integraal kankercentrum Nederland

The future use of radiotherapy in the treatment of NSCLC and SCLC: a prediction

Jan-Willem Bauhuis

Master thesis Health Sciences

Track Optimization of Healthcare Processes

Faculty of Science and Technology (TNW)

University of Twente

First supervisor: Dr. C.G.M Groothuis-Oudshoorn, University of Twente

Second supervisor: Prof.Dr. S. Siesling, University of Twente, IKNL

Supervisor IKNL: J. Evers, MSc, IKNL

January 28, 2022

Abstract

Background: In the past, problems regarding the capacity of radiotherapy occurred in the Netherlands. This resulted in the opening of various satellite locations of existing radiotherapy facilities between 2008 and 2014. Over the last decade, the annual number of stage I-III NSCLC and SCLC diagnoses has increased. Also, the use of (stereotactic body) radiotherapy became more prominent in lung cancer treatment. This could lead to capacity problems again in the future. Therefore, accurate predictions are needed regarding the use of radiotherapy in 2030 for stage I-III NSCLC and SCLC.

Methods: A framework of four steps was created. First, the annual number of stage I-III NSCLC and SCLC diagnoses in the period 2030-2034 was predicted. Second, a prediction model for the use of radiotherapy was made by using a random forest model. Third, a synthetic patient cohort was created for 2030 based on the results of step 1. Fourth, scenario analyses were performed by modifying the variables according to the expected results and creating new synthetic cohort. One example scenario was tested in which an additional radiotherapy facility was opened.

Results: The annual number of stage I-III NSCLC and SCLC diagnoses is predicted to increase to 7640 patients. The prediction model provided an accuracy of 74.5%, a sensitivity of 82%, and a specificity of 65.8%. 3930 patients are expected to receive radiotherapy in 2030. The example scenario provided similar results in terms of the number of predicted patients receiving radiotherapy.

Conclusion: The number of patients receiving radiotherapy increased from 3288 patients in 2019 to 3930 predicted patients in 2030. Other scenarios that affect important predictors, e.g. stage or tumor grade, will most likely affect the use of radiotherapy. The framework performed in this study allows to work out these scenarios by altering the variables accordingly to the expected results.

Table of contents

Introduction

Lung cancer significantly impacts the capacity of oncological care, as it is one of the most common types of cancer in western countries. [1] In the Netherlands, the incidence of lung cancer was approximately 14.000 in 2019 and most patients were aged 75-85 years at the time of diagnosis. Over the last decades, the incidence rates among men have been decreasing from 72.6 (per 100.000 person-years) in 1989 to 38.6 in 2019. However, incidence rates for women have been increasing from 12.0 (per 100.000 person-years) in 1989 to 33.1 in 2019 [2]. Overall, the incidence is still slightly increasing. [3]

Lung cancer diagnoses can be split into small cell lung cancer (SCLC) (~15%) and non-small cell lung cancer (NSCLC) (~85%). Treatment strategies differ by subtype and TNM stage. [4] Also, the performance status of the patient plays a role in the choice of treatment, as well as other patient and environmental factors. [5, 6] In stage I and II NSCLC, surgical resection traditionally has been the preferred choice of treatment. However, stereotactic body radiotherapy (SBRT) provides an alternative curative treatment option for stage I and its use rapidly increased in inoperable (mainly elderly) patients in the Netherlands. [7] Later, a shift from surgery to SBRT was seen in operable patients and nowadays radiotherapy is more often used in the treatment of stage I than surgery [5]. In stage II NSCLC, surgery remained the most delivered treatment modality, applied to more than half of the cases in 2018. [5] While radiotherapy is used in one-third of the stage II NSCLC patients. In unresectable stage III NSCLC, treatment options include concurrent or sequential chemoradiation, of which concurrent is preferred. For the period 2008-2018, a marginal increase in the combined use of chemotherapy and radiotherapy was demonstrated in the Netherlands. Only one-fifth of patients with stage III received concurrent chemoradiation in 2018. [5] In selected stage III cases, surgery is a treatment option, potentially after successful induction therapy. [8]

In stage I SCLC, surgery has shown promising results in selected tumors (T1-2N0) and may nowadays be considered a treatment option. In the period 2008-2018, surgery was increasingly used in stage I SCLC in the Netherlands. [6] In the rest of the stage I tumors and in stage II and III SCLC, treatment with concurrent chemoradiation is the preferred option. Over the last decade, an increase of about 10% could be seen in the overall use of radiotherapy, excluding prophylactic cranial irradiation (PCI), in stage III SCLC. [6] Also, concurrent chemoradiation therapy shows an upward trend in the same period. On the other hand, the use of sequential chemoradiation therapy remains more or less the same. The use of PCI shows a decreasing trend among all stages of SCLC. An explanation for this trend seems to be the publication of the Takashi trial in 2017, which concluded that prophylactic cranial

irradiation did not result in longer overall survival in patients with extensive disease small-cell lung cancer. [9]

Considering the trends in treatment applied over the last decades [5, 6], it can be concluded that there is an overall increase in the use of radiotherapy in lung cancer. In addition to these figures, the use of radiotherapy is expected to increase even further as a result of demographical changes. The number of elderly [10], and also the number of elderly that are diagnosed with lung cancer is increasing [3]. The demand for radiotherapy could further increase if a national lung cancer screening program is implemented, as screening is expected to change the distribution of stages, which results in a larger proportion of early-stage lung cancer diagnoses eligible for stereotactic body radiotherapy (SBRT) [6].

