Bayesian spatio-temporal modelling of tobacco-related cancer data in Switzerland

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

von

Verena Jürgens

aus Oldenburg, Deutschland

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. Jürg Utzinger, PD Dr. P. Vounatsou and Prof. Dr. Annibale Biggeri.

Basel, den 23. April 2013

Prof. Dr. Jörg Schibler
Dekan

The whole is greater than the sum of its parts.

Aristotle

Acknowledgements

This PhD thesis was carried out under the supervision of PD Dr. Penelope Vounatsou and funded by Oncosuisse. I am grateful for any support I received in writing this thesis.

My first lines go to the person this journey started with at Swiss TPH, my supervisor. Penelope, I am grateful for your support I received during the last years. I greatly appreciate that you gave me the opportunity to benefit from your scientific expertise and thank you for helping me to grow in dealing with challenges in research.

Σε ευχαριστώ μέσα από τα βάθη της καρδιάς μου.

This work was done within a project build on several collaborations. Thank you Silvia for contributing to this work with your expert knowledge. My grateful wishes also go to the other collaborators: Thomas Cerny, Harish C. Phuleria, Martin Früh, Matthias Schwenkglenks and Harald Frick. Thank you for your input and support.

I also would like to dedicate a few words to somebody who was not directly involved in my PhD project, but contributed in a special way. Thank you, Marcel, for being the "captain" of an institute, which I feel part of. Many thanks for this motivating spirit.

I am very thankful for the finanical support I received from Oncosuisse. I gratefully acknowledge the financial support from the Dissertationenfonds der Universität Basel/ Basler Studienstiftung for printing this thesis.

Looking back at the last years, I remember many, many people I met during my time at the institute. I very much enjoy(ed) all the exchange with open minded people, that allowed me to learn more about different cultures and life. Thank you so much!

I am not a person of "big words" and don't want to miss out anybody... therefore, an honest appreciation: Thank you all for making this time that exciting!

However, there is no avoiding of listing a few names: Aboud, Alex, Evi, Fede, Fred, Friedi, Laura, Sara, Stephi, Tamara, Vreni – thank you for being my friends!

My deepest dedication goes to the four most important people in my life – there are no words to express my feelings to thank you for your infinite patience and support.

Summary

Tobacco use is the leading cause of preventable death worldwide. Each year the tobacco epidemic accounts for 6 million deaths and costs hundreds of billions of dollars to the economy. Cigarette smoking accounts for more deaths than AIDS, murder, legal and illegal drugs, road accidents and suicide combined. Around 85–90% of all lung cancer deaths are estimated to be attributed to active or passive smoking. In Switzerland, lung cancer is the first cause of cancer mortality in men and second in women (after breast cancer). Gender-specific smoking patterns differ essentially in time as well as in space. In the $19th$ and the beginning of the following century, smoking was restricted to the male population, finding its peak in the 1970s in most European countries. In the past, the image of female tobacco use experienced an essential turn. In the middle of the $20th$ century the smoking prevalence among women increased due to the changes in gender roles and the subsequent effect on female smoking reputation. Before, female smoking had not been socially accepted. After strong gender-related developments, female smoking was associated with independence, emancipation and freedom. This movement was exploited to a great extent by the tobacco industry by adjusting their marketing strategies regarding this new target audience. In many developed countries the gap between gender and smoking prevalence is closing since the last decades, as males are smoking less, while female tobacco smoking is increasing steadily.

Information on spatial as well as temporal patterns and trends of a disease are essential for health planning and intervention purposes. The Swiss Federal Office of Public Health (FOPH) has launched the National Programme Tobacco 2008–2012 aiming to reduce the proportion of smokers, targeting a decline of tobacco-related morbidity and mortality in the country as a final result. Cancer mapping visualizes geographical and temporal patterns and trends. Maps of estimated mortality serve as helpful tools to identify high risk areas and therefore enable focused intervention planning at a higher geographical scale than the national one.

Disease maps of crude rates can be non-informative and might even lead to misinterpretation, as rare diseases or small populations might dominate the map and result in large variability in the estimated rates. Distinction between chance and real difference of the obtained variability is challenging. Spatial modelling of the rates enables the assessment of covariate effects to explain observed patterns and highlight them by obtaining smooth maps. Bayesian methods are the state-of-the-art modelling approach for spatio-temporal analysis. Model formulations improve the estimation of sparse, unstable rates by borrowing strength from their neighbours. In addition, they allow risk factor analysis which takes into account potential spatial correlation. Apart flexible modelling, Bayesian inference provides computational advantages via the implementation of Markov chain Monte Carlo (MCMC) simulation methods.

This thesis aimed (i) to assess geographical differences and trends of age- and gender-specific lung and all tobacco-related cancer mortality in Switzerland; (ii) to project tobacco-related cancer mortality in Switzerland at different geographical levels accounting for spatial variation; (iii) to develop Bayesian age-period-cohort (APC) models for projecting cancer mortality; (iv) to develop Bayesian back-calculation models to estimate age- and gender-specific incidence from sparse mortality data; and (v) to develop models to indirectly approximate gender-specific smoking patterns in space and time by unadjusted and adjusted lung cancer mortality rates with non-smoking risk factors.

In *Chapter 2* Bayesian spatio-temporal models were developed to estimate age- and gender-specific lung and other tobacco-related cancer mortality in Switzerland from 1969 up to 2002. Age-standardized mortality ratios were estimated at municipality level, adjusted for urbanization level and language region. Furthermore, models included random effects to account for spatial variation. Nationwide smooth maps of lung as well as other tobacco-related cancer mortality were produced, identifying high and low risk areas and illustrating the temporal trend for each subgroup.

In *Chapter 3* Bayesian APC models were applied to project Swiss gender-specific tobacco-related cancer mortality at national level and for each language region in Switzerland, i. e. German, French and Italian. The models were extended by including spatial random effects to predict cancer mortality due to tobacco use at cantonal level.

Moreover, nationwide projections were obtained for the age groups 50–69 and ≥ 70 years. Maps as well as future estimates of the disease-specific mortality are helpful tools in intervention planning and identification of region-specific disparities.

In *Chapter 4* a Bayesian APC model with power-link function was developed to overcome

limitations of existing models such as estimation of extreme rates based on the log-link function. The power parameter was considered to be random, estimated by the dataset. Model performance was assessed by comparing observed age- and gender-specific lung cancer mortality rates with one-step ahead projections. The model was applied to project lung cancer death counts in Switzerland up to 2018.

In *Chapter 5* a back-calculation model was applied to estimate incidence from mortality. Existing models were further developed to allow for estimation of rare cancer incidence. Data from Eastern Switzerland were analysed to obtain age- and gender-specific survival parameters. Incidence was estimated by linking subgroup-specific mortality with the corresponding survival distribution. Countrywide maps of gender-specific lung cancer incidence were produced, assuming a constant lung cancer survival across the cantons.

In *Chapter 6* proxies of spatial patterns of smoking prevalence in Switzerland were developed. In particular, Bayesian spatial logistic regression models were applied to analyse smoking data provided from the Swiss Health Survey (1992) and to study gender-specific trends. Furthermore, lung cancer mortality data of 2008–2010 have been modelled using negative binomial regressions with spatial random effects to obtain smoking proxies (i. e. smooth mortality rates and spatial random effects). Validation was done (i) graphically, by comparing maps of model-based smoking patterns and each proxy; (ii) numerically, by estimating Kendall's τ_b coefficient and by applying logistic regression models to quantify the magnitude of association. Results indicated that both, the spatial random effect as well as the mortality rate approximated well the smoking patterns. Our approach provides an indirect smoking approximation which in the absence of tobacco use survey data relies on the availability of lung cancer mortality and non-smoking risk factors.

The main contributions of the research are (i) improved statistical models for projecting cancer mortality. These models allow forecasting of sparse count data; (ii) improved statistical models for estimating tobacco-related incidence from mortality data. These models enable estimation of incidence rates of the disease in the parts of the population, which are not covered by cancer registries; (iii) smooth maps of gender-, age- and site-specific patterns of lung and other tobacco-related cancer mortality over time. These maps identify discrepancies of disease burden and assist implementation and evaluation of the National Programme Tobacco; (iv) a better insight into the differences in tobacco-related cancer mortality rates between linguistic regions and urbanisation by gender; (v) smooth maps of gender-specific patterns of lung cancer incidence for a recent period; (vi) estimates of the geographical patterns of tobacco-related cancer mortality for the next 10 years: this

information is useful for planning purposes such as resource allocation and costing of medical supplies for diagnosis and treatment; (vii) smooth maps of lung cancer mortality, adjusted/unadjusted for non-smoking risk factors, which allow to find a good proxy of smoking behaviour in order to study gender-specific geographical patterns of smoking in Switzerland from limited smoking surveys.

Zusammenfassung

Tabakkonsum ist die führende Ursache vermeidbarer Krankheiten weltweit. Jedes Jahr sterben ca. 6 Millionen Menschen an den Folgen der Tabakepidemie und sie verursacht Wirtschaftskosten in Höhe von hunderten Milliarden. Rauchen ist verantwortlich für mehr Todesfälle als die Summe aus AIDS, Mord, legale und illegale Drogen, Verkehrsunfälle und Suizid. Schätzungen zufolge sind rund 85–90% aller Lungenkrebstodesfälle auf aktiven oder passiven Tabakkonsum zurückzuführen. Männlicher Lungenkrebs ist für den größten Anteil der gesamten Krebsmortalität in der Schweiz verantwortlich, wobei diese Position bei den Frauen vom Brustkrebs, gefolgt von Lungenkrebs, eingenommen wird.
Geschlechtsspezifisches Rauchverhalten weist wesentliche Unterschiede in räumlichen und zeitlichen Trends auf. Im 19. und Anfang des 20. Jahrhunderts war Rauchen ein Habitus des männlichen Teils der Gesellschaft. In den meisten europäischen Ländern fand das männliche Tabakkonsumverhalten seinen Höchstpunkt in den 1970ern. Im Gegenzug dazu erlebte das Gesellschaftsbild der rauchenden Frau in der Vergangenheit eine drastische Kehrtwende. Infolge der Veränderungen der Geschlechterrollen und dem anschließenden Wandel des Ansehens des rauchenden weiblichen Geschlechts stieg die Prävalenz Mitte des 20. Jahrhunderts an. Zuvor wurden rauchende Frauen als nicht gesellschaftsfähig angesehen. Nach starken Entwicklungen bezüglich der Geschlechterrollen wurde das Rauchen der weiblichen Bevölkerung mit Unabhängigkeit, Emanzipation und Freiheit assoziiert. Diese Bewegung wurde im großen Maße von der Tabakindustrie ausgenutzt, indem sie ihre Marketingstrategien an die neue Zielgruppe angepasst haben. In vielen Entwicklungsländern hat sich die Lücke der geschlechtsspezifischen Raucherprävalenz in den letzten Jahrzehnten verkleinert. Dies ist darauf zurückzuführen, dass der Tabakkonsum der Männer zurückgegangen ist und zeitgleich das weibliche Pendant einen steten und starken Anstieg erlebt habt.
Wissen über räumliche als auch zeitliche Verhaltensmuster und Trends einer Krankheit ist wesentlicher Bestandteil für die Gesundheits- und Interventionsplanung. In der Schweiz

hat das Bundesamt für Gesundheit (BAG) das Nationale Programm Tabak 2008–2012 ins Leben gerufen. Zielsetzung des Programms ist die Reduzierung des Raucheranteils und längerfristig ein Rückgang der landesweiten tabakbedingten Morbidität als auch Mortalität. Kartierung des Krebsvorkommens visualisiert geografische und zeitliche Muster und Verläufe. Karten der geschätzten Mortalität dienen als nützliche Hilfsmittel um Gebiete mit hohem Risiko zu identifizieren. Dadurch wird eine zielgerichtete Interventionsplanung auf einem höheren geografischen Maßstab als der Landesebene ermöglicht.
Kartierung von Krankheiten mittels rohen Raten kann uninformativ sein und sogar zu Fehlinterpretation führen, da seltene Krankheiten oder eine kleine Bevölkerungszahl die Karte dominieren können und in großer Variabilität der geschätzten Raten resultieren. Differenzierung zwischen Zufall und wirklichem Unterschied der beobachteten Variabilität ist problematisch. Räumliche Modellierung von Raten ermöglicht die Abschätzung von Effekten der Kovariaten, um beobachtete Muster zu erklären und hebt diese durch geglättete (*smooth*) Karten hervor. Bayes'sche Methoden sind der neueste Stand der raum-zeitlichen Analysen. Sie ermöglichen flexible Modellierung und Schlussfolgerung und bieten rechnerische Vorteile durch die Implementierung von Markov chain Monte Carlo (MCMC). Modellformulierungen verbessern die Schätzungen von instabilen Raten, indem diese durch deren Nachbarn stabilisiert werden. Dadurch heben die resultierenden geschätzten geglätteten Raten Muster hervor und erlauben die Einschätzung der Signifikanz der einzelnen Risikofaktoren unter Berücksichtigung der geografischen Korrelation.
Das Ziel dieser Dissertation war es (i) Bayes'sche raum-zeitliche Modelle für die Analyse der Schweizer tabakbedingten Krebsmortalitätsdaten zu entwickeln und landesweite geglättete Karten der alters-, geschlechts- und standortspezifischen Verhaltensmuster dieser Krebsarten seit 1969 zu erstellen; (ii) tabakbedingte Krebsmortalität in der Schweiz für unterschiedliche geografische Einheiten bis zum Jahre 2018 unter Berücksichtigung der geografischen Korrelation vorherzusagen; (iii) Modelle zur Vorhersage von geschlechtsspezifischen tabakbedingter Krebsmortalität zu entwickeln; (iv) Bayes'sche Rückrechnungsmodelle zu entwickeln, um mittels Mortalität Inzidenz nach Alter und Geschlecht zu schätzen; und (v) Modelle zu entwickeln, welche das geschlechtsspezifische Rauchverhalten in der Schweiz indirekt approximieren, um mittels angepasster Lungenkrebsmortalität landesweite Trends der Raucherprävalenz zu schätzen.
In *Kapitel 2* wurden Bayes'sche raum-zeitliche Modelle zur Schätzung von alters- und geschlechtsspezifischen Lungenkrebs- und sämtlicher tabakbedingten Krebsmortalität in der Schweiz von 1969 bis 2002 entwickelt. Altersstandardisierte Mortalitätsraten wurden

auf Gemeindeebene geschätzt und hinsichtlich des Grads der Urbanisierung und der lokalen Sprache angepasst. Des Weiteren wurde räumliche Variation mittels *random effects* auf Gemeindeebene berücksichtigt. Es wurden landesweite geglättete Karten der Lungenkrebs- als auch sämtlicher tabakbedingter Krebsmortalität erstellt, welche Gebiete mit hohem als auch niedrigem Risiko identifizieren und zeitliche Trends der jeweiligen Untergruppen illustrieren.

In *Kapitel 3* wurde mittels Bayes'sche Alter-Periode-Kohorte (APC) Modelle geschlechtsspezifische tabakbedingte Krebsmortalität auf kantonaler und nationaler Ebene sowie für die Sprachregionen Deutsch, Französisch und Italienisch prognostiziert. Eine Erweiterung des Modells für die Schätzung kantonaler Krebsmortalität bedingt durch Tabakkonsum beinhaltete räumliche random effects. Zusätzlich wurden landesweite Projektionen für die Altersgruppen 50–69 und ≥ 70 bestimmt. Sowohl Karten als auch Schätzungen zukünftiger krankheitsspezifischer Mortalität sind unterstützende Hilfsmittel hinsichtlich der Interventionsplanung und der Identifizierung regionaler Unterschiede.

In *Kapitel 4* wurde ein Bayes'sches Power Model zur Projektion von Krebsmortalität und -inzidenz entwickelt. Schätzungen basierten auf einer Power Link-Funktion zwischen der Mortalitätsrate und den Effekten von Alter, Periode und Kohorte, indem ein fixer Power-Parameter mit dem Wert fünf angenommen wurde. Als eine Erweiterung dessen wurde ein Model entwickelt, welches einen nicht-fixen Power-Parameter annimmt, um dessen Grad durch die Daten schätzen zu können. Modellvalidierung wurde anhand Vorhersage bekannter Schweizer alters- und geschlechtsspezifischer Lungenkrebsmortalität durchgeführt. Jene Modelle, welche die beste prädiktive Leistung aufwiesen, wurden angewandt, um Lungenkrebstodesfälle bis 2018 zu prognostizieren. Die entwickelten Modelle adressieren Kritik, welche gegen das etablierte Power- als auch das Bayes'sche APC Modell geäußert wurden, indem deren jeweiligen Stärken kombiniert wurden.

In *Kapitel 5* wurde das Rückrechnungsmodell angewandt, welches die Schätzung der Inzidenz auf Basis von Mortalität ermöglicht. Das Modell wurde weiterentwickelt, um verlässliche Schätzungen auch für seltene Krebsarten zu ermöglichen. Die Analyse der Inzidenzdaten der Ostschweiz umfassten Follow-up-Zeit und -Status und ermöglichten die Schätzung der alters- und geschlechtsspezifischen Überlebensdauer in der Region. Schätzungen der Inzidenz wurden durch Verbindung untergruppenspezifischer Mortalität mit der entsprechenden geschätzten Überlebensdauer ermittelt.

