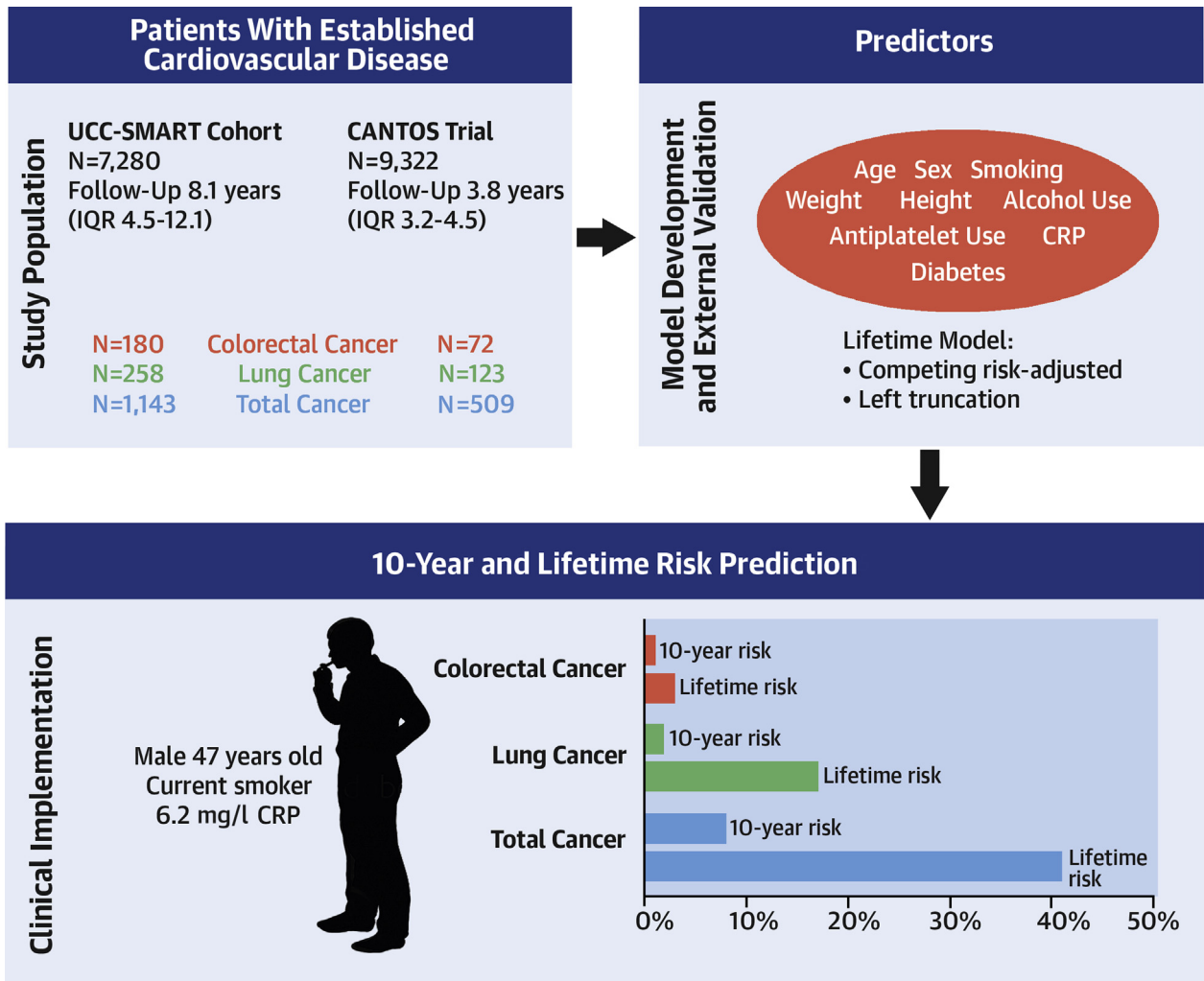
## DISCUSSION

The present study demonstrates that the lifetime and 10-year risk of total, colorectal, and lung cancer can be estimated reasonably well in individual patients with established CVD ([Central Illustration](#)). Although discrimination was moderate with C-statistics of 0.63 to 0.74, calibration of the total, colorectal, and lung cancer models was reasonable. Given the wide distribution of predicted lifetime risks for total cancer and lung cancer ([Figures 2A to 2C](#)), these models can enable the identification of patients at the highest risk for cancer. Innovative and notable aspects of our work include the applicability to patients with established CVD specifically; the relative ease of use with readily clinically available predictors; the prediction of the combined endpoint total cancer; the external validation; and the estimation of lifetime risks with adjustment for competing risks.