The demand and capacity for radiotherapy have been growing over the last decades. In 1990, about one-third of all radiotherapy departments in the Netherlands had waiting lists causing a delay in treatment. [11] In the period 1998-2010, the number of external beam treatments furthermore increased with an annual percentage of 3.5-4.0%. [12] To meet this increasing demand, the Ministry of Health, Welfare and Sport stressed the need to expand the number of accelerators in 2009. [13] This resulted in the opening of various satellite locations of existing radiotherapy facilities between 2008 and 2014 [14]. To ensure that future demand for radiotherapy can be met, it is important to estimate the required radiotherapy capacity. Therefore, the aim of this study is to create a framework to predict the use of radiotherapy for patients with NSCLC or SCLC up to 2030 by means of an estimation regarding the number of stage I-III NSCLC or SCLC patients receiving radiotherapy in 2030. Furthermore, an example scenario will be performed to demonstrate the application of this framework.

Methods

Patients

Patients diagnosed and registered with SCLC or NSCLC stage I, II, or III, between 2017 and 2019 in the Netherlands were selected from the Netherlands Cancer Registry (NCR). The NCR is a nationwide population-based registry containing information regarding the patient, tumor, and the delivered first-line treatment of all diagnosed cancer patients. [15] The information on the medical records from the Dutch hospitals is extracted by trained registrars. NSCLC stage IV is excluded from this study, as in most cases it is not considered curable. Palliative care is the preferred choice of treatment.

Data analyses

To predict the use of radiotherapy in lung cancer patients diagnosed in the Netherlands in 2030, the following four steps were subsequently performed: (1) the annual number of stage I-III NSCLC and SCLC patients diagnosed in the period 2030-2034 was predicted, (2) predictors for the use of radiotherapy in 20017-2019 identified, (3) a synthetic patient cohort for 2030 was created, (4) scenario analyses were performed.

Predicting the number of lung cancer patients in 2030

The Age-Period-Cohort (APC) model developed by NORDPRED was used to predict the incidence [16]. Sensitivity analyses comparing the predicted number of patients with lung cancer diagnosed in 2019 using the APC model and the short-base model showed the best results for the APC model (Supplementary Document 1).

The APC model uses the following input: population's age distribution, sex, the population's cancer incidence distribution, and the year of observation. The results of the model are presented by age groups and for all age groups combined. Graphical figures of the results are also provided. [17]

Identifying predictors for the use of radiotherapy

A classification model that states whether the patient received any form of radiotherapy was built using the random forest model. A random forest model is a tree-based ensemble with each tree depending on a collection of variables that are randomly selected [18]. The original data set, containing 20,289 observations and 17 variables (Supplementary Document 1), was randomly divided into a training set (80%) and a test set (20%). The training set was used to build the random forest model. The test set was used to generate an accuracy estimate. Variables that were used for the random forest model, were selected based on the possible influence on treatment choice. The goal was to create a prediction function to predict the probability of receiving any form of radiotherapy as a function of the dependent variable (Y), which in this model was binary: the patient received any form of radiotherapy, either as the primary treatment or as a (neo)adjuvant therapy. Patients were classified based on several variables, such as age, gender, stage, and region.

Creating a synthetic cohort for 2030

A synthetic patient cohort for 2030 was created, based on the predicted sex and age distributions for 2030 in the first step. This cohort is based on the patient cohort of 2018-2019. This was achieved by predicting the values of the variables and their distribution by means of the values in the patient cohort 2018-2019. Implicitly, this assumes that no other changes occur in the distribution of patient characteristics (e.g. stage, region, morphology) until 2030 and that the only changes that occur are demographical. The incidence prediction of the first step resulted in an overview of the total incidence per age category for men and women. These results were used to determine the size of each age category in the synthetic cohort for 2030. The synthetic cohort was used as input for the random forest model, to predict the number of patients receiving radiotherapy in 2030.

Scenario-analyses

The framework provided in this study, allows to test several scenarios. One example scenario was tested. In this example scenario an additional radiotherapy facility was opened. The aim of this scenario was to test the effect of a reduction in travel time on the use of radiotherapy. Over the last decade, 11 additional radiotherapy facilities were added [14]. Therefore, it was a plausible scenario to test. Two different locations were tested. In both locations, most patients had a travel time of 45+ minutes to the nearest radiotherapy facility. The first location was in Zeeland (postal code 4561AB). The second location was in Drenthe (postal code 8381AB). The data set was modified according to the expected result and new synthetic cohorts were created for this scenario.