In *Kapitel 6* wurden indirekte Schätzer für geschlechtsspezifisches Rauchverhalten in der Schweiz entwickelt. Vorerst wurden logistische Modelle angepasst, um Daten der Schweizer

Gesundheitsbefragung bzgl. der Raucherprävalenz zu analysieren. Des Weiteren wurden räumliche Negativ-Binomiale Regressionsmodelle eingesetzt, welche hinsichtlich der Nichtraucherrisikofaktoren angepasst und auf Daten der Lungenkrebsmortalität in 2008–2010 angewandt wurden. Umweltrisikofaktoren beinhalteten Exposition zu Radon und Luftverschmutzung. Resultierende räumliche random effects als auch Mortalitätsraten basierend auf unangepassten Regressionsmodellen wurden als Schätzer des Rauchverhaltens genutzt. Validierung erfolgte (i) grafisch durch Vergleich der räumlichen Verteilung des modellierten Rauchverhaltens und den indirekten Schätzern; (ii) numerisch, durch Schätzung des Kendall's τ_b Koeffizienten und Verwendung von logistischen Regressionsanalysen. Ergebnisse zeigten, dass sowohl die Mortalitätsraten als auch die räumlichen random effects die Raucherprävalenz gut approximierten. Unsere Methodik ermöglicht es die Verteilung von Raucherprävalenz indirekt abzuschätzen und benötigt Daten über Risikofaktoren und Lungenkrebsmortalität.

Die wichtigsten Beiträge dieser Dissertation sind (i) entwickelte statistische Modelle zur Vorhersage von Krebsmortalität; (ii) erweiterte Modelle zur Rückberechnung von Krebsinzidenz anhand gegebener Krebsmortalitätsdaten; (iii) die Erstellung landesweiter geglätteter Karten über die alters- und geschlechtsspezifische Lungen- und sämtliche tabakbedingte Krebsmortalität in der Schweiz von 1969–2002. Hierfür wurden raum-zeitliche Bayes'sche Negativ-Binomiale Regressionsmodelle verwendet; (iv) einen besseren Einblick in tabakbedingte Krebsmortalität nach Sprachgebiet und Urbanisierungsniveau; (v) geglättete Karten über aktuelle Lungenkrebsinzidenz; (vi) die Schätzung der zukünftigen alters- und geschlechtsspezifischen Lungen- und sämtlicher tabakbedingten Krebsmortalität in der Schweiz bis zum Jahre 2018; und (vii) die Entwicklung von Bayes'schen Modellen zur indirekten Annäherung an das Rauchverhalten, welche die Schätzung von landesweiten Tabakkonsumverhalten für jedes Geschlecht ermöglicht, indem die Modelle hinsichtlich der Nichtraucherrisikofaktoren angepasst wurden.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Tobacco – a global epidemic

Every day more than 15 billion cigarettes are smoked worldwide. Manufactured cigarettes are the predominant form of tobacco-consumption. Several other types of tobacco use exist as for example pipes and cigars. China is the leading cigarette manufacturer – its production increased by 600% from 1960 to 1995 (MacKay and Eriksen, 2004). On a global scale, each year over five trillion cigarettes are produced. Over time, the composition of a manufactured cigarette has changed to a great extent. In the middle of the 20*th* century strong health concerns on smoking caused first efforts to reduce the harm of cigarettes by reducing its tar and nicotine yields. First attempts introduced filtered cigarettes which were followed by *light* cigarettes. However, till today no clear evidence for a reduced health risk is given for these tobacco products. Overall, former attempts to lower the harm of smoking made cigarettes more addicting by adding chemicals or the use of flavours.

1.1.1 Gender- and age-specific tobacco use

Up to the beginning of the 20*th* century smoking was only socially accepted for males. After the First World War the tobacco industry targeted female smoking. While before that time, smoking by women was judged as misbehaviour, the tobacco companies created a new image of female smoking by their advertising strategy. This new image suggested a changing, more independent role of women. Philip Morris even offered lessons on how to smoke for women in London. Media played a fundamental role on the way to tobacco addiction. Figure 1.1 shows an advertisement of Philip Morris in the middle of the 20*th* century. It clearly targets women by suggesting smoking as a sign of emancipation and freedom.

While female smoking is declining in some developed countries as Australia, the USA, Canada and the UK, the number of women smoking remains stable or still increases in many European countries. More than 10% of women are smoking during pregnancy. In addition, they are exposed to many sources of second-hand smoke as public places, at work or even in the private environment at home. Youth smoking is a big challenge. The majority of overall smokers start to consume tobacco before they reach the age of adulthood. About 25% of them smoked their first cigarette before having their 10*th* birthday (MacKay and Eriksen, 2004). Tobacco industry advertisement is the major risk of youth smoking. As the industry applied their strategy to boost female smoking by emphasizing on emancipation and gender equality, it is also applied to younger males and females. While cigarette

Figure 1.1: Philip Morris Tobacco Company advertisement in 1951 (source: tobacco-standford.edu).

smoking is associated with coolness among teenagers and children, it plays an important role in the socializing process. Peer pressure, allegiance, stress management, suppress appetite and self-confidence are all aspects which have a great impact during youth and are also directly linked with tobacco use.

> 'It is important to know as much as possible about teenage smoking patterns and attitudes. Today's teenager is tomorrow's potential regular customer, and the overwhelming majority of smokers first begin to smoke while still in their teens... The smoking patterns of teenagers are particularly important to Philip Morris.'(Philip Morris Companies Inc. 1981)

Studies have shown that even candy cigarettes (Fig. 1.2) double the risk of smoking in older age. In many countries, i. e. Canada, Great Britain, Norway and Australia, candy cigarettes are not allowed to be sold anymore. However, in Switzerland, and also many other countries, they are still available in shops. Lopez et al. (1994) have developed a model to describe the cigarette epidemic in developed countries. The authors used lung cancer death rates as an index of total smoking-attributable mortality. They point out three sources of information which can be used to describe the transition through the four stages of their model: (i) How many adults smoke regularly? (ii) How much does an adult smoke in a given period? (iii) How many people die from smoking (by age, sex and cause

Figure 1.2: Candy cigarette advertisement (source: tobacco-standford.edu).

of death)?

The first stage of the model describes the origin of the smoking epidemic, characterized by low male smoking prevalence. Driven by social conventions, tobacco use among women is rarely seen. The health risks of smoking are not even recognized and lung cancer cases among smokers are not yet inflated. The subsequent stage is influenced by a rapid increase of male smoking prevalence, finding its peak around 50–80%. Furthermore, smoking prevalence among women is rising as well, but lags behind by 10–20 years.

Stage III is driven by a decline in male smoking prevalence, as many middle-aged and older men stopped smoking. In addition, an initial decrease of female prevalence occurs by the end of this phase, which lasts for around three decades. The essential development during this stage is the rapid increase in smoking-attributable mortality. Anti-tobacco campaigns are established and smoking undergoes another change in social acceptance to be rather seen as socially aberrant.

The final stage (IV) describes a continuing, but slow decrease in overall smoking prevalence. Death due to smoking is expected to fall below 30% for males. On the contrary, female smoking-attributable mortality is increasing sharply, finding its peak at around 20–25% of all deaths after 30 years. Furthermore, protection against passive smoking at an individual level is discussed and policies also need to focus on the problem of failing attempts

of smoking cessation by smokers who are addicted to nicotine. Figure 1.3 illustrates the process of the epidemic as defined by Lopez et al.

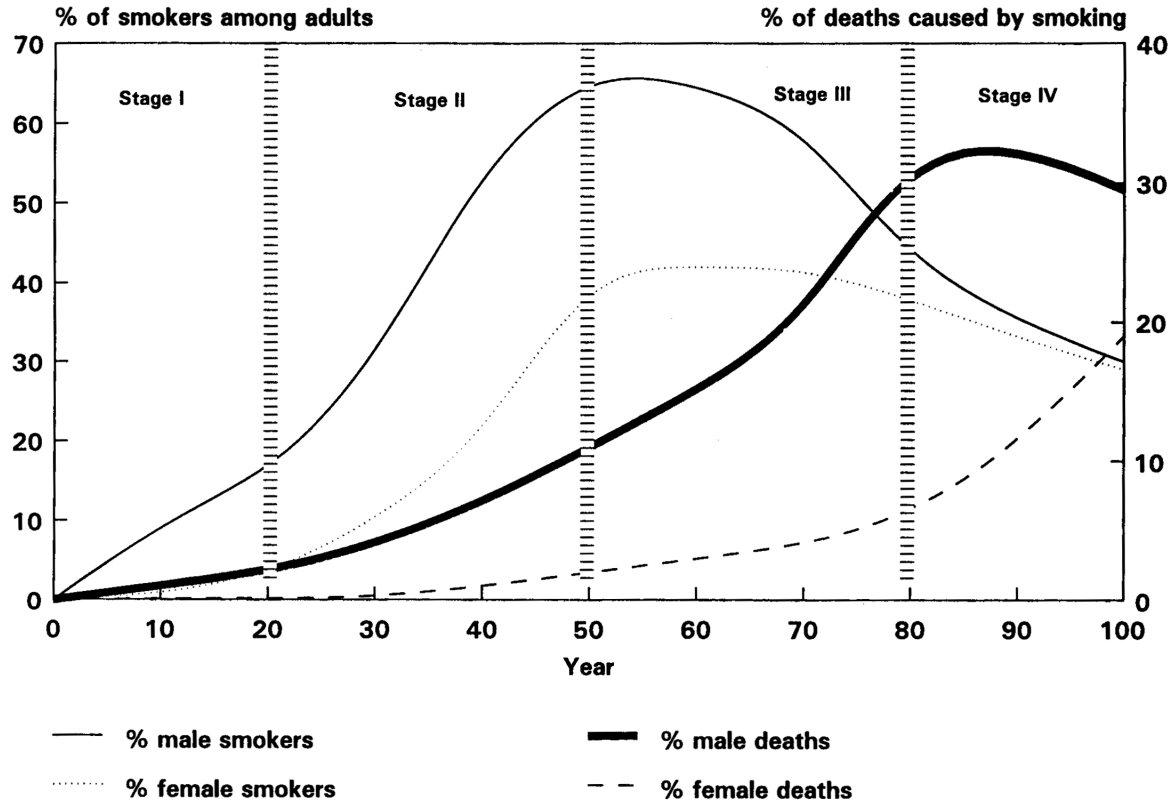


Figure 1.3: Cigarette epidemic model (source: (Lopez et al., 1994)).

1.1.2 Passive smoking

In the last decades, after assessing the significant relation between smoking cancer and lung cancer, several studies are focused on assessing the impact of passive smoking, so-called Environmental Tobacco Smoke (ETS) on health. Association of ETS exposure with several risks has been reported, i. e. increased risk of lower respiratory tract infections (such as bronchitis and pneumonia) and asthma in children as well as lung cancer in adults (United States Environmental Protection Agency, 1992). A frequently studied source of exposure to ETS are non-smoking spouses of smokers. A meta-analysis reported a pooled risk ratio of 1.27 (95% CI 1.17–1.37) for never-smoking women to passive smoking from spouses. However, this risk might be confounded by other factors (Taylor et al., 2007).
For European countries, strong correlation between strength of tobacco control policies

and concern about passive smoking has been reported (Willemsen et al., 2012). Therefore, public information and education on the harmful health consequences from passive smoking are essential.

1.2 Tobacco use in Switzerland

While in the beginning of the last century female smoking was socially not accepted, its prevalence increased constantly with each following birth cohort. In the 1950–1959 cohort, smoking prevalence among both sexes was nearly balanced (37% and 41% for females and males, respectively) (La Vecchia et al., 1988). Another trend of female smoking is the earlier age of starting smoking for younger ones, which in addition is characterized by heavier tobacco consumption. Furthermore, it has been stated, that older men are more likely to quit smoking compared to older women. Since the early 1970s cigarette sales increased steadily with a strong rise, which levelled off around 3 000 cigarettes per adult/year on average (La Vecchia et al., 1988). Gmel (2000) reported increasing trends in current smoking in 1993–1997 for both genders. Costanza et al. (2006) analysed gender-specific habits in tobacco use and concluded a stable smoking prevalence for males and females based on a surveillance study in Geneva during 1993–2003. In 2010, the smoking population was around 30% for males and 24% for females. Almost half of them (48%) intented to quit smoking, while this willingness is highest in the French-speaking part of the country (Keller et al., 2011).

1.2.1 Tobacco-related policies

Lack of knowledge regarding the harms of tobacco use has been observed among smokers as well as non-smokers. Many people are unaware of the range of risks implied and diseases caused by tobacco-smoking (WHO, 2008).

Light cigarettes are one example of a wrong perception, as they are seen to be less adverse compared to regular ones. Cigarettes labelled as *light* are often used as a pre-step before quitting. Studies showed that there is a positive association of switching to light cigarettes and the attempt to quit smoking. However, so far there is no evidence that changing to light-tar cigarettes increases the chance to stop smoking (Weinstein, 2001).

Different strategies exist to prevent and control the burden of tobacco use and related diseases. In Switzerland, the National Programme Tobacco 2008–2012 has been launched by the Federal Office of Public Health (FOPH) with the long-term goal to reduce the number of tobacco-related morbidity and mortality. In 2012, it was extended till the end

of 2016 and determines the national strategy regarding tobacco-prevention.
As the country is splitted into 26 cantons, each being independent and relying on a local government and decision making processes, no strict national agreement or law exists. In 2007, the majority (64%) of the Swiss population (aged 14–65) favoured a smoking ban in restaurants, cafés and bars (Krebs et al., 2008). The acceptance was highest in the Italian-speaking region, followed by the French- (74%) and German-speaking (61%) part of the country. In 2010, the Federal law to protect against passive smoking was conducted. The smoking ban regulation aimed to achieve smoke-free interiors, covering public or working places as restaurants, bars, schools, hospitals etc. As a refinement of this minimal protection, further details have been defined independently at cantonal level. While some cantons allowed service in rooms labelled as *smoke-rooms*, some prohibit any service in rooms where smoking is allowed. This regulation, not offering any service in smoke-rooms, protects the waiting staff, which is no longer exposed to the smoke of the guests and being therefore passive-smokers.
In 2012, the Swiss population had to vote regarding an initiative for stricter smoking ban regulations to protect against passive-smoking. This policy would have implied no service in smoke-rooms for the whole country. With a turnout of 42.3% the initiative failed. Around 34.0% voted for the new initiative, while the lowest number of proponents was observed in the German-speaking part (31.5%), followed by the French-speaking area (39.7%) and almost half of the population living in the Italian-speaking region (48.8%). Geneva was the only canton that voted for the initiative (51%).
Other federal smoking regulations include health warning labels on cigarette packs. Labels either show statements of the health consequences of smoking or illustrate them, as a metastatic lung.
In 2011, the FOPH launched a rather contrary campaign called SmokeFree, focusing on advantages from non-smoking instead of highlighting disadvantages and risks from smoking. The initiative attempts to emphasize on the coolness of non-smoking and the positive implications.

1.2.2 The tobacco lobby

Worldwide, the tobacco industry influences tobacco-related interventions and public health efforts. In Switzerland, the tobacco market is mainly driven by two tobacco companies: Philip Morris and British American Tobacco, controlling the market of cigarette sales of 45–50% each. It has been stated that Switzerland is a paradise for the tobacco industry, being influenced by the tobacco lobby to a greater extent than anywhere else.

> 'We have now a clear view of the tobacco industry's strategies in Switzerland as a result of lawsuits in the United States which have made millions of pages of previously secret tobacco industry documents public. These documents reveal that [...] the tobacco industry made a large, and largely invisible, effort in Switzerland to prevent implementation of meaningful tobacco control legislation and policies in Switzerland.'(Lee and Glantz, 2001)

In addition, the lowest tobacco excise tax in Western Europe can be found in Switzerland. One essential component of an effective intervention strategy planning is to discuss in public and finally aim to eliminate the counter movement by the tobacco industry. In Switzerland, this factor has been identified as a major obstacle of tobacco prevention. However, interventions in the past have rather ignored this influence (Lee and Glantz, 2001).

1.3 Tobacco-related cancer

Tobacco use is the single most preventable cause of death in the world today. It is estimated to reduce normal life expectancy by 20–25 years (United States Department of Health and Human Services, 2010). In the $20th$ century around 0.1 billion deaths are due to tobacco use. This number is estimated to be ten times higher for the $21st$ century (MacKay and Eriksen, 2004). Smoking is a risk factor of cardiovascular disease (CVD). It accounts for more than 10% of all death cases from CVD in 2000 (Ezzati et al., 2005). Around 25% of ischaemic heart disease cases and 75% of chronic bronchitis and emphysema are attributable to smoking. Furthermore, tobacco use is the predominant risk factor for cancer – worldwide, it accounts for 22% of overall cancer mortality. It causes a range of cancer sites, as cancer of the lung, larynx, oral cavity, pharynx, esophagus, pancreas, bladder, kidney, cervix, stomach and acute myeloid leukaemia. Furthermore, a causal relation between tobacco smoking and liver and colorectal cancer was suggested (United States Department of Health and Human Services, 2004). However, no official definition of tobacco-related cancer exists. According to Doll et al. (1976), the following cancer sites are defined as related to tobacco: lung (C33–C34), oesophagus (C15), rectum (C20), pancreas (C25), bladder (C67) and cancer of other respiratory sites (C00–C14) including cancer of the lip, tongue, mouth, larynx, trachea and pharynx (excluding nasopharynx). Numbers in brackets refer to the $10th$ revision of the International Classification of Disease (ICD-10). It has been shown, that quitting smoking at various ages reduces the lifetime risk. The risk for tobacco-related diseases remains for smokers quitting after 44 years of age. However,

several conditions have to be considered regarding this threshold of age, i. e. age of starting and the intensity of smoking.