Several risk prediction models with clinical predictors have previously been published for specific types of cancer, including lung (5,6), colorectal (6,25-28), and breast (6-9) cancers. None of these models were developed for patients with established CVD specifically, even though these patients are at higher risk for total and lung cancer compared with the general population, with standardized incidence ratios of 1.19 (95% CI: 1.10 to 1.29) for total cancer and 1.56 (95% CI: 1.31 to 1.83) for lung cancer (1), due to similar risk factors for CVD and cancer (2). Furthermore, the endpoint total cancer has a different distribution of cancer types in patients with established CVD (1), and patients with established CVD are at higher risk for the competing risks (i.e., dying from CVD) compared with the general population (29), emphasizing the need for a prediction model in patients with established CVD specifically. It has even been hypothesized that CVD itself influences cancer development, for example through cardiac excreted factors in heart failure (30,31), potentially leading to a higher baseline risk independent of traditional risk factors. Even though cancer is a very heterogeneous disease and prognoses are divergent for the various cancer types, from a patient's perspective, risk of any cancer is relevant, with respect to the potential mortality and morbidity associated with the malignancy, frequent hospital visits, demanding treatments (32), and psychological distress (33,34). Furthermore, in patients with CVD, specific cancer types are more common, including cancers of the respiratory tract (1), leading to restricted variation in cancer types.

Our cancer prediction models performed reasonably well, and calibration plots before and after

**CENTRAL ILLUSTRATION** Prediction of Cancer Risk in Patients With Established Cardiovascular Disease



van 't Klooster, C.C. et al. *J Am Coll Cardiol CardioOnc.* 2020;2(3):400-10.

Using data from the UCC-SMART (Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease) cohort, 3 lifetime, competing risk-adjusted prediction models were developed to predict the risk of total, colorectal, and lung cancer with easily clinically available predictors in patients with established cardiovascular disease. Models were externally validated in the CANTOS (Canakinumab Anti-Inflammatory Thrombosis Outcomes Study) data. The models allow for 10-year and lifetime risk predictions. CRP = C-reactive protein; IQR = interquartile range.

recalibration were similar. Only lung cancer risk was slightly underestimated in the CANTOS population before recalibration, probably due to variations in smoking habits, or genetic factors causing a higher baseline risk. The higher discriminative power of the lung cancer model (C-statistic 0.74) compared with the total and colorectal cancer models (C-statistics 0.63 and 0.64, respectively), is possibly due to the strong relation between the predictor smoking status and lung cancer. For the prediction of lung and

colorectal cancer, a simple model with just age, sex, and smoking status could be sufficient; however, for total cancer and the competing risks, the full model was necessary to achieve the most accurate predictions. For lung cancer, even though the calibration plot showed a 4-year risk of  $\pm 3\%$  in the highest risk group, the model allowed for a widespread lifetime risk distribution, assigning lifetime risks up to 37% to a small proportion of patients. As young patients generally have a low 10-year risk of cancer, despite



high-risk factor levels, lifetime risk predictions might provide more accurate estimations of their “true” risk. The lifetime risk of cancer estimated by the total cancer model ranged from 1% to 52%, enabling identification of patients at the highest risk. Median predicted risks for total cancer were higher in the UCC-SMART study population, corresponding with a higher observed incidence rate for total cancer (1.97 vs. 1.48 per 100 person-years), most likely due to more current smokers in UCC-SMART compared with CANTOS (29% vs. 24%). The distribution of colorectal cancer risk predictions is slightly limited, possibly partly due to absence of family history of colorectal cancer as a predictor in the model, and this model might be less appropriate for selecting patients at very high risk for colorectal cancer.

C-reactive protein was included in the risk prediction models based on previous observational research showing a relation between CRP and incident (lung) cancer (3,18,19), and based on results from the CANTOS trial demonstrating that lowering inflammation with an IL-1 $\beta$  inhibitor lowered the incidence of lung cancer and lung cancer mortality (15). Implementing CRP as a marker of low-grade inflammation in risk scores for determining cancer risk could lead to more accurate predictions. In current models for total, colorectal, and lung cancer, CRP improved model fit based on the AIC. Previous research has shown that CRP improved discrimination in a prediction model for lung cancer in the general population, but only for diagnoses within the first 2 years after measuring CRP (18). In the current models for total and lung cancer, an interaction with age resulted in a higher coefficient of CRP with increasing age, potentially representing a higher predictive value of CRP closer to cancer diagnosis.