Statistical software

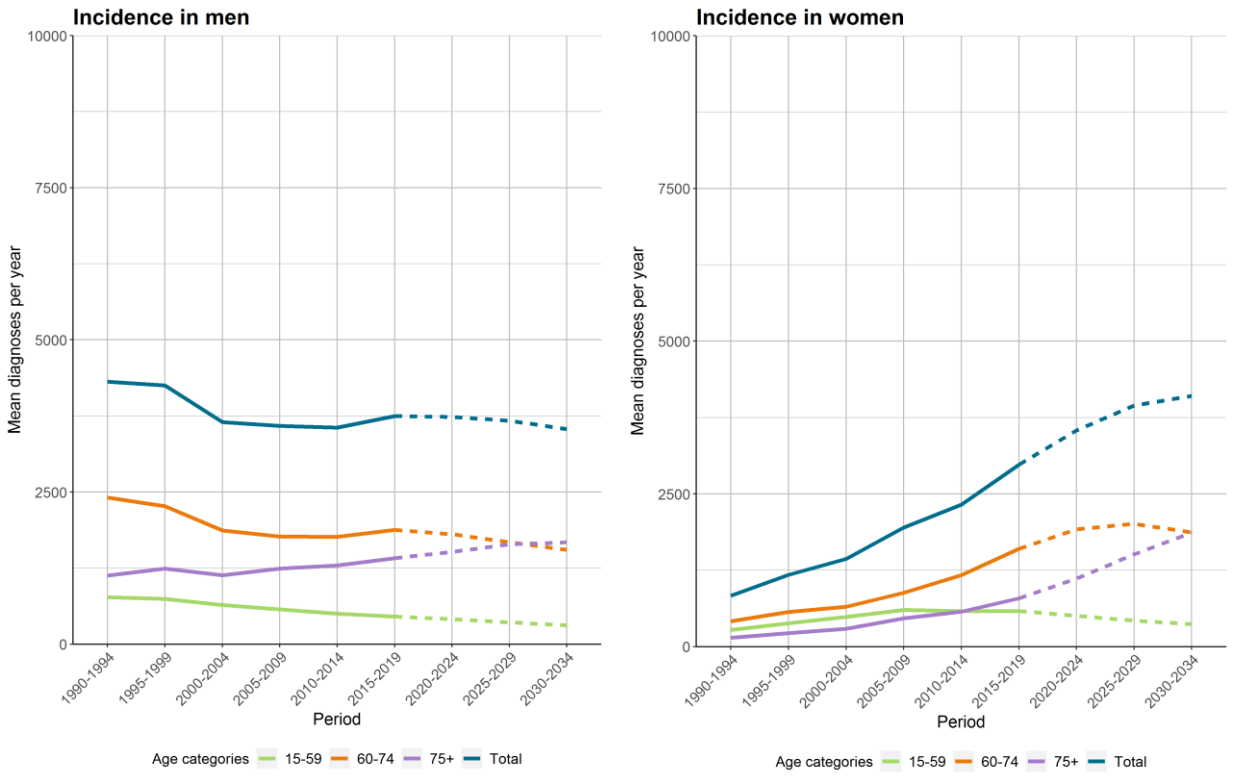
The statistical software used in this study was SAS and RStudio version 4.0.3. SAS was used to prepare the data sets and create the variables. RStudio was used to create the random forest model. The Nordpred package was used in RStudio to predict the number of lung cancer patients in 2030. The Synthpop package was used in RStudio to create the synthetic cohort [19].

Results

NSCLC and SCLC incidence prediction 2030

The prediction results regarding the annual number of patients diagnosed with NSCLC and SCLC stage I-III stratified for men and women are presented in Table 1 and Figure 1. In general, the total annual number of diagnoses of NSCLC and SCLC is expected to increase until the period 2030-2034. For men, the total number of diagnoses per year shows a decreasing trend. This was also observed in age categories ‘15-59’ and ‘60-74’. However, the age category ‘75+’ showed an increasing trend. For women, the total annual number of diagnoses is predicted to increase until the period 2030-2034. In the age category '75 +' an increase is predicted until the period 2030-2034, while in the age category '60-74' the number of diagnoses per year increases until the period 2025-2029, but then starts decreasing. The number of diagnoses per year in the age category '15-59' showed a decreasing trend since the period 2005-2009.

In total, this results in 7640 predicted diagnoses per year in the period 2030-2034. This is a 13.6% increase compared to the period 2015-2019 when the total average number of diagnoses per year was 6728 (Table 1). Stratified for sex, 3537 lung cancer diagnoses for men and 4103 for women are predicted per year in the period 2030-2034.



A. B.

Figure 1: A. Mean observed and predicted NSCLC and SCLC number of diagnoses per year for men. B. Mean observed and predicted NSCLC and SCLC number of diagnoses per year for women.

Table 1
Mean observed (1990-2019) and predicted (2020-2034) NSCLC and SCLC diagnoses per year for men, women, and total.

Period	Men	Women	Total
1990 - 1994	4312	835	5148
1995 – 1999	4253	1178	5431
2000 – 2004	3650	1437	5086
2005 – 2009	3585	1948	5533
2010 – 2014	3558	2325	5883
2015 – 2019	3748	2980	6728
2020 – 2024	3735	3452	7277
2025 – 2029	3673	3947	7620
2030 – 2034	3537	4103	7640

Identifying predictors for the use of radiotherapy

The test set, containing 4027 (20%) observations, was used to test the performance of the random forest model. Due to missing values, 21 observations were excluded from the data set. The confusion matrix of showed an accuracy of 74.5%, 95% CI [0.73-0.76]. Furthermore, the model showed a sensitivity of 82.6%, 95% CI [0.81-0.84] and a specificity of 65.8%, 95% CI [0.64-0.68]. Also, the confusion matrix showed a positive predictive value of 72.3%, 95 CI [0.71-0.74] and a negative predictive value of 77.7%, 95% CI [0.76-0.80]. The variable importance plot provided by the model is shown in figure 2.