Lung cancer

Smoking is the main risk factor of lung cancer. It accounts for 90% of lung cancer morbidity and around 71% of all lung cancer deaths. Passive smoking increases the risk of lung cancer by 20–30%. Figure 1.4 shows estimated male lung cancer mortality for the year 2008. Highest rates are observed for Central-Eastern, Southern Europe, Northern America and Eastern Asia. For females, lung cancer mortality rates at a global scale are highest in Northern America (see Fig. 1.5). Lowest lung cancer mortality for both genders is given in Africa, in particular Middle Africa. It should be mentioned that both maps are characterized by different scales.

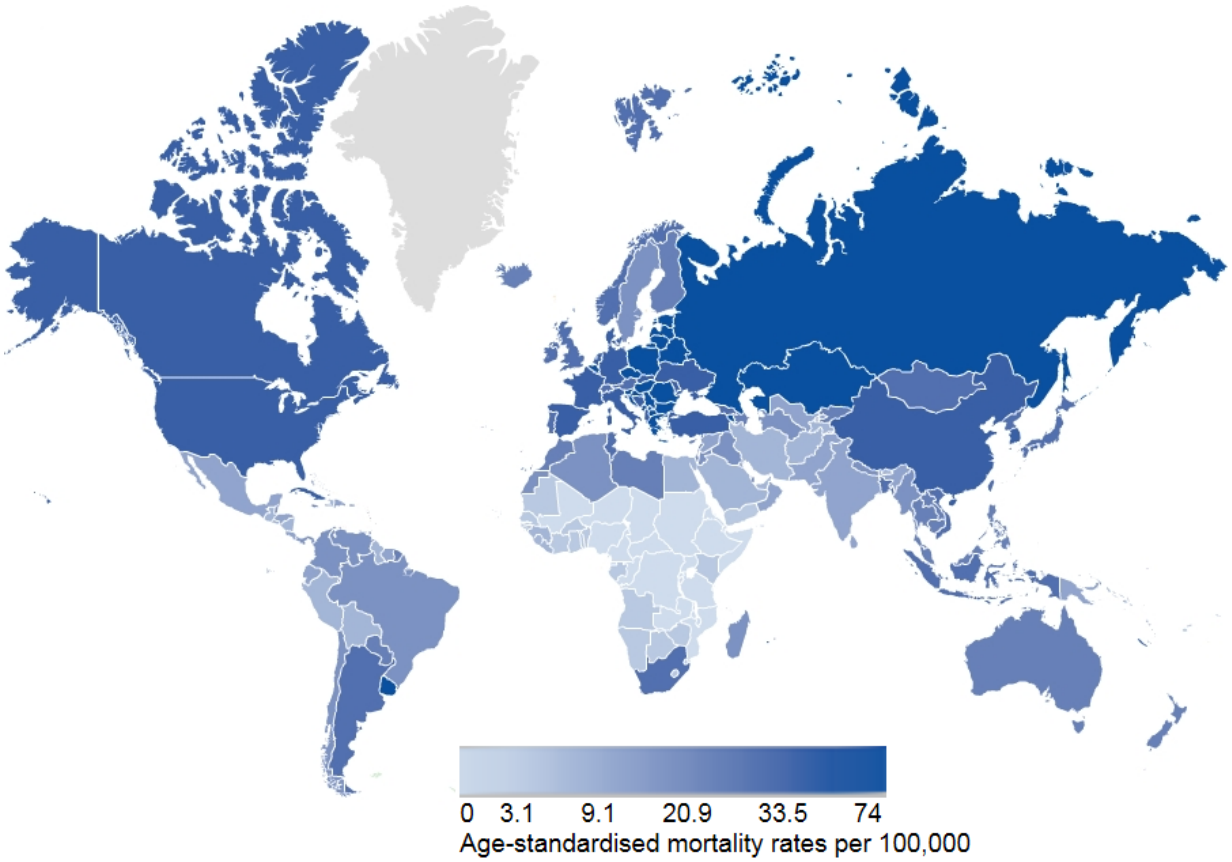


Figure 1.4: Estimated male lung cancer mortality rate per 100 000 (source: GLOBOCAN 2008 (IARC)).

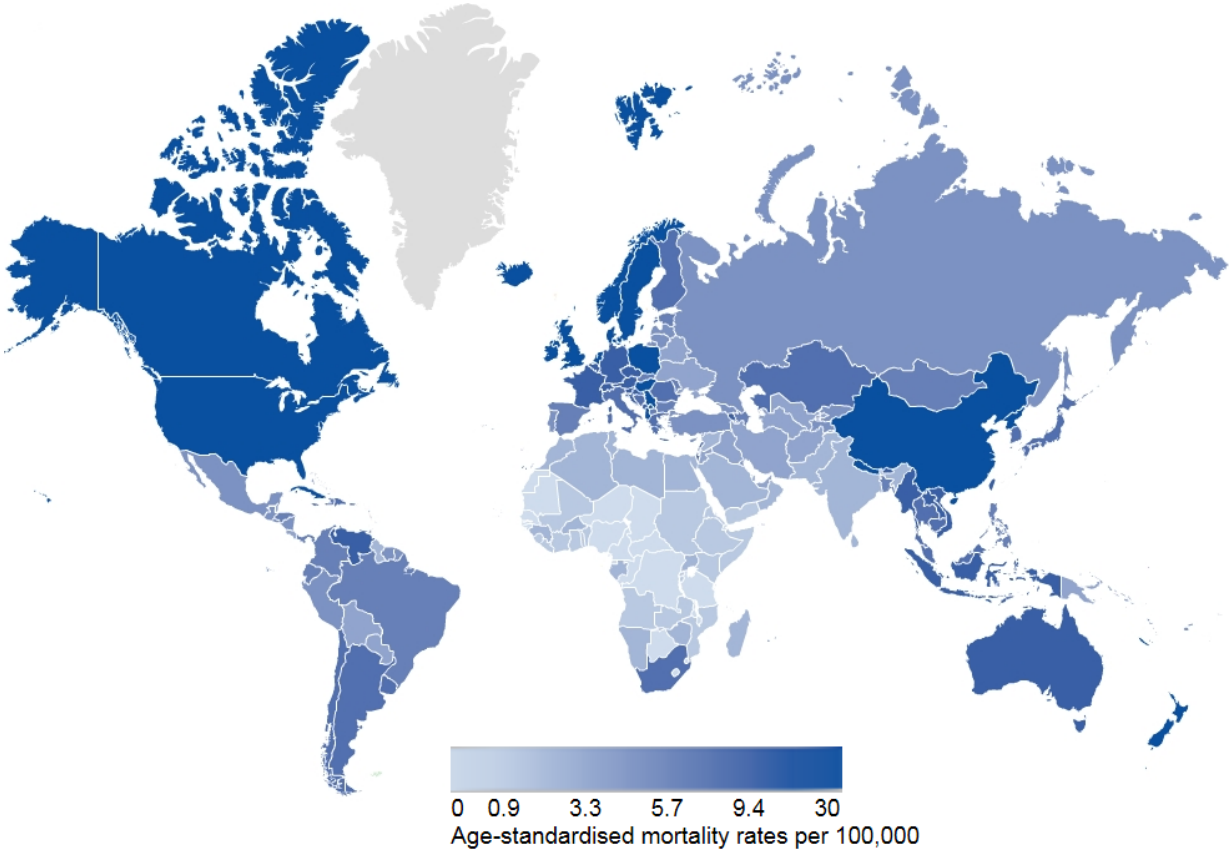


Figure 1.5: Estimated female lung cancer mortality rate per 100 000. (source: GLOBOCAN 2008 (IARC)).

Diagnosis in early lung cancer stage is crucial regarding treatment and survival. While in earlier decades, chest X-rays were applied for screening, nowadays computed tomography (CT) is in the focus as a tool for early lung cancer detection (Keshamouni et al., 2009). Lung cancer can be described by two main types – small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC). The types differ regarding the size of the lung cancer cells. SCLC is the more aggressive type and grows very fast and often metastasized. Around 80–85% of all lung cancer cases are NSCLC, which is less aggressive than the SCLC (Krebsliga Schweiz, 2012). Early detection is essential for the treatment. However, as the symptoms of lung carcinoma in the beginning are not conspicuous, detection in early stages is challenging. Steps before diagnosis cover anamnesis, medical and laboratory examinations, x-ray of the thorax, electrocardiogram and pulmonary function tests. Depending on the results, subsequent examinations, as bronchoscopy, follow for confirmation. Different treatments of lung cancer exist, depending on several factors, i. e. age of the patient or stage of the cancer. The main therapeutic methods are surgery, drug treatment as

chemotherapy and radiotherapy.
At the point of diagnosis, around 75% of all lung cancers are already in an advanced stage or the tumour has metastasized by then. Consequently, survival rates for both genders are very low and there was no significant change in lung cancer survival over the last decades (Keshamouni et al., 2009). In Europe, the lowest survival has been reported for cancer of the lung (beside cancers of the pancreas, oesophagus, brain and liver) (Coleman et al., 2003), where an average 5-year survival of 20% was estimated.
Other risk factors are known – radon is the main risk factor of lung cancer in non-smokers and has been classified as human carcinogen (group 1) by the International Agency for Research on Cancer (IARC) (International Agency for Research on Cancer, 1988). It is a radioactive gas, resulting as an intermediate product of the decay chain of an element in the Earth's crust. Subsequently, radon finds its way through soil and rocks into the houses, where people are exposed to it. In general, its concentration is low and highest concentrations can be found in underground mines (Darby et al., 2001).

1.4 Tobacco-related cancer morbidity and mortality in Switzerland

In Switzerland, smoking causes the largest number of preventable deaths. In 2007, more than 9 000 (6 427 men and 2 774 women) deaths were due to tobacco use, which is 14% of all deaths. Respiratory diseases accounted for 1 669 of those cases, while 3 800 died from cardiovascular diseases. Cancer was responsible for 3 729 smoking-attributable deaths, while 2 485 of them were lung cancer cases (Federal Statistical Office, 2009). Lung cancer is the most common cancer cause of death among men and the second frequent one for women (after breast cancer). Its main risk factor is tobacco smoking, accounting for 85–90% (including passive smoking) (Hirrle, I/2005). Nowadays, around 2 000 males and 900 females die from this cancer every year. About 10% of the cases among males are younger than 55, while the median age is 71. For women the corresponding numbers are 52 and a median of 69 years of age. Since the last three decades, female lung cancer mortality is growing steadily. While in the '80s, the disease-specific standardized rate was 5 per 100 000 inhabitants, it increased more than threefold up to 16 cases (in 2004). On the other hand, male lung cancer mortality peaked in the '80s and decreased since then from 70 to 45 cases per 100 000 in 2004 (Berrut and Junker, 2008).
The country is distinguished into three main languages – German, French and Italian. Assuming that regions defined by the language spoken represent a corresponding cultural

behaviour, this linguistic partition transforms Switzerland into a multicultural unit.
In Switzerland, cancer incidence data are available from the 11 cancer registries, which cover 15 cantons. The largest gap can be found in the German-speaking part of the country, where only 45.2% of its population was covered by cancer registries in 2005–2009 (Federal Statistical Office, 2013a). On the other hand, data on death cases are available at the Federal Statistical Office. The cause of death statistics provide reliable information on the age, gender, civil status, occupation, commune of residence, nationality and cause of death (ICD-10 code) of the reported case. The statistics cover the whole country and lack only by 3% missing cause of death.
Analysis of data at a high geographical resolution allows to assess the significance of risk factors, to produce smooth maps to understand the temporal and spatial distribution of the disease and to identify endemic areas. However, Swiss cancer morbidity and mortality data have not yet been analysed to such extend.

1.5 Bayesian spatio-temporal modelling of cancer data

Spatial data can be broadly classified into three categories viz areal unit data, geostatistical data and point processes (Bailey and Gatrell, 1995). The latter defines data observed at random locations and thus the outcome of interest is the location itself. Areal data arise when the outcome of interest is observed over an area, e. g. mortality rates within administrative regions. In this case the entire study area is partitioned into a finite number of mutually exclusive and exhaustive regions each having a positive area with a measurement attached to it. Geostatistical data arise when the outcome is observed at a finite number of geocoded locations, e. g. village level surveys. Geographical dependency violates the independence assumption of standard statistical models resulting in biased estimates of the parameters if it is not taken into account (Ver Hoef et al., 2001).
Disease maps based on crude rates can be non-informative or misleading when the sizes of the population for some of the areas are small, resulting in large variability in the estimated rates, and making it difficult to distinguish chance variability from real differences.
Bayesian methods have been applied extensively in recent years for modelling spatial data because they allow flexible modelling and inference and provide computational advantages via the implementation of Markov chain Monte Carlo (MCMC) methods (Gelfand and Smith, 1990). The spatial structure is commonly introduced in a hierarchical fashion via the prior distribution of area- or location-specific random effects, although spatial dependence can be built directly on Gaussian response data. The choice of prior distributions

or spatial models depends on the type of data.
In areal data, simultaneously autoregressive (SAR) models (Whittle, 1954), conditional autoregressive (CAR) models (Clayton and Kaldor, 1987; Bernardinelli and Montomoli, 1992; Gelfand and Vounatsou, 2003; Best et al., 2005) and modifications (Besag et al., 1991; Marshall, 1991; Sun et al., 2000) have been suggested as prior specifications.
The models *smooth* or improve the estimate of an unstable rate by borrowing strength from its neighbours. They enable smoothing of the rates to highlight patterns but also assess the significance of risk factors taking into account the geographical correlation. In addition, these models are able to show spatial patterns after adjusting for geographical differences in certain risk factors. By adding a time dimension, Bayesian spatio-temporal models indicate changes of geographical patterns over time and determine how the disease evolves over time in different regions and groups of the population (age or gender-related groups).

1.5.1 Cancer mapping

Graphical presentation of the distribution a disease or so-called disease mapping is essential in spatial epidemiology. It allows to assess the spatial as well as temporal distribution of a disease under consideration and identifies high/low risk areas. Cancer mapping has been applied frequently in the last decades for different cancer sites, countries and geographical units.
Several measures exist to express disease-specific incidence or mortality. Most commonly used quantities are the Age Standardized Rate (ASR) and the Age Standardized Incidence (SIR) or Mortality Ratio (SMR). Both are differing in the way of standardization, which can be distinguished as *direct* and *indirect*. In the case of direct standardization, common reference populations are the Segi world population (Segi, 1960) or the European (Scandinavian) one (Doll and Cook, 1967). The approach is to weight observed age group specific mortality rates by the age distribution of the reference population. In the case of indirect standardization, the overall area under consideration is used as the reference group. Observed sub-regional age group-specific mortality rates are then related to the one expected based on the overall scale, i. e. countrywide. While direct standardization allows for comparison between several countries, the indirect approach gives a clear idea of the distribution within a country in an absolute sense.
Geographical information systems (GIS) software has been used frequently for handling and illustrating geographically referenced information in public health and epidemiology.

Since the last decades numerous investigations have been done on cancer mapping studies applying Bayesian hierarchical spatio-temporal models. First analyses include work by Xia et al. (1997), mapping lung cancer mortality rates in Ohio based on Bayesian spatial models. As a follow-up, these models were developed further by Waller et al. (1997) and Sun et al. (2000) to assess space-time patterns of lung cancer rates.
Knorr-Held and Best (2001) introduced a shared component model in 2001, which allows for identification of similar geographical patterns in multiple cancer sites in a joint spatial analysis. These models aim to distinguish between the unobserved shared component for all cancers (e. g. smoking) and the component related to the disease.
Richardson et al. (2006) jointly mapped lung cancer rates for males and females to extract common risk factors. The authors assessed the temporal evolution of the gender-specific risk factors and studied weather there is evidence of a differential risk between males and females beyond that expected from tobacco use. Another joint model approach was presented by Downing et al. (2008), mapping lung, oesophagus, stomach, pancreas, bladder and kidney cancer incidence rates in England. Numerous other applications of Bayesian space-time modelling on mapping cancer related outcomes are available.
Comprehensive data on Swiss cancer mortality are available from death certificates, which have been collected and coded by the Federal Office of Statistics since the end of the $19th$ century, but they are available in electronic form (indispensable for analysis) since 1969. They are geo-referenced and cover the whole country. Studies have been conducted to obtain crude cancer incidence and mortality estimates including lung cancer and other tobacco-related cancers by region and selected years during 1986–2005 (Luthi et al., 2005; Pury et al., 2007a,b). Few studies have explored temporal trends in lung cancer mortality (La Vecchia et al., 1988; Levi et al., 2006). Detailed analyses to assess spatio-temporal patterns of site-specific cancers and their associated risk factors are scarce. The only map of cancer mortality rates including lung and other tobacco-related cancers is that of Schüler and Bopp (1997) depicting geographical variation in mortality during 1970–1990 in a rather descriptive way.
So far, no covariate-adjusted and smooth nationwide maps of tobacco-related cancer mortality and incidence rates, determining the changes over time in Switzerland, exist. The wealth of information provided in these databases has not been fully explored to assess space-time patterns and trends of lung and tobacco-related cancer at different administrative levels in Switzerland.

1.5.2 Projection of cancer mortality

Cancer projections aim to assess past trends of a disease-specific incidence or mortality and to predict them into the future. Projections might be done for short- (≤ 5 years) or long-term, both requiring different analyses and modelling approaches.

The most commonly used approach for cancer projections are age-period-cohort (APC) models. Within this method, the dataset is stratified into *age*, indicating the age group, *period*, indicating in which time period the death occurred and *cohort*, giving information on the birth cohort of the person died. The relationship of the three components *age* - *period* = *cohort* results in the well-known identification problem (Bray, 2002). Many solutions to deal with this problem have been proposed. However, they are often based on strong assumptions. Projected mortality rates are not influenced by this identification problem.