There are multiple potential applications of this work, which each require further study. Personalized risk assessment is considered informative and motivational by patients (35), and effective risk communication can lead to changes in behavior (36). Although observed effects of personalized risk communication on healthy behavior changes have been small and evidence is inconsistent (37), effects are dependent on risk information (36). Lifetime risk predictions for cancer, especially in patients at a younger age, could potentially aid in discussions on the importance of healthy lifestyle habits, including smoking cessation. Future prospective studies are needed to evaluate lifestyle improvements and clinical outcomes in patients at high risk for cancer identified by these current models. Moreover, we hypothesize that these models could be used to further inform screening. Results from a recent lung

cancer screening trial (NELSON [Nederlands-Leuvens Longkanker Screenings Onderzoek]) showed that screening for lung cancer could reduce lung cancer mortality in men (cumulative rate ratio for death from lung cancer at 10 years of 0.76; 95% CI: 0.61 to 0.94) (38). The NELSON trial included 50- to 74-year-old current or former smokers who had smoked more than 15 cigarettes a day for more than 25 years or more than 10 cigarettes a day for more than 30 years, and showed a 10-year risk for lung cancer of approximately 6% in the screening group (incidence rate of 5.58 cases per 1,000 person-years) (38). Similarly, it could be hypothesized that patients with stable CVD with a high 10-year predicted risk of lung cancer may benefit from screening computed tomography imaging of the chest. A predicted 10-year lung cancer risk of 6% (close to the 90th percentile in CANTOS) that corresponds to the observed risk in the NELSON study, could potentially be used as one threshold. In addition, application of the predicted lung cancer risk could be used to inform thresholds for targeted diagnostics in patients with early symptoms and high predicted 10-year risks, potentially leading to earlier detection and treatment of cancer.

**STUDY STRENGTHS AND LIMITATIONS.** Strengths of the present study include the large study populations for both the derivation and external validation of the cancer risk prediction models. Another important strength is the competing risk-adjusted analyses, preventing overestimation of the event of interest, especially in a population of patients with established CVD. Furthermore, by using age as the underlying time scale in the models, predictions are not limited by follow-up time in the derivation cohort and lifetime predictions are enabled. Last, the prediction model is available in the [Supplemental Appendix](#). Limitations, however, should be considered. These include the smaller number of lung cancer and colorectal cancer in the development and validation study populations. Furthermore, external validation in the CANTOS trial could be performed only up to 4 years due to limited length of follow-up, although internal validation of 10-year predictions in UCC-SMART showed good calibration. Previous studies have shown that lifetime predictions based on the current methodologies provide adequate estimates for up to at least 17 years (10), and the advantage of CANTOS is the large number of patients with CVD and detailed information on incident cancer. C-statistics for the total cancer, colorectal cancer, and lung cancer models were moderate (0.62 to 0.74), comparable to previous cancer risk predictions models (5,7,25) and recurrent CVD risk prediction models in patients with

established vascular disease (24,39,40). However, evaluation of discrimination with the C-statistic is not optimal in assessing performance of risk prediction models. Calibration is a more clinically relevant performance measure for risk prediction accuracy (24), and calibration of the total, colorectal, and lung cancer predictions models in the CANTOS trial population were all reasonable. Although patients were included in stable phase after a qualifying cardiovascular event, patients potentially changed lifestyle habits, such as smoking, during follow-up, and the single baseline measurement might not reflect such time varying covariates. Last, several potentially important predictors, including level of education, socioeconomic status, race, and family history of cancer were unavailable in the derivation cohort and could not be included in the prediction models, possibly limiting model performance.

## CONCLUSIONS

Lifetime and 10-year risk of total cancer, colorectal cancer, and lung cancer can be estimated reasonably well with easy clinically available predictors in patients with established CVD. The wide distribution of predicted lifetime risks for total and lung cancer enables identification of patients at the highest risk for cancer. With additional study, the lifetime total and lung cancer models could be used in clinical practice to further promote healthy lifestyle changes, and application of these models, particularly the 10-year lung cancer risk model, could potentially lower thresholds for targeted diagnostics and screening.