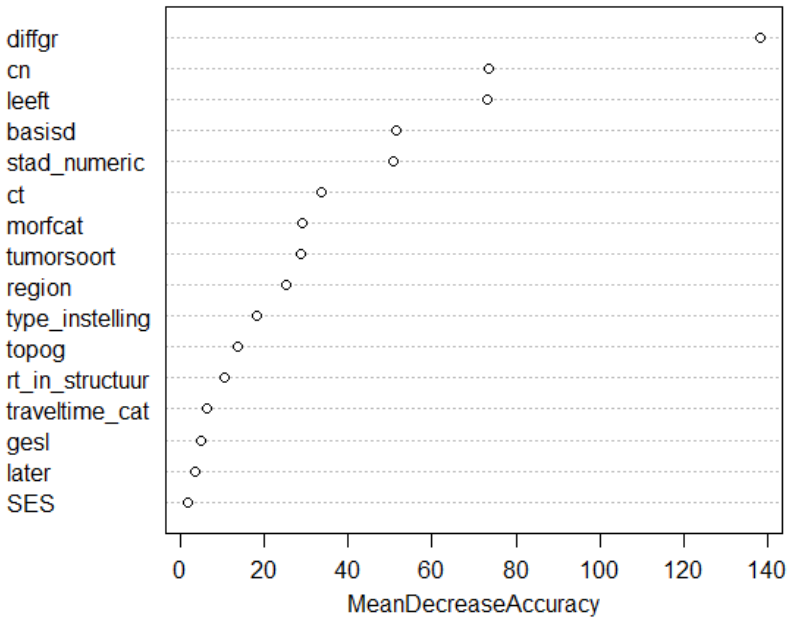


Figure 2: Variable importance plot provided by the random forest model

Creating a synthetic cohort for 2030

The synthetic patient cohort for 2030 was created using data from the patient cohort 2018-2019. The number of synthesized patients was based on the results of the incidence predictions for every age category stratified for men and women. This resulted in a patient cohort for 2030 consisting of 7640 patients. A comparison of the distribution between the synthetic cohort and the cohort 2018-2019 is shown in Table 2. The random forest model was used to calculate the use of radiotherapy for this synthetic cohort. 57 observations were excluded due to missing values. 3930 of the total 7583 patients were predicted (51.8%) to receive radiotherapy in 2030.

Table 2

Comparison of patient characteristics between the cohort 2018-2019 and the synthesized cohort 2030 for patients with NSCLC and SCLC in the Netherlands.

		Cohort 2018-2019		Synthetic cohort 2030	
		N = 13,622		N = 7640	
		N	(%)	N	(%)
Sex	Men	7389	54.2	3537	46.3

	Women	6233	45.8	4103	53.7
Age	15-59	1995	14.6	680	8.9
	60-74	7095	52.1	3418	44.7
	75+	4532	33.3	3542	46.4
Stage	I	5732	42.1	3234	42.3
	II	2132	15.7	1259	16.5
	III	5758	42.3	3147	41.2
cT	0	52	0.4	30	0.4
	1	35	0.3	20	0.3
	1A	703	5.2	383	5.0
	1B	2828	20.8	1592	20.8
	1C	1977	14.5	1104	14.5
	1MI	12	0.1	7	0.1
	2A	1783	13.1	1057	13.8
	2B	860	6.3	486	6.4
	3	2178	16.0	1251	16.4
	4	2996	22.0	1626	21.3
	X	198	1.5	84	1.1
cN	0	8142	59.8	4636	60.7
	1	1009	7.4	584	7.6
	2	2739	20.1	1465	19.2
	3	1476	10.8	786	10.3
	X	256	1.9	169	2.2
Tumor type	NSCLC	9544	70.1	5128	67.1
	SCLC	959	7.0	504	6.6
	Other	3119	22.9	2008	26.3
Region	Zuid-Holland	2736	20.1	1529	20.0
	Noord-Holland	2157	15.8	1228	16.1
	Noord-Brabant	2297	16.9	1247	16.3
	Gelderland	1729	12.7	974	12.7
	Limburg	1104	8.1	614	8.0
	Utrecht	859	6.3	507	6.6
	Groningen	509	3.7	296	3.9
	Zeeland	259	1.9	131	1.7
	Flevoland	291	2.1	177	2.3
	Drenthe	422	3.1	268	3.5
	Friesland	413	3.0	237	3.1
	Overijssel	846	6.2	432	5.7
SES	1	2037	15.0	1166	15.3
	2	1659	12.2	954	12.5
	3	1504	11.0	869	11.4
	4	1371	10.1	733	9.6
	5	1322	9.7	699	9.1
	6	1284	9.4	715	9.4
	7	1188	8.7	665	8.7
	8	1143	8.4	654	8.6
	9	1073	7.9	613	8.0
	10	901	6.6	515	6.7
	Missing	140	1.0	57	0.7
Travel time	0-15 min	5660	41.6	3189	41.7
	15-30 min	6955	51.1	3904	51.1

	30-45 min	949	7.0	518	6.8
	45+ min	57	0.4	29	0.4
	Missing	1	0.0	0	0.0
Hospital type	Academic	1718	12.6	873	11.4
	Top Clinical hospitals	4839	35.5	2719	35.6
	General	7064	51.9	4048	53.0
	Missing	1	0.0	0	0.0
Radiotherapy in house	Yes	3012	22.1	1619	21.2
	No	10610	77.9	6021	78.8
Morphology category	NSCLC Squamous carcinoma	3351	24.6	1787	23.4
	NSCLC Adenocarcinoma	4336	31.8	2376	31.1
	NSCLC Large/Other	4976	36.5	2973	38.9
	SCLC	959	7.0	504	6.6
Topography	C340	598	4.4	348	4.6
	C341	7840	57.6	4394	57.5
	C342	550	4.0	307	4.0
	C343	3985	29.3	2252	29.5
	C348	228	1.7	119	1.6
	C349	421	3.1	220	2.9
Lateralization	1	5604	41.1	3137	41.1
	2	7895	58.0	4444	58.2
	3	1	0.0	0	0.0
	4	3	0.0	4	0.1
	X	119	0.9	55	0.7
Tumor grade	1	92	0.7	47	0.6
	2	2451	18.0	1293	16.9
	3	1040	7.6	538	7.0
	4	85	0.6	48	0.6
	9	9954	73.1	5714	74.8
Basis for diagnosis	1	3	0.0	3	0.0
	2	3103	22.8	1998	26.2
	4	6	0.0	5	0.1
	5	2134	15.7	1091	14.3
	6	632	4.6	310	4.1
	7	7744	56.8	4233	55.4

Scenario analyses

Adding an additional radiotherapy facility in Zeeland resulted a travel time reduction for 40 patients. Adding this facility in Drenthe resulted a travel time reduction for 26 patients. The shift in the distribution of travel time was summarized in table 3. In the baseline scenario, the mean travel time was 16 minutes and 52 seconds, with a median of 16 minutes (p25-p75: 10-22). The median was did not change by adding an additional radiotherapy facility. Mean travel time also showed similar results for both locations compared to the baseline situation. Adding an additional facility in Zeeland resulted in

3906 predicted patients receiving radiotherapy in 2030. Adding the facility in Drenthe resulted in 3898 patients.