Two main approaches have been well-established – the power model implemented in the Nordpred software package developed by the cancer registry in Norway (Møller et al., 2003) and on the other side the Bayesian APC model, mainly driven by the work of Bray et al. (Bray, 2002; Baker and Bray, 2005). Both methods are based on the classical APC approach, assessing and extrapolating the effects of age, period and cohort. Main differences are the statistical inference framework the models are built in and, as a more methodological difference, the link-function applied for regressing the effects on the outcome of interest.

1.5.3 Back-calculation of cancer incidence

Data on mortality are available from Statistical Offices, which often compile information from death certificates at national level. Reliable data on incidence can be obtained from cancer registries, which often do not cover the whole country. In Switzerland, local cancer registries exist at cantonal level, however, they only cover half of the Swiss population. Registration of cancer cases is incomplete mainly for the German-speaking cantons in the middle of the country. Furthermore, the cancer registries have been established at different time points and therefore may only cover short time periods.

Different approaches have been developed to estimate incidence from mortality data in regions without registries. A Bayesian back-calculation model has been proposed (Mezzetti and Robertson, 1999), linking mortality data with survival probabilities to estimate incidence in the past. The model considers a fixed one-year survival, followed by an exponentially distributed survival up to five years. The incidence-mortality ratio is an alternative

approach (Colonna et al., 1999) to estimate incidence based on observed mortality.

1.5.4 Indirect approximation of smoking patterns in space

Information on gender-specific smoking prevalence is rather sparse. Frequently, survey data are not sufficient to assess space-time patterns of smoking, which is essential to estimate the burden of tobacco use. In Switzerland, smoking data are provided from the Swiss Health Survey (SHS), which was initiated in 1992 and conducted every five years. Another source is the Tobacco Monitoring Switzerland, a survey on tobacco consumption among those aged 14–65, which was originated in 2001. Both sources have their limitations as they do not cover the whole country and do not provide data from the past to assess longterm trends and effects.
Lung cancer rates have been used as an indicator of tobacco use (Peto et al., 1992; Chellini et al., 2006). This approach only depends on the availability of information about lung cancer mortality, which is often provided by death certificates.

1.6 Objectives of the thesis

The overall objective of this research was to identify spatio-temporal patterns in tobacco-related cancer in Switzerland and forecast future trends in cancer mortality in the country.

1.6.1 Specific objectives

The specific objectives of this thesis were:

(i) Assess geographical differences and trends of age- and gender-specific lung and all tobacco-related cancer mortality in Switzerland (Chapter 2);

(ii) Project tobacco-related cancer mortality in Switzerland at different geographical levels accounting for spatial variation (Chapter 3);

(iii) Develop Bayesian APC-models for projecting cancer mortality data (Chapter 4);

(iv) Develop Bayesian back-calculation models to estimate age- and gender-specific incidence from sparse mortality data (Chapter 5);

(v) Develop models to indirectly approximate gender-specific smoking patterns in space and time by adjusting lung cancer mortality rates with non-smoking risk factors (Chapter 6).

Chapter 2

Bayesian spatio-temporal modelling of tobacco-related cancer mortality in Switzerland

Verena Jürgens[1,2], Silvia Ess[3], Harish C. Phuleria[1,2], Martin Früh[4], Matthias Schwenkglenks[5], Harald Frick[6], Thomas Cerny[4], Penelope Vounatsou[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Cancer Registry of St. Gallen & Appenzell, St. Gallen, Switzerland
[4] Medical Oncology-Hematology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland
[5] Institute of Pharmaceutical Medicine (ECPM), University of Basel, Basel, Switzerland
[6] Cancer Registry Grisons & Glarus, Kantonsspital Graubünden, Chur, Switzerland

This paper has been published in *Geospatial Health* 2013, 7(2):219–236.

Abstract

Tobacco smoking is a main cause of disease in Switzerland; lung cancer being the most common cancer mortality in men and the second most common in women. Although disease-specific mortality is decreasing in men, it is steadily increasing in women. The four language regions in this country might play a role in this context as they are influenced in different ways by the cultural and social behaviour of neighbouring countries. Bayesian hierarchical spatio-temporal, negative binomial models were fitted on subgroup-specific death rates indirectly standardized by national references to explore age- and gender-specific spatio-temporal patterns of mortality due to lung cancer and other tobacco-related cancers in Switzerland for the time period 1969–2002. Differences influenced by linguistic region and life in rural or urban areas were also accounted for. Male lung cancer mortality was found to be rather homogeneous in space, whereas women were confirmed to be more affected in urban regions. Compared to the German-speaking part, female mortality was higher in the French-speaking part of the country, a result contradicting other reports of similar comparisons between France and Germany. The spatio-temporal patterns of mortality were similar for lung cancer and other tobacco-related cancers. The estimated mortality maps can support the planning in health care services and evaluation of a national tobacco control programme. Better understanding of spatial and temporal variation of cancer of the lung and other tobacco-related cancers may help in allocating resources for more effective screening, diagnosis and therapy. The methodology can be applied to similar studies in other settings.

Keywords: Bayesian inference; Conditional Autoregressive models; disease mapping; lung cancer mortality; Markov chain Monte Carlo simulation; tobacco smoking; Switzerland

2.1 Introduction

In Switzerland, lung cancer is the first cause of cancer mortality in men and the second in women (after breast cancer), accounting for 2 900 deaths per year (Berrut and Junker, 2008). In 1970, the mortality rate per 100 000 inhabitants was around 60 cases for males, whereas it was only five for females but, by 2004, it had decreased to 45 for males and increased to 16 for females (Berrut and Junker, 2008). Also, the geographical distribution of the disease differs remarkably by gender with female lung cancer showing high concentrations in the cities, whereas male lung cancer is clustered in the central, more rural and industrialized part of the country (Schüler and Bopp, 1997). Tobacco smoking is the main cause (Tyczynski et al., 2003), but can also result in cancers of the larynx, oral cavity, pharynx and oesophagus (International Agency for Research on Cancer, 2004). Even passive exposure to tobacco smoke raises the risk of disease (Subramanian and Govindan, 2007). The Swiss Federal Office of Public Health (FOPH) launched a tobacco prevention programme for the period 2001–2008, later extended to also cover the following years (2009–2012). The main objective of this programme is to reduce the number of tobacco-related deaths and diseases in Switzerland. To achieve appropriate public health decisions, differences in regions influenced by certain factors, e. g. gender and whether people live in urban or rural neighbourhoods, have to be identified. Maps of the spatial patterns and trends of lung cancer and other tobacco-related cancers can assist the national tobacco programme by identifying high-priority areas for particularly intensive anti-tobacco campaigns.

Over the last 15 years, Bayesian hierarchical spatio-temporal models have been widely used for disease mapping (Sue Bell and Broemeling, 2000; Xia et al., 1997; Johnson, 2004). This methodology enables the estimation of covariate-adjusted smooth maps highlighting patterns of the disease and exploring covariate effects. In Switzerland, only a few studies exploring the temporal differences and trends in lung cancer mortality have been published (La Vecchia et al., 1988; Levi et al., 2006). Available estimations of geographical disparities are limited to raw rates at the regional level and cover only selected years (Luthi et al., 2005; Pury et al., 2007a,b; Bouchardy et al., 2011). Mapping the raw rates of the disease mortality can be non-informative and even misleading by overestimating the mortality in small populations (Clayton and Kaldor, 1987).

The aim of this paper was to explore age- and gender-specific spatio-temporal patterns of lung cancer and other tobacco-related cancer mortality rates in Switzerland using Bayesian hierarchical spatio-temporal models. We also wanted to use this approach to determine differences between rural or urban living and, as Switzerland is a multilingual country, extend the investigation into age- and gender-specific mortality by linguistic region.

2.2 Material and methods

2.2.1 Data sources

Mortality data were provided by the Federal Statistical Office (FSO). The data covered yearly death records during 1969–2008 at the individual level. Information about year of birth, age at death, sex, municipality of residence, nationality and cause of death was available. During the period covered in this study, there was a transition from the 8th revision of international classification of diseases (ICD-8) to ICD-10 in the period 1994/1995. We therefore chose to standardize the whole database according to ICD-10.

Our analysis was separated into two parts. On the one hand we focused exclusively on lung cancer mortality as its smoking-attributable fraction is very high and therefore reflects trends in smoking behaviour. On the other hand we analysed mortality of all tobacco-related cancers, including lung cancer. Following the definition of Doll et al. (1976), the following cancers were considered as tobacco-related: lung (C33-C34), oesophagus (C15), rectum (C20), pancreas (C25), bladder (C67) and cancer of other respiratory sites (C00–C14), including cancer of the lip, tongue, mouth, larynx, trachea and pharynx (excluding nasopharynx).

Population size by age, sex and municipality was obtained from the Swiss census. The municipality was identified by name and the given FSO municipality number. During the study period, a census was carried out every 10 years starting in 1970. Due to the missing population data for the interim years, we divided the study into four periods, each covering the four years closest to the census year, i. e. 1969–1972, 1979–1982, 1989–1992 and 1999–2002. The population size was assumed to be constant within each period. The four main language regions of Switzerland (German, French, Italian and Romansh – see Fig. 2.1) were taken into account because the language can be used as proxy for culture and behaviour. Language information at the municipality level was obtained from the FSO.

Digital information on the structure of the municipality boarders of Switzerland in 2009, given by geospatial data in shapefile format, was extracted from the FSO website (Federal Statistical Office, 2012b). Municipalities were considered as rural if the population density was less than 150 people per km^2 and otherwise as urban, following the definition of the Organisation for Economic Co-operation and Development (OECD) (Schrader, 1997). Information on the geographical size of each municipality was extracted from the shapefile offered by the FSO. The current number of Swiss municipalities amounts to about 2 700, more than 1 500 of which have less than 1 000 inhabitants. The number of municipalities with population size greater than 10 000 is less than 120 (census of 2000). There have been numerous fusions and partitions of municipalities over the time of the study.

The death records of the mortality dataset were updated by the FSO to fit the municipality structure of October 2009. To align the data spatially, we converted the population data of the last

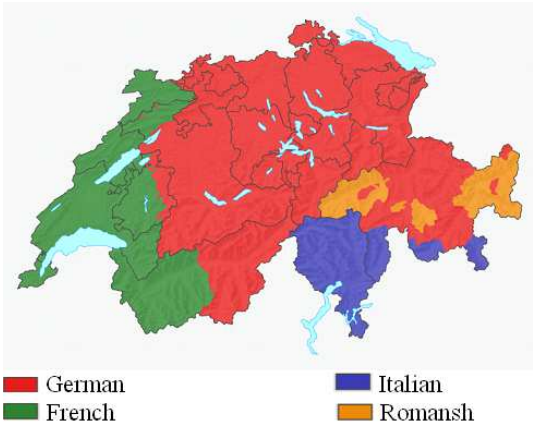


Figure 2.1: Language regions of Switzerland in the year 2000 (source: (Federal Statistical Office, 2012a)).

four censuses into this municipality structure. Municipalities without population were identified and merged with one of their immediate neighbours. The shapefiles were modified in R (R Development Core Team, 2011) to restructure the municipality boarders.

2.2.2 Statistical Analysis

Spatio-temporal models for all tobacco-related cancer mortality rates were fitted separately for each gender and age group. Age was grouped in two categories: middle aged (35–59 years old) and elderly ($\geq$60 years old) according to Peto et al. (1992). Mortality before age 35 years was excluded. Time trends were modelled as a categorical covariate with four levels corresponding to the periods: 1969–1972, 1979–1982, 1989–1992 and 1999–2002.

Multivariate negative binomial regression models were used to 1) assess the patterns of age- and gender-specific lung mortality due to cancer and other tobacco-related cancers in space and time; and 2) determine differences by language and type of dwelling (urban or rural). Spatial correlation was considered by introducing random effects at the municipality level. Models were formulated following a Bayesian inferential framework, considering age- and gender-specific interactions of time periods with language region and type of dwelling.

Mortality rates are often standardized to compare observed and expected rates with reference to a standard population (Oberaigner and Vittadello, 2010). Standardisation may be direct or indirect (Breslow and Day, 1987), the latter usually done by calculating the standardized mortality ratios (SMRs) (Johnson, 2004). This is recommended if the mortality variation within-country is of primary interest (Estève et al., 1994; PAHO, 2002). In this case, age- and gender-specific SMRs were used to measure mortality. Standardisation over the entire study population was not appropriate because of large absolute differences in disease occurrence. Therefore, indirect

standardisation was done using the subgroup-specific national death rate averaged across the whole study period from 1969 to 2002. However, this meant that the estimates of the temporal trend for each period had to be interpreted in relation to the remaining periods for the specific age group and gender. Since the resulting subgroup maps cannot be compared directly because of the specific standardisation for each subgroup, a separate analysis was carried out for the subset of mortality data during the 1999–2002 period, standardizing only over the corresponding time. To take into account disparities in age distributions among municipalities, all SMRs were additionally standardized by 5-year age groups, starting with age group 35–39 years (Sturtz and Ickstadt, 2007). An SMR of 1.0 implies that there is no difference between the mortality rate of the municipality under consideration and the national rate for the specific subgroup. An SMR that is lower or higher than 1.0 stands for a mortality rate that is lower or higher than the national average.

We assumed the observed age- and gender-specific counts of lung and all tobacco-related cancer deaths Y_{ijlt} in municipality i ($i = 1, \ldots, N$), gender j ($j = 1, 2$) and age group l ($l = 1, 2$) in period t to follow a negative binomial distribution $Y_{ijlt} \sim \text{NegBin}(p_{ijlt}, r)$ with parameters p_{ijlt} and r, where p_{ijlt} relates to average number of cases via the formula $p_{ijlt} = r/(r + \mu_{ijlt})$. Age- and gender-specific spatial random effects as well as possible non-linear trends were modelled on the log of the mean age SMR following model formulations according to (Xia et al., 1997):

$$log(\mu_{ijlt}) = log(E_{ijlt}) + \alpha + X_{is}^T \beta_s + \phi_{ijl} \tag{2.1}$$

where E_{ijlt} is the age- and gender-specific expected number of deaths, X_{is} the vector of covariates s related to municipality i and β_s the coefficients of associated covariates.

Spatial correlation was introduced by age- (l) and gender-specific (j) random effects ϕ_{ijl} at municipality level (i), modelled via a Conditional Autoregressive (CAR) process (Besag et al., 1991; Bernardinelli and Montomoli, 1992). Spatial dependency among the municipalities was introduced by the conditional prior distribution of the age- and gender-specific ϕ_i with

$$\phi_i \mid \boldsymbol{\phi}_{-i} \sim \mathcal{N}\left(\frac{\gamma \sum_{\substack{q=1 \\ q \neq i}}^{N} c_{iq}\phi_q}{w_i}, \frac{\sigma^2}{w_i}\right) \tag{2.2}$$

where $\boldsymbol{\phi}_{-i} = (\phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_N)$ and c_{iq} indicates the degree of spatial influence of municipality i to the remaining municipalities, taking the value 1 if they are adjacent and 0 otherwise, and γ quantifying the overall spatial dependence, assumed to be 1 following Besag et al. (1991), and w_i being the number of neighbours of municipality i. As the Bayesian formulation requires, specification of prior distributions for the model parameters were specified. The variance parameter r, controlling overdispersion, was assumed to follow a gamma prior distribution. We assumed

Year	Rate*	Male	Female
1970	68	1 568 (90%)	175 (10%)
1980	77	2 111 (87%)	311 (13%)
1990	76	2 245 (82%)	491 (18%)
2000	69	2 044 (72%)	781 (28%)

Table 2.1: Lung cancer mortality in Switzerland: death counts and proportion of males and females. (*Per 100 000 inhabitants)

a flat prior for the intercept α and uniform prior distributions $U(-\infty, \infty)$ for the regression coefficients β_s.

After a burn-in of 1 000 iterations the model was run for another 100 000 iterations achieving convergence, which was assessed by graphical inspection of the Markov chain output. Finally another 10 000 iterations were run to save a sample of 1 000 with thinning 10 to obtain summaries of the posterior distribution of the parameters.

Preliminary analyses fitting negative binomial models were carried out in Stata/IC (Stata Corporation; College Station, USA). Bayesian models were fitted using OpenBUGS (Imperial College and MRC, London, UK). The resulting maps of all tobacco-related cancer mortality rates and results of the interaction analysis are given in the Apppendix (Figs. 2.8–2.13 and Tables 2.5–2.6). The Kulldorff's spatial scan statistic implemented in the DCluster package within the R software was used to detect significant mortality clusters in the most recent period (Kulldorff and Nagarwalla, 1995; Gómez-Rubio et al., 2005).

2.3 Results

2.3.1 Lung cancer mortality

Males

Table 2.1 reports a decline of lung cancer deaths in males between 1990 and 2000. Table 2.2 shows the results of the spatial negative binomial regression analysis carried out on the mortality data over the whole study period. Lung cancer SMR in younger men (35–59 years old) decreased by 40% during 1999–2002 in comparison to 1969–1972. This trend is illustrated in Figure 2.2. A similar result has not been confirmed for older men (≥ 60 years). On the contrary, SMR increased by around 25% in the 1980s and 1990s and was not significantly different in 1999–2002, compared to the 1970s. Also, results indicate no significant differences in lung cancer mortality between the language regions (Table 2.2, Figs. 2.2–2.3).