**ACKNOWLEDGMENTS** We thank the registration team of the Netherlands Comprehensive Cancer Organisation (IKNL) for the collection of data for the Netherlands Cancer Registry as well as IKNL staff for scientific advice. We gratefully acknowledge the contribution of the research nurses: R. van Petersen (data-manager); B. van Dinther (study manager); and

the members of the Utrecht Cardiovascular Cohort-Second Manifestations of ARterial disease-study group (UCC-SMART-study group): F.W. Asselbergs and H.M. Nathoe, Department of Cardiology; G.J. de Borst, Department of Vascular Surgery; M.L. Bots and M.I. Geerlings, Department of Epidemiology; M.H. Emmelot, Department of Geriatrics; P.A. de Jong and T. Leiner, Department of Radiology; A.T. Lely, Department of Obstetrics/Gynaecology; N.P. van der Kaaij, Department of Cardiothoracic Surgery; L.J. Kappelle and Y. Ruigrok, Department of Neurology; M.C. Verhaar, Department of Nephrology, F.L.J. Visseren (chair) and J. Westerink, Department of Vascular Medicine, University Medical Center Utrecht and Utrecht University.

**ADDRESS FOR CORRESPONDENCE:** Dr. Frank L.J. Visseren, Department of Vascular Medicine, University Medical Centre Utrecht, P.O. Box 85500, 3508 GA Utrecht, the Netherlands. E-mail: [f.l.j.visseren@umcutrecht.nl](mailto:f.l.j.visseren@umcutrecht.nl). Twitter: [@TCilie](https://twitter.com/TCilie).

## PERSPECTIVES

**COMPETENCY IN PATIENT CARE:** Patients with established CVD are at higher risk for cancer compared with the general population, due to similar risk factors, such as smoking and obesity. The currently developed and externally validated prediction models estimate the 10-year and the lifetime risks for total, colorectal, and lung cancer in patients with established CVD.

**TRANSLATIONAL OUTLOOK:** Careful consideration of cancer risks in CVD clinical practice could lead to improved patient care by increasing patients' efforts to improve lifestyle habits, such as smoking cessation, and potentially by lowering thresholds for targeted diagnostics and screening. Prospective studies are needed to evaluate lifestyle improvements and clinical outcomes in patients at high risk for cancer identified by these models.

## REFERENCES

1. van Kruijsdijk RC, van der Graaf Y, Peeters PH, Visseren FL, Second Manifestations of ARterial disease (SMART) Study Group. Cancer risk in patients with manifest vascular disease: effects of smoking, obesity, and metabolic syndrome. *Cancer Epidemiol Biomarkers Prev* 2013;22:1267-77.
2. Blaes A, Prizment A, Koene RJ, Konety S. Cardio-oncology related to heart failure: common risk factors between cancer and cardiovascular disease. *Heart Fail Clin* 2017;13:367-80.
3. Van't Klooster CC, Ridker PM, Hjortnaes J, et al. The relation between systemic inflammation and incident cancer in patients with stable cardiovascular disease: a cohort study. *Eur Heart J* 2019;40:3901-9.
4. Dagenais GR, Leong DP, Rangarajan S, et al. Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. *Lancet* 2020;395:785-94.
5. Katki HA, Kovalchik SA, Petito LC, et al. Implications of nine risk prediction models for selecting ever-smokers for computed tomography lung cancer screening. *Ann Intern Med* 2018;169:10-9.
6. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open* 2015;5:e007825.
7. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. *Breast Cancer Res Treat* 2012;132:365-77.