Table 3.

Comparison of the distribution in travel time as a result of adding an additional radiotherapy facility.

Travel time category	Baseline situation*	Location Zeeland	Location Drenthe
< 15 minutes	3189	3229 (+40)	3215 (+26)
<30 minutes	3904	3894 (-10)	3872 (-32)
<45 minutes	518	503 (-15)	522 (+4)
≥45 minutes	29	14 (-15)	31 (+2)

* reference situation

Discussion

Based on our incidence predictions, the total number of NSCLC and SCLC diagnoses is expected to increase to 7640 diagnoses on average per year in the period 2030-2034. The number of diagnoses for men shows a decreasing trend until 2030-2034, while for women the number of diagnoses for women is expected to increase until the period 2030-2034. 3930 of the 7583 patients (51.8%) are predicted to receive any type of radiotherapy (excluding PCI) as part of their treatment. This figure is an increase of 19.5% compared to 3288 patients receiving radiotherapy (excluding PCI) in 2019.

Comparing the findings of this study to the literature, similar incidence predictions were made for 2020 in the Netherlands [20]. The predicted number of new cancer patients for 2020 was 66.000 men and 57.000 women. According to the Netherlands Cancer Registry [21], the observed number of cancer patients for men and women in 2020, was 60.725 and 54.322 respectively. The observed number of diagnoses is lower than the prediction. However, this difference could be explained by the COVID-19 pandemic, which resulted in a decrease in cancer diagnoses in 2020. [22] Overall, these incidence predictions proved to be relatively accurate.

With the increase in cancer incidence, the use of radiotherapy as a treatment option also increased. In 2010, 120 linear accelerators were available in the Netherlands. This increased to 131 accelerators in 2016. Furthermore, the number of external beam treatments increased from approximately 54.000 in 2010 to almost 60.000 per year in 2016. [23]. In the treatment of lung cancer,

similar trends in the use of radiotherapy were reported by Evers et al. for the period 2008-2018 [5, 6]. The findings of this study suggest that this increasing trend will continue for the period 2020-2034.

However, multiple plausible scenarios in the future may impact the use of radiotherapy. One scenario was tested as an example, i.e. opening an additional radiotherapy facility in the region Zeeland or Drenthe, showing similar results for the number of predicted patients receiving radiotherapy compared to the current situation. Another scenario that is currently discussed is the implementation of a national screening program [24]. In the Netherlands, a screening program was implemented for breast cancer in 1990, which led to earlier detection and a shift in the distribution of stages [25]. With the framework presented in this study a similar scenario could be simulated for lung cancer by changing the variables according to the expected results of the screening. This shift towards lower stages is expected to affect the use of radiotherapy. Since the use of radiotherapy is preferred over surgery in stage I NSCLC and applied in one-third of the stage II NSCLC patients. Another plausible scenario that could be simulated is a change in smoking behaviour. Although the prediction in this study is too short to see the expected results of a recent change in smoking behaviour. The effect of this change can be simulated by applying a percentage reduction on the incidence prediction equal to the percentage reduction change in smoking behaviour.

Strengths and limitations

One of the main strengths of this study involves the use of a comprehensive and transparent approach to predict the number of patients receiving radiotherapy in 2030, which allows to easily modify the predictions based on novel insights or changes in clinical practice. The framework that includes these four steps allows using possible future scenarios as input for the model, by altering the values of the synthetic cohort according to the expected results.

Also, the Nordpred model used in this study is a modified version of the age-period-cohort model. This model is chosen since the exponential function of the original model produces predictions that grow exponentially over time. This results in overestimated predictions of some cancer types. By using the power function in the Nordpred model (x5 instead of exp(x)) this growth can be reduced. [16]

The predicted incidence rates need to be interpreted with caution because these are based on the assumption that past trends will continue into the future. A sudden change in the future trends

cause over-or underestimation of these projections; however, these changes are difficult to predict. Besides, this always challenges attempts to predict future events. To account for this, a function provided in the APC model by Nordpred is used, in which the predicted trendline can be adjusted to prevent overestimating the incidence. This can be done by cutting the trend line: each year in the five-year period can be trimmed with a percentage relative to the year before. This way expected future effects could be incorporated in the incidence prediction which could not be seen yet in the input data.

Another limitation of the study is that SES scores were only available for patients diagnosed before 2019. For patients diagnosed since 2019, the SES scores were imputed based on the scores from 2018, 2017 or 2016 in the corresponding postal codes. For future research using Performance Status (PS) instead of social-economic status is recommended, since SES will not be available anymore, and PS is probably a better predictor for receiving radiotherapy.

Furthermore, it should be noted that stage IV is excluded from this study, as in most cases it is not considered curable. Palliative care is the preferred choice of treatment. Therefore, the results of some scenarios, such as the implementation of a screening program will most likely not be as realistic. Because any shift in the distribution from stage IV to III cannot be measured.