The SMR was 7% higher for older men in urban areas. Models including interaction terms (Table 2.5) indicate that the mortality in younger men had decreased everywhere in Switzerland

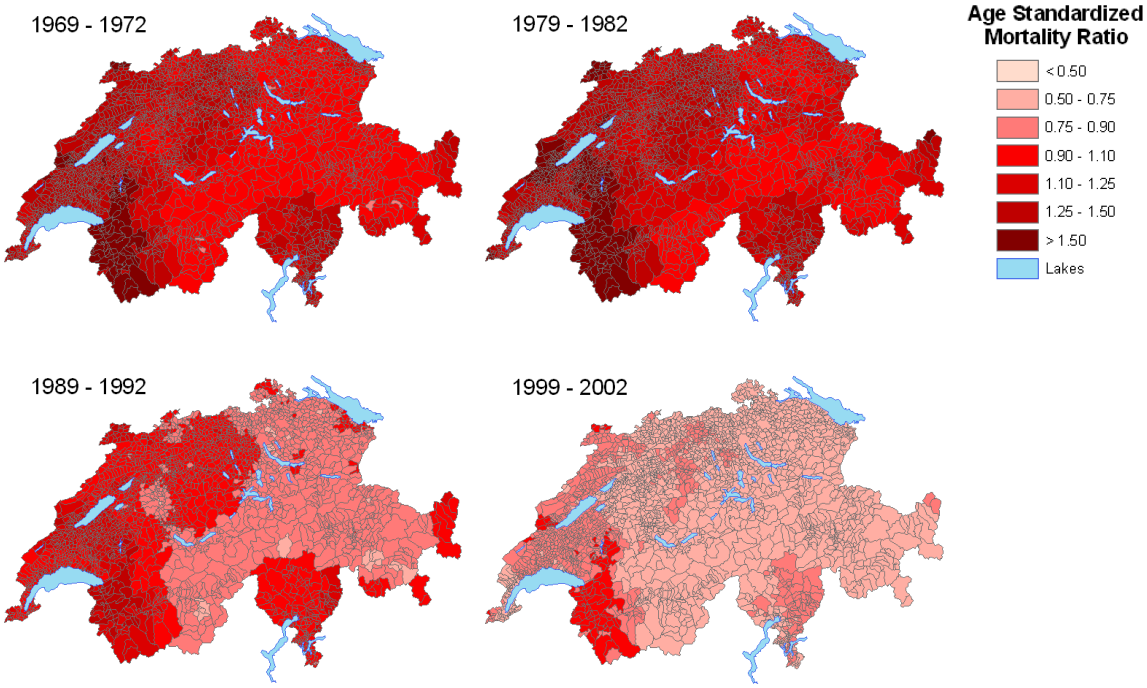


Figure 2.2: Geographical distribution of lung cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for males aged 35–59 years.

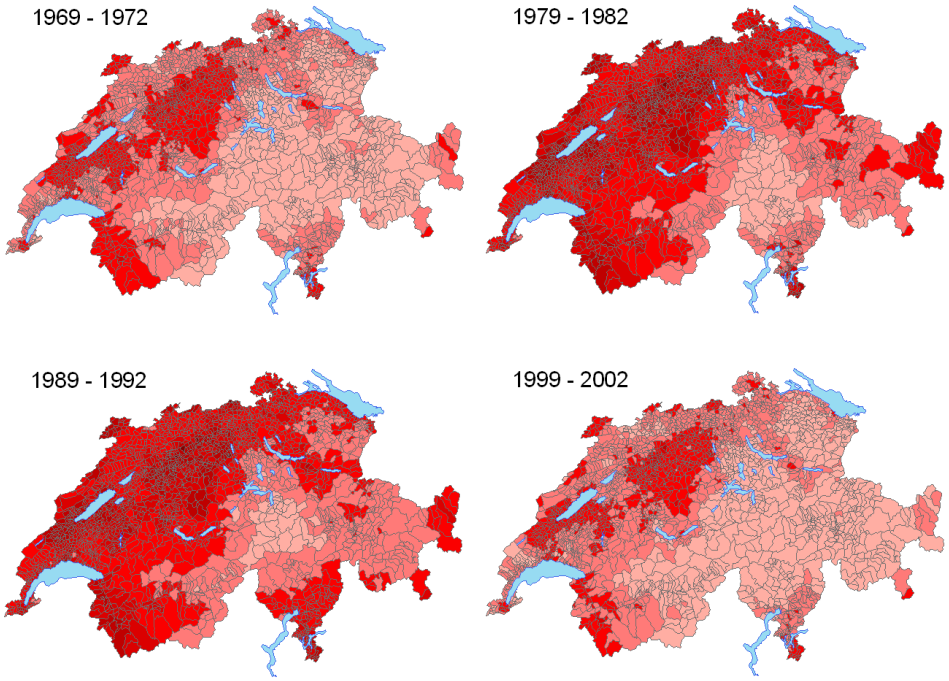


Figure 2.3: Geographical distribution of lung cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for males aged 60 years and above.

	Male		Female	
Age group (years)	35–59	≥ 60	35–59	≥ 60
	SMR Ratio (95% BCI*)			
Period				
1969–1972	1.00	1.00	1.00	1.00
1979–1982	1.07 (1.00; 1.15)	1.26 (1.20; 1.31)	1.64 (1.31; 2.05)	1.32 (1.16; 1.50)
1989–1992	0.81 (0.75; 0.88)	1.23 (1.18; 1.29)	2.48 (2.03; 3.04)	1.78 (1.58; 2.01)
1999–2002	0.60 (0.55; 0.64)	0.96 (0.92; 1.01)	3.50 (2.87; 4.28)	2.52 (2.25; 2.86)
Language				
German	1.00	1.00	1.00	1.00
French	1.16 (0.99; 1.35)	1.07 (0.94; 1.17)	1.36 (0.98; 1.74)	1.30 (1.07; 1.59)
Italian	1.25 (0.93; 1.66)	1.15 (0.82; 1.41)	0.93 (0.58; 1.39)	1.23 (0.71; 2.06)
Romansh	0.95 (0.58; 1.44)	0.94 (0.72; 1.25)	0.76 (0.23; 1.87)	0.54 (0.25; 1.01)
Dwelling				
Rural	1.00	1.00	1.00	1.00
Urban	0.97 (0.91; 1.03)	1.07 (1.03; 1.12)	1.20 (1.04; 1.36)	1.28 (1.18; 1.40)
Spatial variation	0.31 (0.26; 0.37)	0.31 (0.27; 0.35)	0.39 (0.28; 0.49)	0.35 (0.29; 0.42)

Table 2.2: Spatio-temporal model estimates of age- and gender-specific lung cancer mortality in Switzerland from 1969–1972 to 1999–2002. (*Bayesian credible interval)

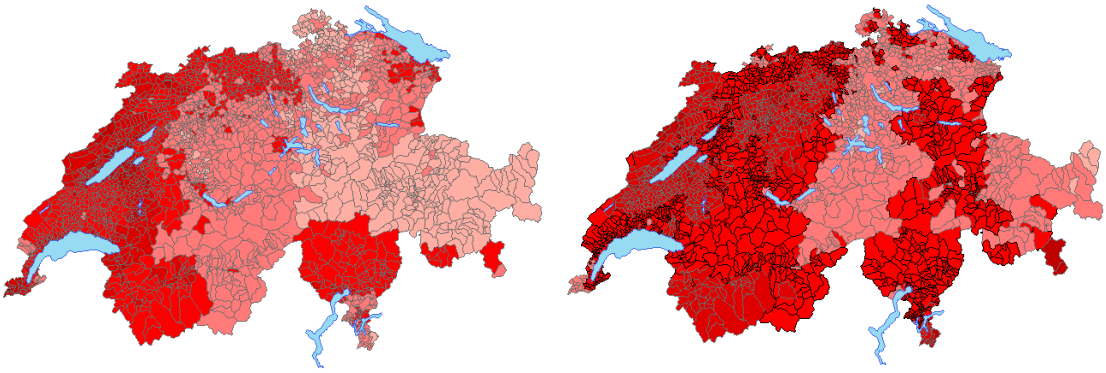
by the year 2000, with the exception of the Romansh-speaking and urban French-speaking regions. Older males showed an increasing trend in the rural parts but the urban areas showed no difference in SMR except for the German ones, where SMRs decreased between the 1970s and 2000. Figure 2.4 illustrates the most recent period (1999–2002) revealing a lower lung cancer mortality in the alpine regions. While the highest mortality rate for young men was concentrated in the rural French-speaking region, high mortality in older men was more common in the generally rural German-speaking and Italian-speaking parts.

Females

Raw data (Table 2.1) indicate that female mortality increased between 1970 and 2000. Also, Bayesian spatial models suggest an increasing trend during the last 30 years. In particular, the SMR of younger females was 3.5 times higher in 2000 than in the 1970s. The SMR for older women was 2.5 times higher in 1999–2002 compared to 1969–1972 (Table 2.2). The SMR in urban regions was more than 20% higher than in the rural areas (Fig. 2.5, 2.6). In the French-speaking part, older females had significantly higher mortality compared to the German one (SMR ratio of 1.30). The lowest mortality was estimated in the Romansh part although the difference seen is not statistically significant, probably due to the small number of deaths (0.13% out of the total

(a) Male, age 35–59 years

(b) Male, age ≥ 60 years



(c) Female, age 35–59 years

(d) Female, age ≥ 60 years

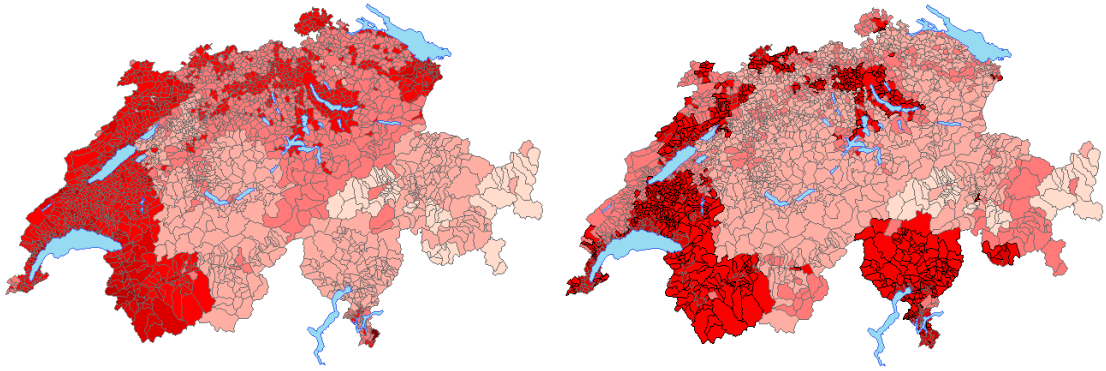


Figure 2.4: Geographical distribution of age- and gender-specific lung cancer mortality risks in Switzerland in the period 1999–2002.

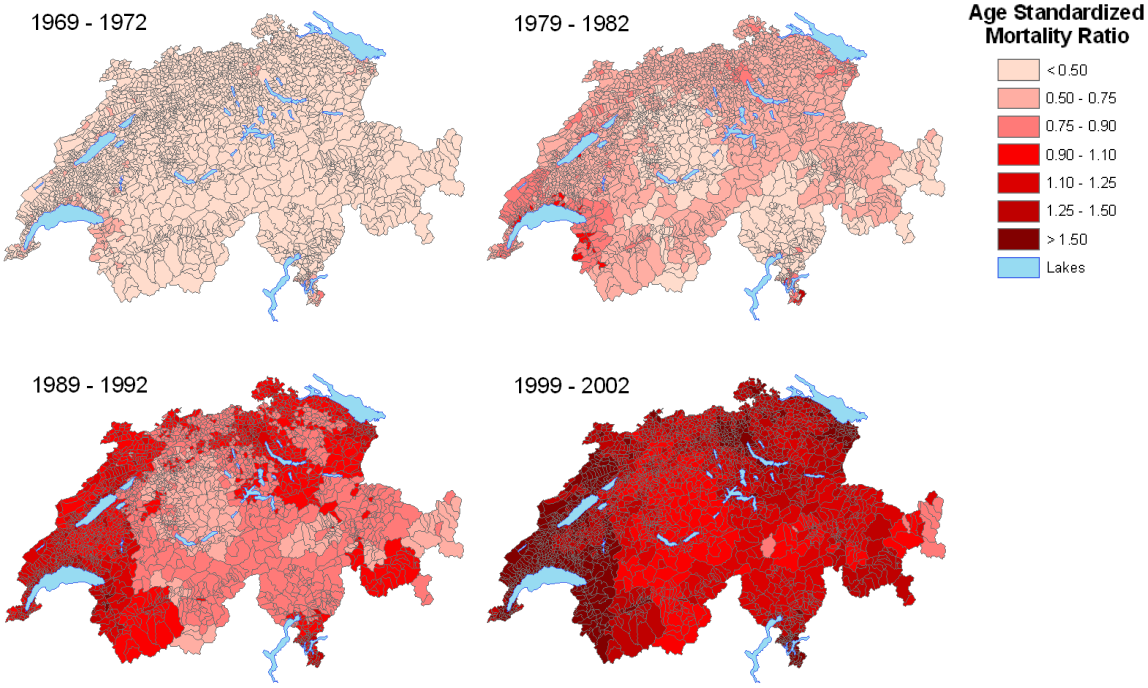


Figure 2.5: Geographical distribution of smooth lung cancer mortality in Switzerland from 1969–1972 to 1999–2002 for females aged 35–59.

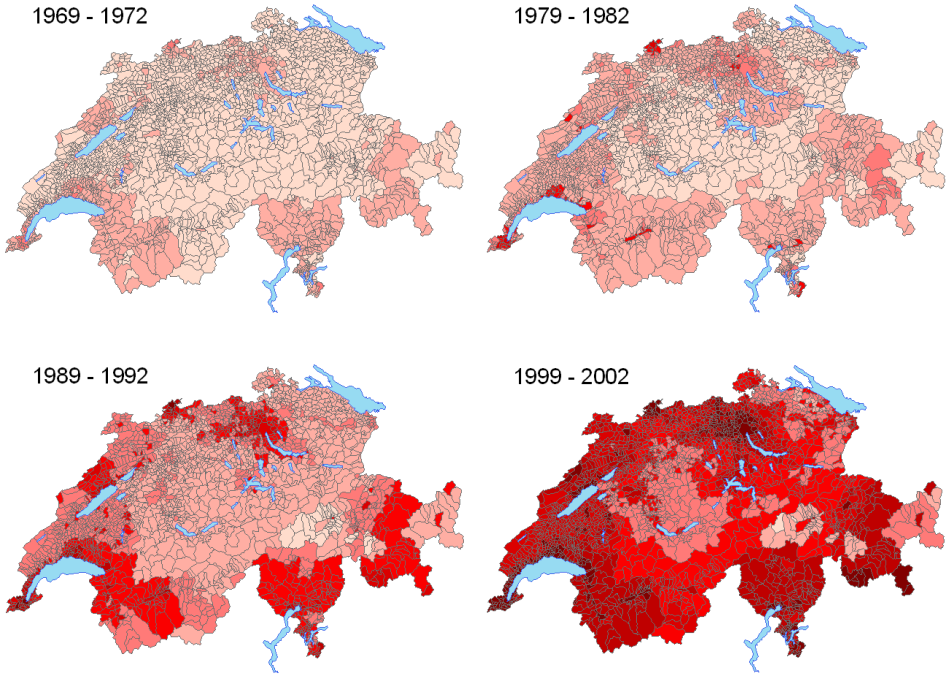
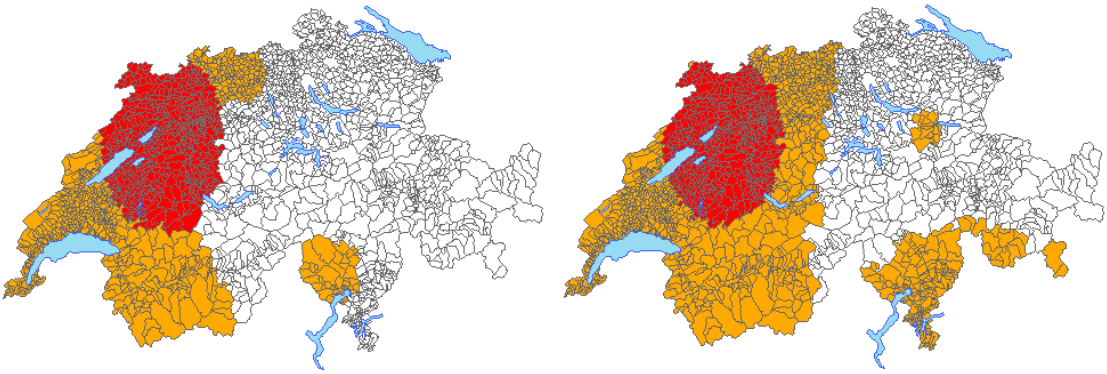


Figure 2.6: Geographical distribution of lung cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for females aged 60 years and above.