8. Cintolo-Gonzalez JA, Braun D, Blackford AL, et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast Cancer Res Treat* 2017;164:263-84.
9. Maas P, Barrdahl M, Joshi AD, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol* 2016;2:1295-302.
10. Dorresteijn JA, Kaasenbrood L, Cook NR, et al. How to translate clinical trial results into gain in healthy life expectancy for individual patients. *BMJ* 2016;352:i1548.
11. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. *BMJ* 2010;341:c6624.
12. Ridker PM, Everett BM, Thuren T, et al. Anti-inflammatory therapy with canakinumab for atherosclerotic disease. *N Engl J Med* 2017;377:1119-31.
13. Ridker PM, Thuren T, Zalewski A, Libby P. Interleukin-1beta inhibition and the prevention of recurrent cardiovascular events: rationale and design of the Canakinumab Anti-inflammatory Thrombosis Outcomes Study (CANTOS). *Am Heart J* 2011;162:597-605.
14. Simons PC, Algra A, van de Laak MF, Grobbee DE, van der Graaf Y. Second manifestations of ARterial disease (SMART) study: rationale and design. *Eur J Epidemiol* 1999;15:773-81.
15. Ridker PM, MacFadyen JG, Thuren T, et al. Effect of interleukin-1beta inhibition with canakinumab on incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *Lancet* 2017;390:1833-42.
16. World Health Organization International Agency for Research on Cancer. Available at: <http://gco.iarc.fr/today/data/factsheets/cancers/39-All-cancers-fact-sheet.pdf>. Accessed October 2019.
17. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation and updating*. New York, NY: Springer, 2009.
18. Muller DC, Larose TL, Hodge A, et al. Circulating high sensitivity C reactive protein concentrations and risk of lung cancer: nested case-control study within Lung Cancer Cohort Consortium. *BMJ* 2019;364:k4981.
19. Allin KH, Bojesen SE, Nordestgaard BG. Inflammatory biomarkers and risk of cancer in 84,000 individuals from the general population. *Int J Cancer* 2016;139:1493-500.
20. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation* 2016;133:601-9.
21. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;94:496-509.
22. Beiser A, D'Agostino RB Sr., Seshadri S, Sullivan LM, Wolf PA. Computing estimates of incidence, including lifetime risk: Alzheimer's disease in the Framingham Study. The Practical Incidence Estimators (PIE) macro. *Stat Med* 2000;19:1495-522.
23. de Vries TI, Eikelboom JW, Bosch J, et al. Estimating individual lifetime benefit and bleeding risk of adding rivaroxaban to aspirin for patients with stable cardiovascular disease: results from the COMPASS trial. *Eur Heart J* 2019;40:3771-8.
24. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115:928-35.
25. Usher-Smith JA, Harshfield A, Saunders CL, et al. External validation of risk prediction models for incident colorectal cancer using UK Biobank. *Br J Cancer* 2018;118:750-9.
26. Steffen A, MacInnis RJ, Joshy G, Giles GG, Banks E, Roder D. Development and validation of a risk score predicting risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2014;23:2543-52.
27. Shin A, Joo J, Yang HR, et al. Risk prediction model for colorectal cancer: National Health Insurance Corporation study, Korea. *PLoS One* 2014;9:e88079.
28. Wei EK, Colditz GA, Giovannucci EL, et al. A comprehensive model of colorectal cancer by risk factor status and subsite using data from the Nurses' Health Study. *Am J Epidemiol* 2017;185:224-37.
29. van Kruijsdijk RC, van der Graaf Y, Koffijberg H, et al. Cause-specific mortality and years of life lost in patients with different manifestations of vascular disease. *Eur J Prev Cardiol* 2016;23:160-9.
30. Hasin T, Gerber Y, Weston SA, et al. Heart failure after myocardial infarction is associated with increased risk of cancer. *J Am Coll Cardiol* 2016;68:265-71.
31. Meijers WC, Maglione M, Bakker SJL, et al. Heart failure stimulates tumor growth by circulating factors. *Circulation* 2018;138:678-91.
32. Hurria A, Togawa K, Mohile SG, et al. Predicting chemotherapy toxicity in older adults with cancer: a prospective multicenter study. *J Clin Oncol* 2011;29:3457-65.
33. Zabora J, BrintzenhofeSzoc K, Curbow B, Hooker C, Piantadosi S. The prevalence of psychological distress by cancer site. *Psychooncology* 2001;10:19-28.
34. Aass N, Fossa SD, Dahl AA, Moe TJ. Prevalence of anxiety and depression in cancer patients seen at the Norwegian Radium Hospital. *Eur J Cancer* 1997;33:1597-604.
35. Manuel DG, Abdulaziz KE, Perez R, Beach S, Bennett C. Personalized risk communication for personalized risk assessment: real world assessment of knowledge and motivation for six mortality risk measures from an online life expectancy calculator. *Inform Health Soc Care* 2018;43:42-55.
36. Ahmed H, Naik G, Willoughby H, Edwards AG. Communicating risk. *BMJ* 2012;344:e3996.
37. French DP, Cameron E, Benton JS, Deaton C, Harvie M. Can communicating personalised disease risk promote healthy behaviour change? A systematic review of systematic reviews. *Ann Behav Med* 2017;51:718-29.
38. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med* 2020;382:503-13.
39. Dorresteijn JA, Visseren FL, Wassink AM, et al. Development and validation of a prediction rule for recurrent vascular events based on a cohort study of patients with arterial disease: the SMART risk score. *Heart* 2013;99:866-72.
40. Wilson PW, D'Agostino R Sr., Bhatt DL, et al. An international model to predict recurrent cardiovascular disease. *Am J Med* 2012;125:695-703.e1.

---

**KEY WORDS** colorectal cancer, lung cancer, risk prediction

---

**APPENDIX** For supplemental tables and figures, please see the online version of this paper.