Conclusions

This study has found that the annual number of diagnoses for stage I-III NSCLC and SCLC is expected to increase until the period 2030-2034. 3930 patients are predicted to receive any form of radiotherapy in 2030, either as the main treatment modality or as a (neo)adjuvant therapy.

This comprehensive and transparent framework consisting of four steps to predict the number of patients receiving radiotherapy in 2030, allows future modifications of the predictions based on novel insights or changes in clinical practice. The impact of these changes is measured by the number of patients that receive radiotherapy in 2030. The results of this study may guide the planning of radiotherapy in stage I-III lung cancer treatment in the future. The framework presented in this study also allows scientists and policymakers to explore the consequences of interventions aimed at the impact and treatment of lung cancer.

For future research, stage IV should be included. This makes the effect and impact of scenarios as realistic as possible. Furthermore, the same predictions should be made for other types of cancer, to predict the total use of radiotherapy and the number of radiotherapy facilities required in 2030.

References:

[1] Cancer. World Health Organization; 2021 [cited 2021Oct25]. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer

[2] Hendriks LEL, Dingemans A-MC, De Ruysscher DKM, Aarts MJ, Barberio L, Cornelissen R, et al. Lung cancer in the Netherlands. Journal of Thoracic Oncology. 2021Mar1;16(3):355–65.

[3] Netherlands Cancer Registry (NCR), IKNL. [cited 2021Oct25] Obtained from https://iknl.nl/nkr-cijfers?fs%7Cepidemiologie_id=506&fs%7Ctumor_id=200&fs%7Cregio_id=530&fs%7Cperiode_id=545%2C546%2C547%2C548%2C549%2C550%2C551%2C552%2C553%2C554%2C555%2C556%2C557%2C558%2C559%2C560%2C561%2C562%2C563%2C564%2C565%2C566%2C567%2C568%2C569%2C570%2C571%2C572%2C544%2C543%2C542%2C541&fs%7Cgeslacht_id=622&fs%7Cleeftijdsgroep_id=655&fs%7Cjaren_na_diagnose_id=665&fs%7Ceenheid_id=681&cs%7Ctype=line&cs%7CxAxis=periode_id&cs%7Cseries=epidemiologie_id&ts%7CrowDimensions=periode_id&ts%7CcolumnDimensions=&lang%7Clanguage=nl

[4] Lababede O, Meziane MA. The eighth edition of TNM staging of lung cancer: Reference Chart and diagrams. The Oncologist. 2018Apr12;23(7):844–8.

[5] Evers J, de Jaeger K, Hendriks LEL, van der Sangen M, Terhaard C, Siesling S, et al. Trends and variations in treatment of stage I–III non-small cell lung cancer from 2008 to 2018: A nationwide population-based study from the Netherlands. Lung Cancer. 2021Mar20;155:103–13.

[6] Evers J, Hendriks LEL, De Jaeger K, Wijsman R, De Ruysscher D, Terhaard C, et al. Trends and variations in the treatment of stage I-III Small Cell Lung Cancer from 2008 to 2019: A nationwide population-based study from the Netherlands. Lung Cancer. 2021Oct27;162:61–70.

[7] Palma D, Visser O, Lagerwaard FJ, Belderbos J, Slotman BJ, Senan S. Impact of introducing stereotactic lung radiotherapy for elderly patients with stage I non–small-cell lung cancer: A population-based time-trend analysis. Journal of Clinical Oncology. 2010;28(35):5153–9.

[8] Dickhoff C, Dahele M, de Langen AJ, Paul MA, Smit EF, Senan S, et al. Population-based patterns of surgical care for Stage IIIA NSCLC in the Netherlands between 2010 and 2013. Journal of Thoracic Oncology. 2016Jan7;11(4):566–72.

[9] Tomassen ML, Aarts MJ, Peters M, van Lindert A, De Ruysscher DKM, Verhoeff JJC, et al. Prophylactic cranial irradiation in patients with small cell lung cancer in the Netherlands: A population-based study. Clinical and Translational Radiation Oncology. 2021Feb11;27:157–63.

[10] Centraal Bureau voor de Statistiek. Ouderen. Centraal Bureau voor de Statistiek. 2021 [cited 2021Oct25]. Available from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/leeftijd/ouderen

[11] van Daal WAJ, Bos MA. Infrastructure for radiotherapy in the Netherlands: Development from 1970 to 2010. International Journal of Radiation Oncology*Biology*Physics. 1997Jan15;37(2):411–5.

[12] Slotman BJ, Vos PH. Planning of radiotherapy capacity and productivity. Radiotherapy and Oncology. 2013Mar7;106(2):266–70.

[13] Ministerie van Volksgezondheid Wen S. Regeling van de Minister van Volksgezondheid, Welzijn en sport van 23 Oktober 2009, Nr. CZ/TSZ-2963442, Houdende Vaststelling van het planningsbesluit radiotherapie 2009. Staatscourant 2009, 16811 | Overheid.nl > Officiële bekendmakingen. Ministerie van Binnenlandse Zaken; 2009 [cited 2021Oct25]. Available from: https://zoek.officielebekendmakingen.nl/stcrt-2009-16811.html

[14] Concentratie. Nederlandse Vereniging voor Radiotherapie en Oncologie. [cited 2021Oct25]. Available from: https://nvro.nl/werk-opleiding/beroepsbelangen/concentratie

[15] Netherlands Comprehensive Cancer Organisation. Netherlands Cancer Registry (NCR). [cited 2021 December, 24th]; Available from: http://www.iknl.nl/en/ncr.