(a) Male, age 35–59 years

(b) Male, age ≥ 60 years



(c) Female, age 35–59 years

(d) Female, age ≥ 60 years

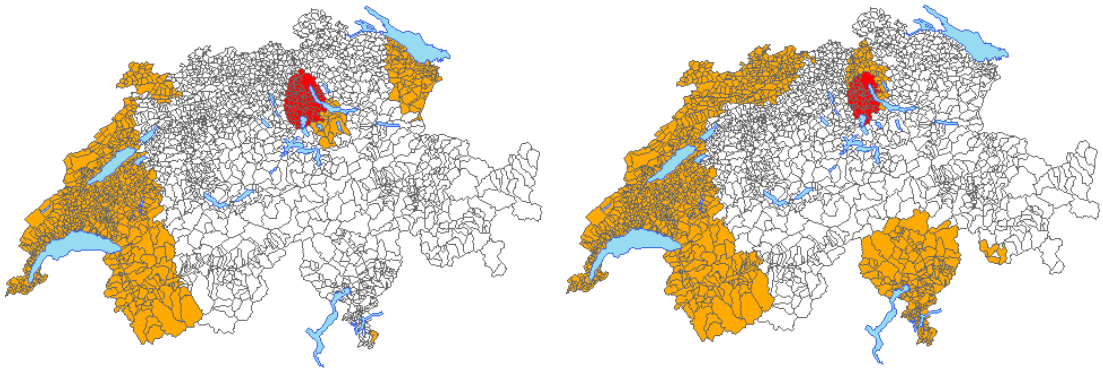


Figure 2.7: Clusters of age- and gender-specific lung cancer mortality risks in Switzerland in the period 1999–2002 (according to Kulldorff's statistic; red colour = highly significant clustering; yellow colour = significant at lower level).

number for females in the year 2000). While the SMR increased from the 1970s to 2000 in all remaining regions, it was not significantly different in 2000 in the rural Italian region. However, this result was not seen in older women. Lung cancer mortality increased in the French-speaking and urban German-speaking regions for the older age group during the 1990s and in the remainder of regions during 2000. Figure 2.4 shows higher mortality in the French-speaking area in 2000, whereas younger females showed the highest mortality in the rural part of this region.

Figures 2.2–2.3 and 2.5–2.6 and 2.8–2.11 show the spatial pattern of lung and all tobacco-related cancer mortality separately for each age and gender group over the four time periods investigated. The maps presented in Figure 2.4 and 2.12 show the distribution of lung cancer and all tobacco-related cancer mortality for each subgroup in 1999–2002. The estimates have to be interpreted independently for each subgroup. Figures 2.7 and 2.13 show the most (red) to the least (yellow) significant clusters of modelled lung and all tobacco-related cancer mortality in 1999–2002 detected by Kulldorff's spatial scan statistic.

Year	Rate*	Male	Female
1970	130	2 947 (79%)	790 (21%)
1980	152	3 747 (78%)	1 050 (22%)
1990	150	3 924 (73%)	1 440 (27%)
2000	130	3 571 (67%)	1 781 (33%)

Table 2.3: Tobacco-related cancer mortality in Switzerland: death counts and proportion of males and females. (*Per 100 000 inhabitants)

2.3.2 Tobacco-related cancer mortality

Males

The estimates of the Bayesian spatial models regarding all tobacco-related cancer mortality (including lung cancer) indicate a decrease of 33% for those 35–59 years old and 20% for those aged 60 and above. For younger males a 21% significantly higher SMR was observed for the French-speaking region in comparison to the baseline German (Table 2.4 and Fig. 2.8). The SMR in older men did not differ between the language regions. Models including interaction terms showed an overall decrease for younger men in 2000 with exception of the urban French part. For the older age group an overall decrease was estimated, except for the rural, French-speaking and Italian-speaking areas.

Females

Mortality increased steadily over the last 30 years. The SMR nearly doubled for the 35–59 age group and increased by 15% for the older women from 1969–1972 to 1999–2002. There were no significant differences between the language regions over the entire study period (Table 2.4). In comparison to the rural areas, mortality in the urban regions was significantly higher both for younger and older women – 14% and 11%, respectively (Figs. S3–S4). In the 1990s, the group aged 35–59 years showed a significantly higher mortality in the German-speaking and rural French-speaking areas. In 1999–2002 an overall increase was observed, except in the rural Italian-speaking areas. In the older age group, there was no significant interaction between time period and language or urban groups observed.

2.4 Discussion

This study explores the spatio-temporal patterns of lung and tobacco-related cancer mortality in Switzerland, for the first time based on state-of-the-art Bayesian conditional autoregressive models. Subgroups defined by age and gender were studied separately. The results show the steep and constant rise, especially in urban regions, of mortality in women due to lung cancer and other

	Male		Female	
Age group (years)	35–59	≥ 60	35–59	≥ 60
	SMR Ratio (95% BCI)			
Period				
1969–1972	1.00	1.00	1.00	1.00
1979–1982	1.08 (1.01; 1.14)	1.10 (1.06; 1.13)	1.26 (1.12; 1.42)	1.07 (1.00; 1.13)
1989–1992	0.89 (0.84; 0.95)	1.05 (1.01; 1.08)	1.57 (1.39; 1.76)	1.12 (1.06; 1.19)
1999–2002	0.67 (0.63; 0.71)	0.80 (0.77; 0.83)	1.95 (1.74; 2.19)	1.15 (1.09; 1.22)
Language				
German	1.00	1.00	1.00	1.00
French	1.21 (1.03; 1.43)	1.12 (0.96; 1.23)	1.14 (0.95; 1.36)	1.04 (0.95; 1.20)
Italian	1.23 (0.87; 1.60)	1.12 (0.96; 1.43)	1.25 (0.92; 1.73)	1.10 (0.76; 1.51)
Romansh	0.86 (0.55; 1.20)	0.87 (0.72; 1.06)	1.15 (0.58; 2.11)	1.10 (0.79; 1.46)
Dwelling				
Rural	1.00	1.00	1.00	1.00
Urban	0.97 (0.92; 1.02)	1.04 (1.01; 1.07)	1.14 (1.04; 1.25)	1.11 (1.05; 1.16)
Spatial variation	0.30 (0.26; 0.36)	0.25 (0.22; 0.28)	0.32 (0.27; 0.40)	0.26 (0.23; 0.32)

Table 2.4: Spatio-temporal model estimates of age- and gender-specific tobacco-related cancer mortality in Switzerland from 1969–1972 to 1999–2002.

tobacco-related cancers. The lung cancer mortality in women was found to be highest in the French-speaking parts of the country, while the language region was not significantly associated with lung cancer mortality in men. The rates of male lung cancer, and all tobacco-related cancer mortality, were higher than those for women but with a decreasing temporal trend. A peak of tobacco-related mortality for males was observed in the 1980s followed by a constant decrease as also seen in other European countries (Levi et al., 2004). Results for lung cancer, and all other forms of tobacco-related cancer mortality, were similar but the temporal trend was particularly striking for lung cancer mortality. In general, around 85% of the lung cancer cases in Switzerland are attributed to the consumption of tobacco (Federal Office of Public Health, 2012a). The temporal trends observed in lung cancer mortality can be explained by gender-specific changes in smoking habits. These facts and the patterns observed reflect how ubiquitous the habit of smoking is.

Several recent reviews of lung cancer in people who never smoked have highlighted the importance of other risk factors such as genetic predisposition and various forms of exposure, e. g. environmental tobacco smoke, outdoor air pollution, ionizing radiation as well as occupational exposures to carcinogens (Couraud et al., 2012; Samet et al., 2009; Torok et al., 2011). The second most important cause of lung cancer in Switzerland (but far behind that of tobacco) is exposure to

radon especially in alpine regions and in the Jura. Radon is believed to be responsible for 200 lung cancer deaths per year (Federal Office of Public Health, 2012b). Due to a strong combined effect of smoking and radon, most of the radon-induced lung cancers occur among smokers, but radon is the primary cause of lung cancer among people who never smoked (WHO, 2009). Radon exposure is assumed to be stable in time, while the temporal and gender-specific trends observed change and are not related to radon exposure.

Schopper and Obrist (Schopper and Obrist, 2005) report that the highest incidence of male lung cancer is found in Western and Southern Switzerland, an outcome which agrees with our findings. Like our results indicate, these authors also found the highest female lung cancer incidence in the western, French-speaking part. Apart from this, Schüler and Bopp (1997) described a cluster of lung cancer in the central part of the German-speaking part of Switzerland, which is a rural but still a highly industrialized region. Socio-economic or environmental factors related to high industrialization might explain this cluster (Fano et al., 2004), which is confirmed by our analysis for older males (Fig. 2.3).

High female mortality in the urban areas is likely to be due to the change in female smoking behaviour of the middle of the twentieth century that took place earlier and was more pronounced in the cities (Curtin et al., 1997). Similar trends have been observed in other populations of the Western World (Graham, 1996). Exposure to environmental risks, such as air pollution and peer-to-peer tobacco smoke (especially in urban areas), may have accentuated the trend (Subramanian and Govindan, 2007). A similar observation has been reported in Germany, where lung cancer mortality rates in women doubled from 1960 to 1996 and increased by 155% in the 45–54 age group (Lienert et al., 2000).

Faeh et al. (2009) describe differences in cause-specific mortality between Swiss regions maintaining that patterns follow disparities among European populations suggesting an important, cultural factor. Others beg to differ, e. g. Tyczynski et al. (2003) who compared lung cancer mortality among European countries and report that female lung cancer mortality rates are highest in Northern Europe followed by the western part of the continent. Interestingly, while they found lung cancer mortality in women to be lower in France compared to Germany, we found the French-speaking parts of Switzerland having a higher mortality rates than the German-speaking regions, similar to the results of Faeh et al. (2009).

Lopez et al. (1994) describe a model of smoking epidemic defined by different stages. The most crucial component of this model is the lag of three to four decades between the rise in smoking prevalence and the increase in smoking-attributable mortality. According to them, female smoking highest prevalence typically lags males by 10–20 years. In males, lung cancer mortality has declined since the 1980s, so that if the predictions of Lopez et al. would take place in Switzerland, a reduction in female smoking-related mortality can be expected to start soon. Indeed, a recent decline of female lung cancer incidence and mortality from 2003–2007 has been reported from

the United States of America (USA) (Kohler et al., 2011). This trend is expected to continue for at least two more decades. Unfortunately, there is a dearth of data on smoking behaviour in Switzerland for the past few decades. One of the first national surveys was not until 1997. Since then a decrease in the number of smokers for both genders has been reported (Marques-Vidal et al., 2011).

What sets this study apart is the methodological approach using Bayesian CAR models on a small spatial unit (the municipality) allowing the identification of areas with high incidence of lung cancer and other tobacco-related cancer mortality. We updated the boundaries, in many places changed during the 40-year period covered, and defined the urbanization level and language on the basis of the most recent structure, i. e. that of October 2009. This may have wrongly classified those that had been restructured. However, less than 1% of the municipalities underwent re-classification and this low percentage should not affect the general results of the study. Furthermore, the Bayesian approach introducing spatial correlation at the municipality level should have reduced the impact of any potential re-classification.

Several approaches for disease mapping exist, e. g. count regression models accounting for spatial correlation by the CAR process are widely used. (Besag et al., 1991) proposed the inclusion of unstructured random effects into the model to account for heterogeneity in the data. Identifiability issues may arise when there is no prior information about the contribution of the spatial and non-spatial variation to the total one (Eberly and Carlin, 2000). Frequently Poisson regression models are used to map mortality rates but they are based on the assumption of the equality of mean and variance, which is not always valid. Negative binomial models account for unobserved variation by including an additional parameter. We did not consider an unstructured random effect, as we assume that additional variation is captured by the variance parameter of the negative binomial distribution.

Before the change of coding system from ICD-8 to ICD-10 in 1995, a priority rule was used by which cancer was rated high-priority and assigned as cause of death when multiple causes were documented on the death certificate. Especially in older people with several morbidities, this priority rule may have led to inflated cancer rates. However, because the rule was applied centrally and to all death certificates similarly, it cannot have affected the spatial patterns. Moreover it is unlikely that this rule could have had much influence on mortality of persons aged ≥ 60 years. A re-analysis taking into account the correction factor of 0.96 for lung cancer (Berrut and Junker, 2008) that was provided by the FSO did not change the results of the lung cancer-related mortality.

A limitation of the present study is the lack of more recent trends, which is due to the missing population data at the municipality level for the more recent years. As soon as data are available we will carry out further analyses to explore the trends for the years after 2002 and to address potential declining trends in female tobacco-related cancer mortality as seen in USA. We

are currently developing models to estimate inter-censual population data to make use of the full mortality dataset. However, in this study we aggregated the years around censuses, excluding the inter-censual years. Although proportional mortality rate estimates do not rely on population size (Sitas et al., 1989), we compared the results of this study using both standardized and proportional mortality rates. The conclusions by both approaches were similar suggesting that ignoring inter-censual years does not bias the results. The results also suggest a cohort effect for both genders, which is known for lung cancer mortality (Lee et al., 1990). Most of these findings can be explained by the gender-specific tobacco epidemics. Including information about the birth cohorts in this analysis might improve the estimation of the temporal trends.
Our results and newer trends would be useful as basis for projections and estimates of lung and all tobacco-related cancer mortality in the future. The insights of spatial differences in mortality as well as visualization of time and regional differences in tobacco-related cancer derived from this study will hopefully give new impetus to tobacco control activities and help evaluate the effectiveness of the national control programme.

Acknowledgements

This work was supported by Oncosuisse [grant number KLS-02393-02-2009]. We would like to thank the Swiss Federal Statistical Office for providing the data.

2.5 Appendix

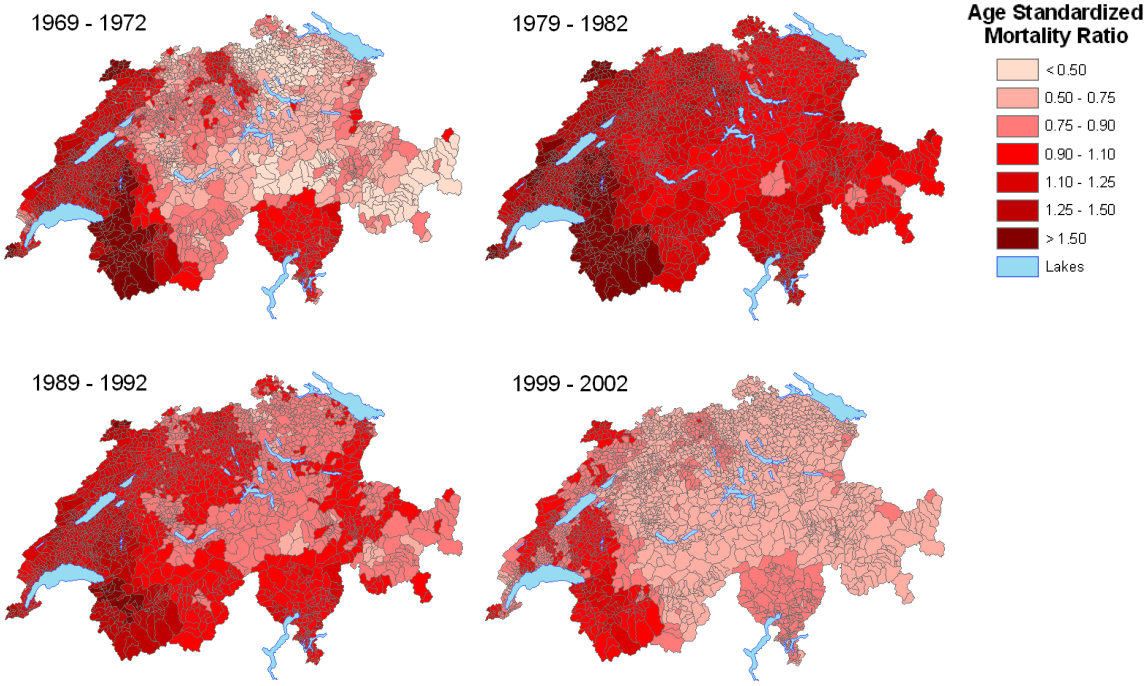


Figure 2.8: Geographical distribution of tobacco-related cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for males aged 35–59 years.

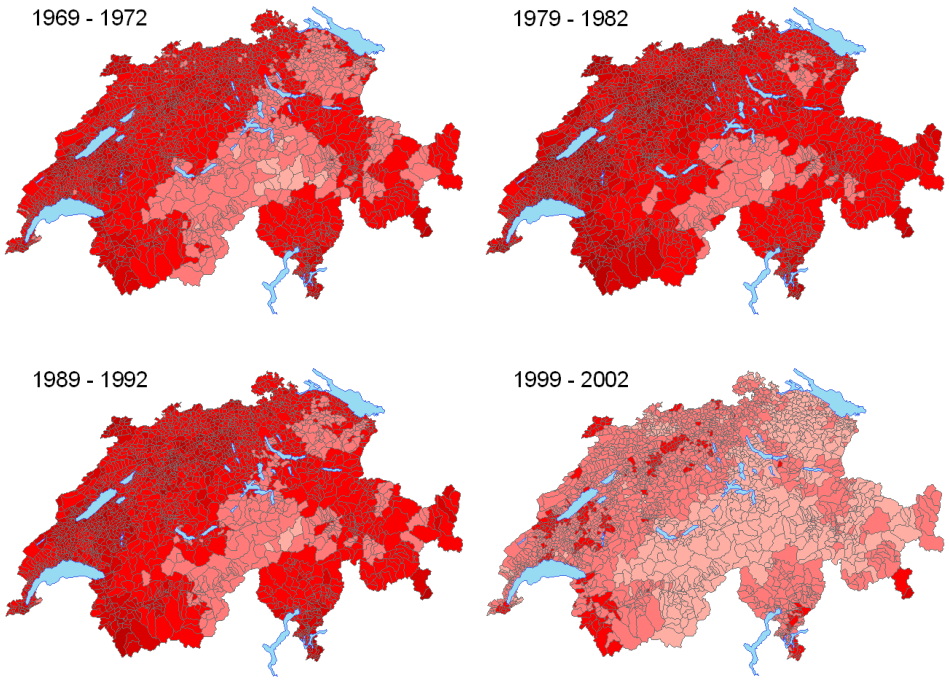


Figure 2.9: Geographical distribution of tobacco-related cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for males aged 60 years and above.