[16] Møller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talbäck M, et al. Prediction of cancer incidence in the Nordic countries: Empirical comparison of different approaches. Statistics in Medicine. 2003;22(17):2751–66.

[17] Larønningen S, Ferlay J, Bray F, Engholm G, Ervik M, Gulbrandsen J, Hansen HL, Hansen HM, Johannesen TB, Kristensen S, Kristiansen MF, Lam F, Laversanne M, Miettinen J, Mørch LS, Ólafsdóttir E, Óskarsson O, Pejicic S, Petterson D, Skog A, Skovlund CW, Tian H, Toorell N, Virtanen A, Aagnes B, Storm HH (2021). NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 9.1 (27.09.2021). Association of the Nordic Cancer Registries. Cancer Registry of Norway. Available from: https://nordcan.iarc.fr/, accessed on 26-10-2021

[18] Breiman L. Random Forests. Machine Learning. 2001;45(1):5–32.

[19] Nowok, B., G.M. Raab & C. Dibben (2016), synthpop: Bespoke creation of synthetic data in R. Journal of Statistical Software, 74:1-26; DOI:10.18637/jss.v074.i11. Available from: https://www.jstatsoft.org/article/view/v074i11

[20] Signaleringscommissie Kanker van KWF Kankerbestrijding. Kanker in Nederland tot 2020; Trends en prognoses, Amsterdam, 2011. Available from:

https://www.medischcontact.nl/web/file?uuid=5e7f00e5-4bb9-4635-8b40-b00aa89805c2&owner=8b0a181f-3a46-40cc-b794-9de61bc0db3f

[21] Netherlands Cancer Registry (NCR), IKNL. [cited 2021Dec26] Obtained from https://iknl.nl/nkr-cijfers?fs%7Cepidemiologie_id=506&fs%7Ctumor_id=1&fs%7Cregio_id=530&fs%7Cperiode_id=545%2C544%2C543%2C542&fs%7Cgeslacht_id=623%2C621%2C622&fs%7Cleeftijdsgroep_id=656&fs%7Cjaren_na_diagnose_id=666&fs%7Ceenheid_id=683&cs%7Ctype=line&cs%7CxAxis=periode_id&cs%7Cseries=geslacht_id&ts%7CrowDimensions=periode_id&ts%7CcolumnDimensions=geslacht_id&lang%7Clanguage=nl

[22] Dinmohamed AG, Visser O, Verhoeven RH, Louwman MW, van Nederveen FH, Willems SM, et al. Fewer cancer diagnoses during the covid-19 epidemic in the Netherlands. The Lancet Oncology. 2020Mar30;21(6):750–1.

[23] Slotman BJ, Vos P, Slot A, Keus R, Verheij M. Radiation oncology in the Netherlands. International Journal of Radiation Oncology*Biology*Physics. 2018;100(1):5–11.

[24] de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. New England Journal of Medicine. 2020;382(6):503-13.

[25] de Munck L, Siesling S, Fracheboud J, den Heeten GJ, Broeders MJ, de Bock GH. Impact of mammographic screening and advanced cancer definition on the percentage of advanced-stage cancers in a steady-state breast screening programme in the Netherlands. British Journal of Cancer. 2020Sep7;123(7):1191–7.

Supplementary document

Predicting the incidence of lung cancer

Background

Various models have been developed for predicting cancer incidence. Møller et al. refer to a number of models in their paper [16], of which the Age-Period-Cohort (APC) model and the short-base model are most commonly used. In this document, these models are compared to select the most appropriate one for our study. In order to make predictions on the incidence of cancer, predictions on demographic developments in the general population are required as input. These predictions are available through Statistics Netherlands (CBS).

Variables

The input variables used for both models are: the population's age distribution, sex, the population's cancer incidence distribution, and the year of observation.

The short-base model uses 11 one-year periods as input, and has the following submodels:

1. PREDNAP: $R_{ap} = A_a(1 + Dp)$
2. PREDAAAP: $R_{ap} = A_a + D_a p$
3. PREDMAP: $R_{ap} = \exp(A_a + Dp)$
4. PREDMAAP: $R_{ap} = \exp(A_a + D_a p)$

R_{ap} is the incidence rate in age category *a* and period *p*. A_a is age category *a*. D is the drift parameter for all age categories. D_a is the drift parameter for age category *a*. *p* is the period. Models 1 and 2 (PREDNAP and PREDAAAP) are both linear in time and are used for cancer sites with an increasing trend. Models 3 and 4 (PREDMAP and PREDMAAP) are both log-linear and are used for cancer sites with a decreasing trend. [16] Models 2 and 4 provide separate slopes for each age group. If these models give a significant improvement compared to model 1 or 3 respectively, then these models should be used.

The APC model uses four to 6 five-year periods and five-year age groups as input, and has the following submodels:

1. Log link: $R_{ap} = \exp(A_a + Dp + P_p + C_c)$

2. Power link: $R_{ap} = (A_a + D * p + P_p + C_c)^5$

R_{ap} is the incidence rate in age category *a* and period *p*. The model being built estimates the incidence rate based on age (Aa), period of diagnosis (Pp) and birth cohort (Cc). Pp is the coefficient associated with period p. In addition to these three variables, a linear component is estimated for the trend in general: This is the so-called drift parameter and represents how much the trend increases (or decreases) each period. It is D*p (drift * period). The power of five in the Power link model is chosen to level off the exponential growth. The outcomes of the APC model are all presented for five-year periods. To get the average number per year, this outcome needs to be divided by 5.

Predictions versus the observed incidence rate

To test the accuracy of the different models, lung cancer incidences up to 2019 were predicted with data from the past and subsequently compared with the observed incidence rates, stratified for sex.

The outcomes of the prediction are expressed as absolute incidence (cases) and standardized incidence rate (ESR). The absolute incidence is the predicted number of cases with all ages combined based on the current age distribution and the predicted age distribution. The ESR is adjusted for the predicted change in age distribution.

Detailed results are shown below. Some formulas did not converge. These rows are filled with 'x'.

Lung cancer incidence in men:

	Prediction 2019 (11 year of input data: 2000-2010)		
Method	Observed	Predicted	Difference (%)
Short-base PREDNAP (cases)	x	x	x
Short-base PREDNAP (ESR)	x	x	x
Short-base PREDAAAP (cases)	7718	8360	8.3
Short-base PREDAAAP (ESR)	59	63	6.8
Short-base PREDMAP (cases)	7718	8340	8.1
Short-base PREDMAP (ESR)	59	64	8.5
Short-base PREDMAAP (cases)	7718	8705	12.8

Short-base PREDMAAP (ESR)	59	66	11.9
	Prediction 2015 – 2019		
Method	Observed	Predicted	Difference (%)
APC power5 (cases)	37974	39864	5,0
APC power5 (ESR)	61.7	66.1	7.1
APC Poisson (cases)	37974	39495	4.0
APC Poisson (ESR)	61.7	65.0	5.3

Lung cancer incidence women:

	Prediction 2019 (11 year of input data: 2000-2010)		
Method	Observed	Predicted	Difference (%)
Short-base PREDNAP (cases)		x	x
Short-base PREDNAP (ESR)		x	x
Short-base PREDAAAP (cases)		x	x
Short-base PREDAAAP (ESR)		x	x
Short-base PREDMAP (cases)	6484	8781	35.4
Short-base PREDMAP (ESR)	49	71	44.9
Short-base PREDMAAP (cases)	6484	9382	44.7
Short-base PREDMAAP (ESR)	49	71	44.9
	Prediction 2015 – 2019		
Method	Observed	Predicted	Difference (%)
APC power5 (cases)	30388	33306	9.6
APC power5 (ESR)	48.1	53.7	11.6
APC Poisson (cases)	30388	36765	21.0
APC Poisson (ESR)	48.1	59.3	23.3

Findings and considerations

The APC model uses a maximum of 6 five-year periods, while the short-base model uses eleven one-year periods. The best APC submodel gives a more accurate estimate than the best short-base submodel for both men and women in 2019. In addition to the results shown in this document suggesting that the APC model gives the most accurate estimations, the APC model has two main advantages making it the preferred model for making predictions of cancer incidences. First of all, the predicted trendline can be adjusted to prevent overestimating the incidence. This can be done by cutting the trend line: each year in the five-year period can be trimmed with a percentage relative to the year before. This could be a useful feature to incorporate expected future effects on the incidence which could not be seen yet in the input data, e.g. changed smoking behaviour in women for the prediction of lung cancer incidence. Secondly, a birth cohort is included in the model, which is valuable when there are exposures or circumstances in the past making some generations more likely to be diagnosed with cancer than future (predicted) generations.

Identifying predictors for the use of radiotherapy: codebook

Variable name	Meaning	Description	Levels
Sex	Sex	The sex of the patient	1 = Men 2 = Women
age	Age	The age of the patient at the time of diagnosis	
Stad_numeric	Stage	Clinical tumor stage according to the eighth edition of Tumor Node Metastases (TNM) classification	1 = stage I 2 = stage II 3 = stage III
RT_total	Radiotherapy Total	States whether the patient received any type of radiotherapy as part of their treatment, either as the main	0 = No 1 = Yes

		treatment modality or as a (neo)adjuvant therapy.	
ct	Clinical T stage	Primary Tumor [4] The eighth edition of the TNM classification will provide a more detailed description of the levels.	0 1 1A 1B 1C 2A 2B 3 4 X
Cn	Clinical N stage	Regional Lymph Nodes [4] The eighth edition of the TNM classification will provide a more detailed description of the levels.	0 1 2 3 X
tumorsoort	Tumor type	Classification of the tumor type	302310 = NSCLC 302320 = SCLC 302340 = Other
Region	Region	The region that the patient lived at the time of diagnosis.	Zuid-Holland Noord-Holland Noord-Brabant Gelderland Limburg Utrecht Groningen Zeeland Flevoland Drenthe

			Friesland Overijssel
SES	SES	Social Economic Score of the patient at the time of diagnosis.	1 2 3 4 5 6 7 8 9 10
Traveltime_cat	Travel time	One way travelling time by car and calculated using the postal code of the nearest radiotherapy facility and the patient's home address postal code at the time of diagnosis.	1 = <15 min 2 = <30 min 3 = <45 min 4 = >45 min
Type_instelling	Hospital type	Classification of the type of hospital diagnosing lung cancer	Academic Top Clinical General
Rt_in_structuur	Radiotherapy in-house	States whether the diagnosing hospital has a radiotherapy department embedded in the organization of the hospital.	0 = No 1 = Yes
Morf_cat	Morphology category	A categorization of morphology into four main categories.	NSCLC Squamous carcinoma NSCLC Adenocarcinoma NSCLC Large Cell/Other SCLC

Topog	Topography	States the specific location of the tumor in the body	C340 C341 C342 C343 C348 C349
later	Lateralization	Describes which side of an paired organ is the origin of the primary cancer	1 2 3 4 X
diffgr	Tumor grade	Describes how abnormal the tumor cells and the tumor tissue look under a microscope	1 2 3 4 9
basisd	Basis for diagnosis	Describes the basis for diagnosis	1 2 4 5 6 7