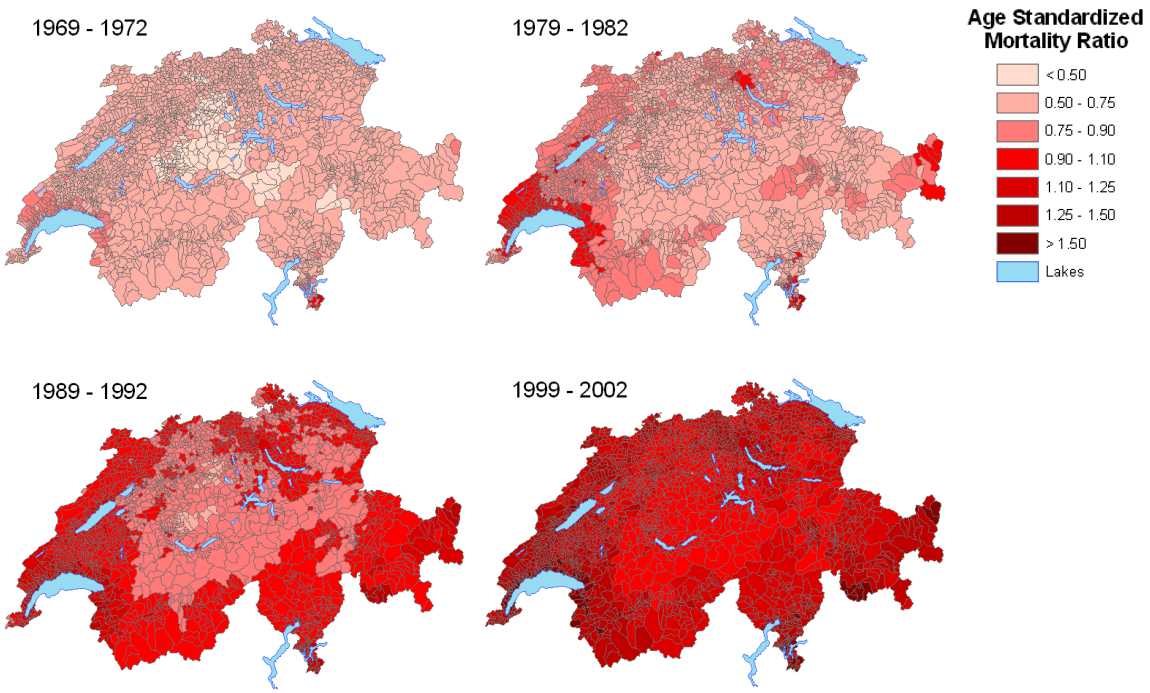


Figure 2.10: Geographical distribution of tobacco-related cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for females aged 35–59 years.

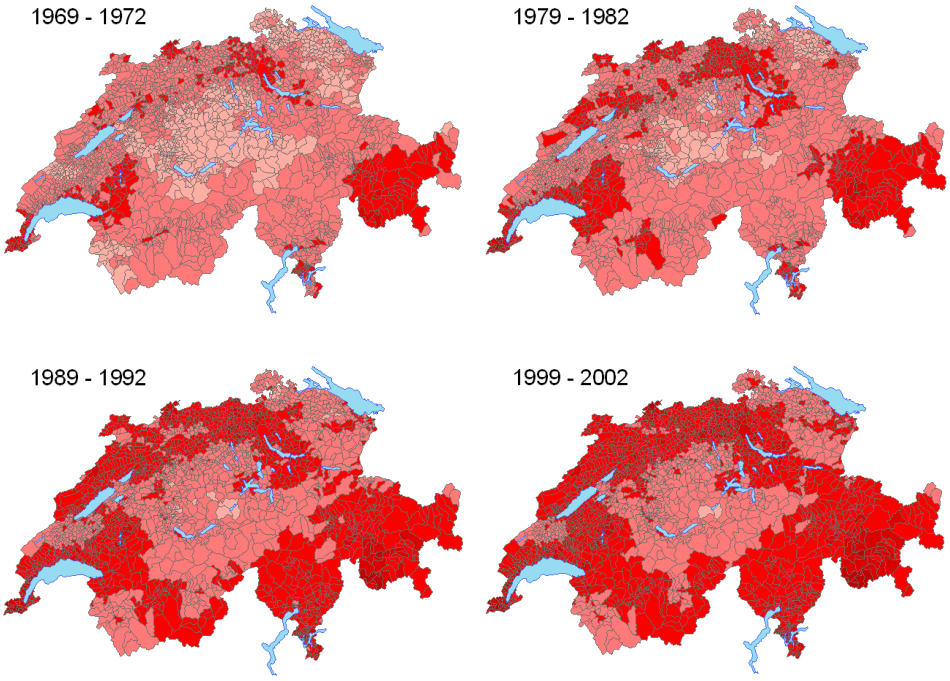
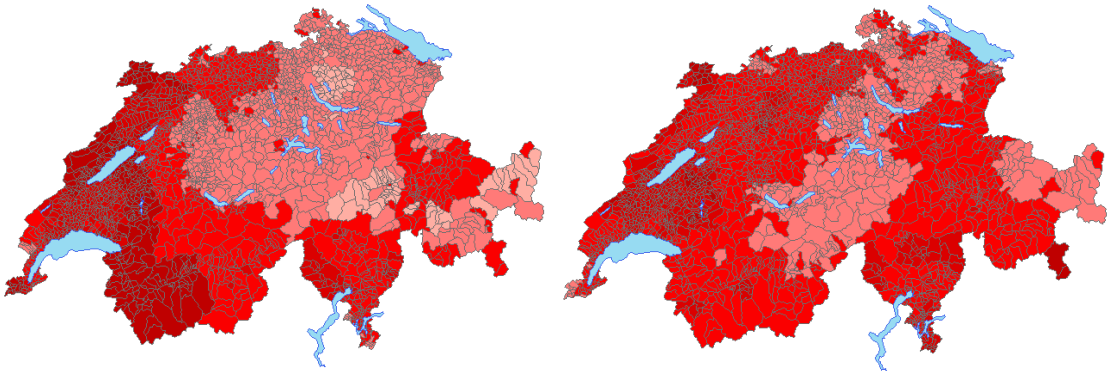


Figure 2.11: Geographical distribution of tobacco-related cancer mortality risks from 1969–1972 to 1999–2002 in Switzerland for females aged 60 years and above.

(a) Male, age 35–59 years

(b) Male, age ≥ 60 years



(c) Female, age 35–59 years

(d) Female, age ≥ 60 years

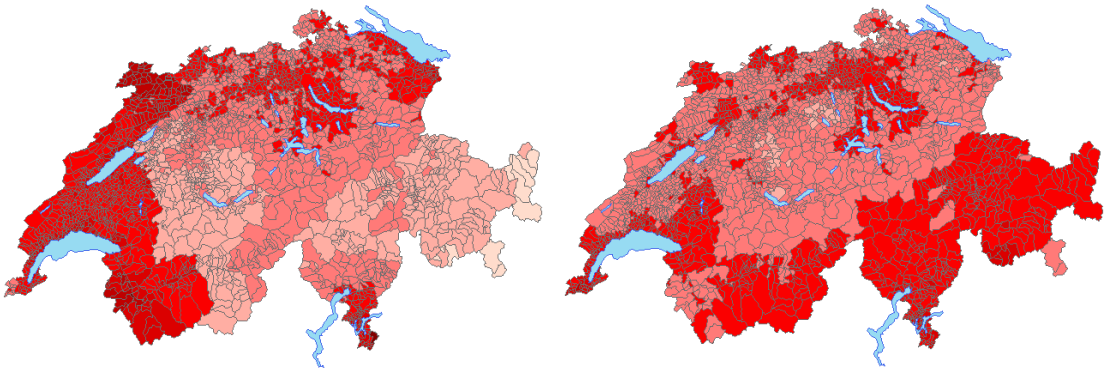
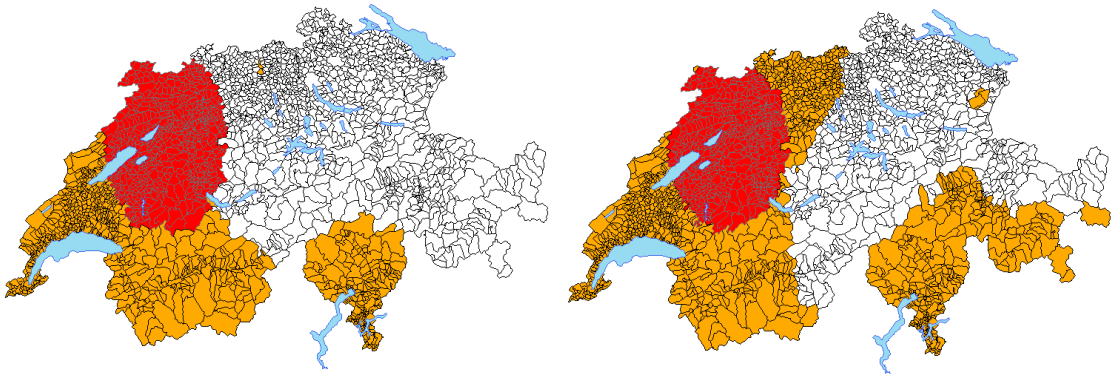


Figure 2.12: Geographical distribution of age- and gender-specific tobacco-related cancer mortality risks in Switzerland in the period 1999–2002.

(a) Male, age 35–59 years

(b) Male, age ≥ 60 years



(c) Female, age 35–59 years

(d) Female, age ≥ 60 years

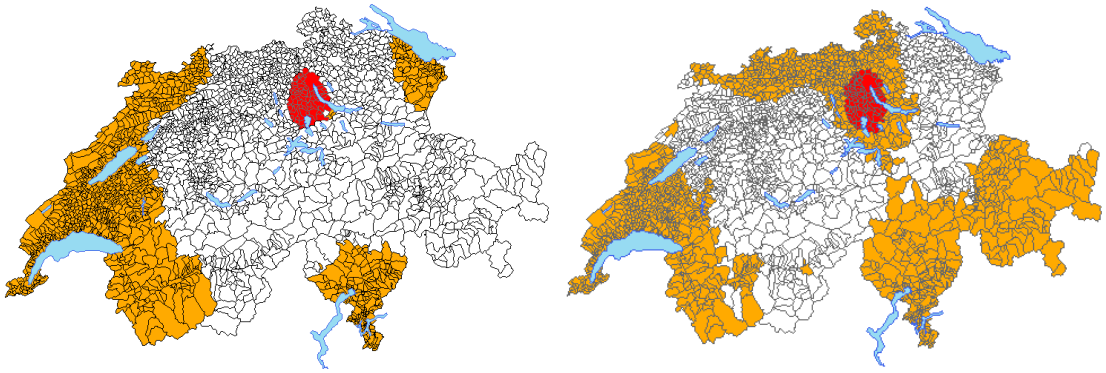


Figure 2.13: Clusters of age- and gender-specific tobacco-related cancer mortality risks in Switzerland in the period 1999–2002 (according to Kulldorff's statistic; red colour = highly significant clustering; yellow colour = significant at lower level).

Factor 1	Factor 2	1979–1982	1989–1992	1999–2002
Dwelling	**Language**		Male, age 35–59 years	
Urban	German	1.04 (0.80; 1.36)	0.98 (0.71; 1.32)	0.60 (0.45; 0.80)
Urban	French	1.26 (0.80; 1.93)	0.98 (0.61; 1.59)	0.78 (0.49; 1.23)
Urban	Italian	0.90 (0.48; 1.64)	0.86 (0.47; 1.62)	0.48 (0.42; 0.62)
Urban	Romansh	1.07 (0.19; 5.73)	/[1]	1.09 (0.19; 5.59)
Rural	German	1.02 (0.91; 1.15)	0.75 (0.65; 0.85)	0.49 (0.43; 0.56)
Rural	French	1.23 (0.92; 1.63)	0.98 (0.71; 1.32)	0.63 (0.46; 0.85)
Rural	Italian	0.88 (0.55; 1.38)	0.86 (0.55; 1.34)	0.46 (0.28; 0.75)
Rural	Romansh	1.04 (0.22; 4.84)	3.42 (1.01; 13.07)	0.89 (0.18; 3.86)
Dwelling	**Language**		Male, age 60 years and above	
Urban	German	1.13 (0.96; 1.34)	1.57 (1.32; 1.89)	0.79 (0.66; 0.93)
Urban	French	1.32 (1.00; 1.76)	1.19 (0.91; 1.58)	0.92 (0.70; 1.23)
Urban	Italian	1.45 (0.99; 2.12)	1.36 (0.93; 2.01)	0.99 (0.97; 1.05)
Urban	Romansh	1.06 (0.45; 2.57)	1.48 (0.67; 3.51)	1.24 (0.56; 2.99)
Rural	German	1.26 (1.17; 1.35)	1.35 (1.26; 1.46)	1.09 (1.01; 1.18)
Rural	French	1.47 (1.22; 1.77)	1.57 (1.32; 1.89)	1.28 (1.06; 1.55)
Rural	Italian	1.61 (1.20; 2.14)	1.80 (1.36; 2.41)	1.48 (1.10; 1.97)
Rural	Romansh	1.19 (0.55; 2.59)	1.96 (0.97; 4.21)	1.73 (0.86; 3.77)
Dwelling	**Language**		Female, age 35–59 years	
Urban	German	1.62 (0.68; 3.83)	2.81 (1.17; 6.14)	3.44 (1.52; 7.10)
Urban	French	2.09 (0.53; 8.03)	2.59 (0.70; 8.36)	3.86 (1.08; 11.91)
Urban	Italian	1.35 (0.22; 8.15)	3.47 (0.73; 16.03)	3.58 (3.71; 6.06)
Urban	Romansh	/	4.27 (1.47; 24.19)	/
Rural	German	1.40 (0.89; 1.99)	2.47 (1.67; 3.39)	3.55 (2.37; 4.89)
Rural	French	1.81 (0.69; 4.16)	2.81 (1.17; 6.14)	3.98 (1.69; 8.20)
Rural	Italian	1.17 (0.28; 4.22)	3.77 (1.20; 11.78)	2.68 (0.87; 8.68)
Rural	Romansh	/	/	/
Dwelling	**Language**		Female, age 60 years and above	
Urban	German	1.44 (0.86; 2.39)	1.71 (1.05; 2.70)	2.54 (1.66; 3.80)
Urban	French	1.66 (0.75; 3.64)	2.51 (1.19; 5.07)	3.08 (1.58; 5.96)
Urban	Italian	1.81 (0.62; 5.17)	2.47 (0.92; 6.32)	2.22 (1.69; 3.48)
Urban	Romansh	/	/	/
Rural	German	0.95 (0.73; 1.18)	1.25 (0.98; 1.53)	2.08 (1.68; 2.50)
Rural	French	1.09 (0.63; 1.80)	1.71 (1.05; 2.70)	2.52 (1.60; 3.92)
Rural	Italian	1.19 (0.52; 2.56)	1.67 (0.82; 3.37)	2.87 (1.47; 5.51)
Rural	Romansh	/	/	/

Table 2.5: Interpretation of the results of the spatial negative binomial regression analysis with interaction terms for age- and gender-specific lung cancer mortality in Switzerland.
[1]Due to very small counts, the estimates are not meaningful.

Factor 1	Factor 2	1979–1982	1989–1992	1999–2002
Dwelling	**Language**		Male, age 35–59 years	
Urban	German	0.98 (0.79; 1.24)	1.05 (0.83; 1.36)	0.63 (0.49; 0.80)
Urban	French	1.16 (0.81; 1.68)	0.98 (0.68; 1.42)	0.77 (0.52; 1.12)
Urban	Italian	1.01 (0.61; 1.63)	0.93 (0.58; 1.51)	0.63 (0.57; 0.73)
Urban	Romansh	1.01 (0.27; 3.62)	2.14 (0.77; 6.99)	0.85 (0.23; 3.27)
Rural	German	1.08 (0.97; 1.19)	0.87 (0.79; 0.98)	0.63 (0.56; 0.70)
Rural	French	1.27 (1.01; 1.62)	1.05 (0.83; 1.36)	0.78 (0.60; 0.98)
Rural	Italian	1.10 (0.75; 1.57)	1.00 (0.71; 1.45)	0.68 (0.47; 0.99)
Rural	Romansh	1.10 (0.34; 3.48)	2.31 (0.95; 6.69)	0.85 (0.26; 2.87)
Dwelling	**Language**		Male, age 60 years and above	
Urban	German	1.03 (0.91; 1.17)	1.22 (1.07; 1.40)	0.71 (0.63; 0.80)
Urban	French	1.11 (0.90; 1.37)	1.02 (0.83; 1.25)	0.79 (0.64; 0.97)
Urban	Italian	1.08 (0.82; 1.43)	1.04 (0.77; 1.35)	0.83 (0.80; 0.87)
Urban	Romansh	1.37 (0.74; 2.51)	1.40 (0.76; 2.57)	1.22 (0.68; 2.21)
Rural	German	1.11 (1.05; 1.17)	1.11 (1.06; 1.18)	0.85 (0.80; 0.90)
Rural	French	1.20 (1.05; 1.37)	1.22 (1.07; 1.40)	0.94 (0.82; 1.08)
Rural	Italian	1.16 (0.95; 1.44)	1.21 (0.99; 1.51)	0.98 (0.80; 1.20)
Rural	Romansh	1.48 (0.86; 2.52)	1.67 (0.98; 2.87)	1.46 (0.87; 2.46)
Dwelling	**Language**		Female, age 35–59 years	
Urban	German	1.19 (0.68; 2.08)	1.82 (1.05; 3.18)	1.84 (1.12; 3.02)
Urban	French	1.65 (0.69; 3.93)	1.84 (0.79; 4.32)	2.37 (1.07; 5.11)
Urban	Italian	1.15 (0.38; 3.67)	2.06 (0.69; 6.25)	1.88 (1.71; 2.74)
Urban	Romansh	/	/	/
Rural	German	1.15 (0.89; 1.51)	1.45 (1.13; 1.87)	1.88 (1.50; 2.36)
Rural	French	1.59 (0.90; 2.85)	1.82 (1.05; 3.18)	2.42 (1.44; 3.99)
Rural	Italian	1.12 (0.50; 2.66)	2.05 (0.92; 4.60)	1.54 (0.97; 4.35)
Rural	Romansh	/	/	/
Dwelling	**Language**		Female, age 60 years and above	
Urban	German	1.09 (0.87; 1.35)	1.15 (0.90; 1.47)	1.16 (0.95; 1.41)
Urban	French	1.10 (0.76; 1.58)	1.32 (0.92; 1.90)	1.25 (0.89; 1.75)
Urban	Italian	1.03 (0.64; 1.68)	1.35 (0.86; 2.09)	1.08 (0.98; 1.26)
Urban	Romansh	1.67 (0.62; 4.66)	0.96 (0.33; 2.77)	1.38 (0.55; 3.47)
Rural	German	1.03 (0.94; 1.15)	0.99 (0.89; 1.09)	1.07 (0.97; 1.17)
Rural	French	1.04 (0.82; 1.35)	1.15 (0.90; 1.47)	1.16 (0.91; 1.45)
Rural	Italian	0.98 (0.69; 1.43)	1.18 (0.84; 1.62)	1.26 (0.91; 1.76)
Rural	Romansh	1.58 (0.67; 3.97)	0.84 (0.33; 2.13)	1.27 (0.57; 2.87)

Table 2.6: Interpretation of the results of the spatial negative binomial regression analysis with interaction terms for age- and gender-specific tobacco-related cancer mortality in Switzerland.

Chapter 3

Tobacco-related cancer mortality – projections for different geographical regions in Switzerland

Verena Jürgens[1,2], Silvia Ess[3], Harish C. Phuleria[1,2], Martin Früh[4], Matthias Schwenkglenks[5], Harald Frick[6], Thomas Cerny[4], Penelope Vounatsou[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Cancer Registry of St. Gallen & Appenzell, St. Gallen, Switzerland
[4] Medical Oncology-Hematology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland
[5] Institute of Pharmaceutical Medicine (ECPM), University of Basel, Basel, Switzerland
[6] Cancer Registry Grisons & Glarus, Kantonsspital Graubünden, Chur, Switzerland

This paper has been published in *Swiss Medical Weekly* 2013, 143:w13771.

Abstract

Switzerland is divided into 26 cantons of variable population size and cultural characteristics. Although a federal law to protect against passive smoking and a national tobacco control program exist, details of tobacco-related policies are canton-specific. This study aimed to project gender-specific tobacco-related cancer mortality in Switzerland at different geographical levels for the periods 2009–2013 and 2014–2018.

In this analysis data on Swiss tobacco-related cancer mortality from 1984 until 2008 have been used. Bayesian age-period-cohort models were formulated to assess past trends of gender-specific tobacco-related cancer mortality and to project them up to 2018 at cantonal and language region level. Furthermore, estimates are provided at national scale by age categories 50–69 and ≥ 70.

Model-based estimates at cantonal level identified regions with low and high tobacco-related cancer mortality rates for observed and projected periods. Our analysis based on language regions showed the lowest mortality in the German-speaking part. Projections at national level, between younger (age 50–69) and older (age ≥ 70) males, indicated an ongoing decreasing trend for males but an upward trend for females. The gap of tobacco-related cancer mortality rates between younger and older males seems to be shrinking. In females, a stronger rise was obtained for the younger age group.

Our findings indicate region-, sex- and age-related differences in tobacco-related cancer mortality in Switzerland and this could be useful for health care planning and for evaluating the impact of canton-specific tobacco-related policies and interventions.

Keywords: Bayesian inference; age-period-cohort models; tobacco-related cancer mortality; projections

3.1 Introduction

Tobacco-smoking is a major cause of premature mortality. In 1990, around three quarters of a million deaths in the age group 35–69 were attributable to tobacco smoking in Europe (Boyle, 1997). In fact, it accounts for more deaths than traffic-accidents, AIDS, alcohol, cocaine, homicide, suicide and fires together (Chaturvedi, 2004). Smoking cessation reduces the hazard substantially in younger age (Doll et al., 2004). Smoking behaviour changed over the last decades. In Switzerland, smoking prevalence decreased during 1992 and 2007 in both genders (Marques-Vidal et al., 2011). Several factors may have contributed to this trend including prevention programs and greater sensitivity to mortality from tobacco-related illness in males (Flandorfer et al., 2010). Smoking can be supposedly used to balance stress and especially younger women use smoking to control or even suppress appetite. On average, the weight of smokers is 3–4 kg lower than for non-smokers (Heishman, 1999).

Reports of the International Agency for Research on Cancer showed that in 2008 21.9% of all male cancer deaths in Switzerland were related to the lung, whereas the proportion was 12.7% for Swiss women. Smoking is the major risk factor for the development of lung cancer, but it also causes cancers at other sites. Smoking may have an effect on cancers of the larynx and the pharynx (Lee et al., 2008). Parental smoking may induce childhood cancer, however studies showing evidence of definite association are still missing. In addition, there is strong evidence of a causal association between passive smoking and lung cancer (Taylor et al., 2007). Doll et al. (1976) considered the following cancer sites to be related to tobacco: lung (C33–C34), oesophagus (C15), rectum (C20), pancreas (C25), bladder (C67) and cancer of other respiratory sites (C00–C14) including cancer of the lip, tongue, mouth, larynx, trachea and pharynx (excluding nasopharynx).

Switzerland is divided into 26 cantons, which vary in population size and cultural characteristics. Regulations, e. g. in health systems, are canton-specific, covering domains which are not controlled by the federal constitution. In 2010 a federal law to protect against passive-smoking was enacted, aiming at smoke-free interiors in enclosed spaces that are publicly accessible, i. e. restaurants, bars, schools, hospitals etc., and workplaces with two or more people. However, smoking-rooms still exist under certain conditions in many cantons as federal law just provides the minimum provisions for protection against passive smoking and stricter smoking-related laws are canton-dependent. Figure 3.1 shows the Swiss cantons according to their smoking ban regulations in restaurants in May 2012. All cantons applied the national law which can be seen as a minimal protection against passive-smoking, whereas some defined further regulations. These refinements either allow service only in those rooms indicated for smoking (smoke-rooms) or prohibit any service in smoke-rooms.

Information about the trend of tobacco-related cancer mortality over time as well as the spatial distribution of the disease are essential for health care planning related to smoking behaviour. They can identify areas with high tobacco-specific cancer mortality. Projections of mortality

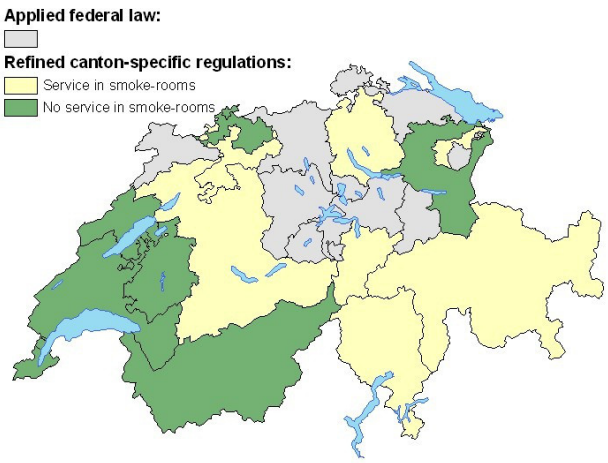


Figure 3.1: Canton-specific smoking ban regulations indicating application of the federal law (*Bundesgesetz zum Schutz vor Passivrauchen*) (in grey), further regulations covering service in smoke-rooms (in yellow) and no service in smoke-rooms (in green).

rates are helpful tools for assessing the future burden of a disease and as a basis for evaluating the impact of interventions carried out. Model-based estimates can then be compared to observed data for the projection periods.

Age-period-cohort (APC) models are commonly used to project cancer mortality and incidence. Within this approach the dataset is stratified into three components separating the age, the period when the event occurred and the birth cohort of the person. Typically, projections based on APC-models are done at country level (Clèries et al., 2006; Eilstein et al., 2008), however estimates at subunits may be useful (Knorr-Held and Rainer, 2001) especially when health care planning is decentralized as in the case of Switzerland.

This study aims to project gender-specific tobacco-related cancer mortality Switzerland at different geographical levels for the periods 2009–2013 and 2014–2018. Projected estimates at cantonal level may identify high- or low risk areas and help evaluating canton-specific smoking-related policies in the future. In addition, mortality data are aggregated over the three language regions (i. e. German, French and Italian) to assess trends among the regions which can be compared with trends in same language neighbouring countries – Germany, France and Italy. Furthermore, countrywide projections are estimated by specific age groups to determine the burden of tobacco-related cancer in the future on a larger scale and possibly identify age-specific differences.

3.2 Methods

3.2.1 Data sources and management

Mortality data at individual level were provided by the Federal Statistical Office (FSO). Cancer death counts related to tobacco, following the definition by Doll et al. (1976) were used for the analysis. Tobacco-related cancer rarely occurs in persons younger than 50. Therefore, we excluded death counts of those aged below. In 1995 priority given to certain causes of deaths was removed from rules related to coding death certificate. This change affected cause-specific reported deaths, however FSO provides disease-related correction factors to adjust for the years before 1995 (Berrut and Junker, 2008). Estimated rates were (age-) standardized directly, using Segi World population (Segi, 1960).

Subgroup-specific population data at cantonal level were obtained from the webpage of the FSO, covering the periods 1984–1988, ..., 2004–2008, as well as 2009–2013 and 2014–2018. The data for the latter time periods were estimated by the FSO considering a mid-way scenario, based on past population trends and assuming no extreme changes (e. g. migration). Age and period were aggregated to five-year groups.

Cantons with small number of cases, which did not allow for reliable estimates, were merged with their neighbours. These included canton Uri, Obwalden and Nidwalden, as well as Appenzell Ausserrhoden and Appenzell Innerrhoden. A map of Switzerland (Fig. 3.8) as well as cantonal abbreviations are provided in the Appendix. ArcGIS, a geographic information system, has been used to create maps illustrating the results. The required shapefile describing the spatial configuration of the cantons was obtained from the FSO and modified.

3.2.2 Statistical analysis

Bayesian Poisson and Negative Binomial APC regression models were formulated and fitted in WinBUGS (Imperial College and MRC, London, UK) using Markov chain Monte Carlo (MCMC) simulations. The mean of the age- and period-specific death counts μ_{ijk} was regressed on the effects of age, period and cohort. The general model applied for projecting at cantonal level reads as follows:

$$log(\mu_{ijk}) = log(n_{ijk}) + \alpha_{ik} + \beta_{jk} + \gamma_{j-i,k} + \phi_k \tag{3.1}$$

where n_{ijk} is the population for the age group (i), period (j) and region (k), α, β and γ account for the effect of age, period and cohort, respectively. Canton-specific random effects ϕ_k modeled via conditional autoregressive (CAR) specifications (Bernardinelli and Montomoli, 1992) were included to allow for spatial correlation.

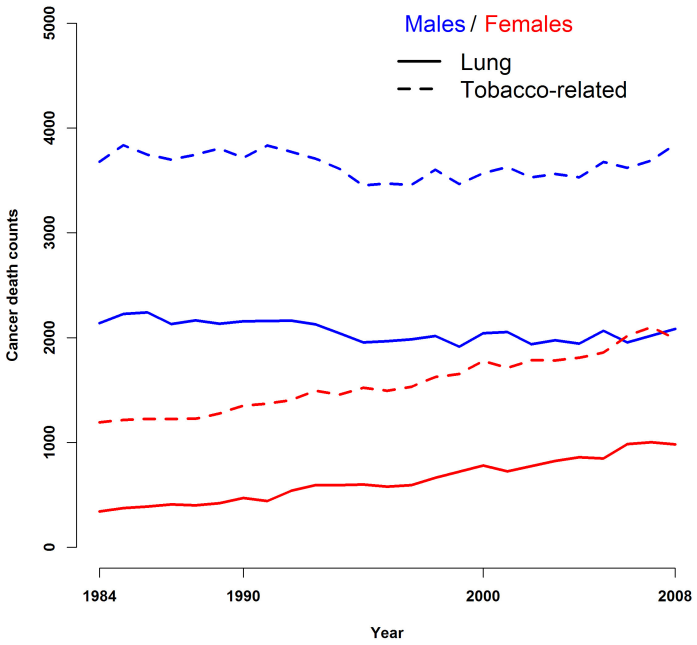


Figure 3.2: Overall lung and all tobacco-related cancer deaths observed during 1984–2008 at country level for males and females.

Bayesian APC-models were applied to project age- and gender-specific tobacco-related cancer mortality for the periods 2009–2013 and 2014–2018 (i) at national level (ii) for the three language regions and (iii) at cantonal level. The models for countrywide and language-specific analyses did not consider spatial random effects. Results are presented for each gender and sub-region. In addition, countrywide estimates are given for the age categories 50–69 and ≥ 70. The predictive performance of the models was assessed by comparing model-based and observed data using the sum of squared residuals. In particular, models were fitted during 1984–1998 and predictions were obtained for the remaining period of 1999–2008. Implementation details and model evaluation can be found in the Appendix.

3.3 Results

Preliminary analysis suggested similar trends between lung and tobacco-related cancer mortality in Switzerland during 1984 to 2008 (see Fig. 3.2). Death counts for males were stable while for females increased steadily.

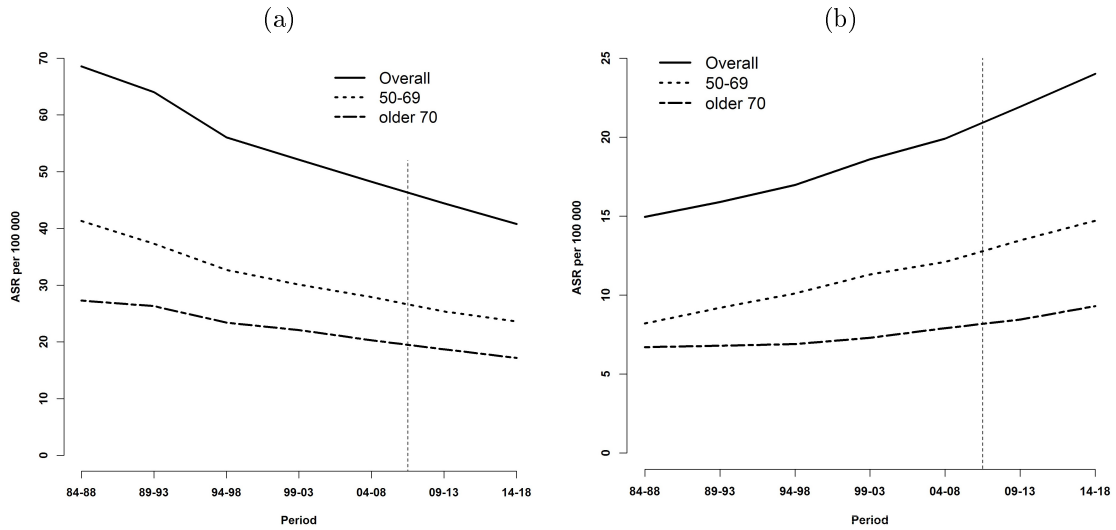


Figure 3.3: Age-specific and overall age-standardized tobacco-related cancer mortality rates (per 100 000), observed (during 1984–2008) and predicted (posterior median during 2009–2013 and 2014–2018) at country level for males (a) and females (b). The vertical line indicates the start of the prediction period.

3.3.1 Age- and gender-specific projections at national level

Figure 3.3 shows observed and predicted tobacco-related cancer mortality rates during 1984–2008 and 2009–2018, respectively, for each gender and age category. For both, males and females, higher rates were observed for the younger age group.

For males, a decreasing trend was obtained which is less pronounced from 1994–1998 onwards. On the other hand, projected absolute death counts increased during 2009–2013 and 2014–2018 (see Table 3.1), reflecting the demographic aging of the population (see Fig. 3.4).

The difference in mortality rates between younger and older males narrowed during the study periods. Overall, the mortality rate in males decreased by approximately 30%.
Tobacco-related cancer mortality for Swiss women steadily increased during the study period. In contrast to males, model-based estimates indicated that the difference of mortality rates between younger and older females increased. The projected rates in 2014–2018 indicate that the female tobacco-related cancer mortality rates rise by more than 60% during the whole study period.

3.3.2 Gender-specific projections for the language regions

Figure 3.5 shows the trends for gender-specific tobacco-related cancer mortality rates by language regions. For males, the lowest rates were observed for the German-speaking region. No significant differences were observed for the rates in the French- and Italian-speaking part for either

	2004–2008	2009–2013	2014–2018
Males			
Overall	17 656	18 340 (17 300, 19 820)	19 270 (17 260, 22 100)
50–69	7 703	8 030 (7 522, 8 690)	8 023 (7 085, 9 333)
$\geq$ 70	9 953	10 310 (9 713, 11 130)	11 260 (10 080, 12 910)
Females			
Overall	9 349	11 080 (10 510, 11 620)	13 270 (11 960, 14 570)
50–69	3 453	4 255 (4 005, 4 534)	5 010 (4 416, 5 643)
$\geq$ 70	5 896	6 824 (6 456, 7 148)	8 261 (7 453, 9 026)

Table 3.1: Raw age- and gender-specific tobacco-related cancer death counts observed (during 2004–2008) and projected (posterior median during 2009–2013 and 2014–2018) at country level. The numbers in the brackets correspond to 95% Bayesian credible intervals.

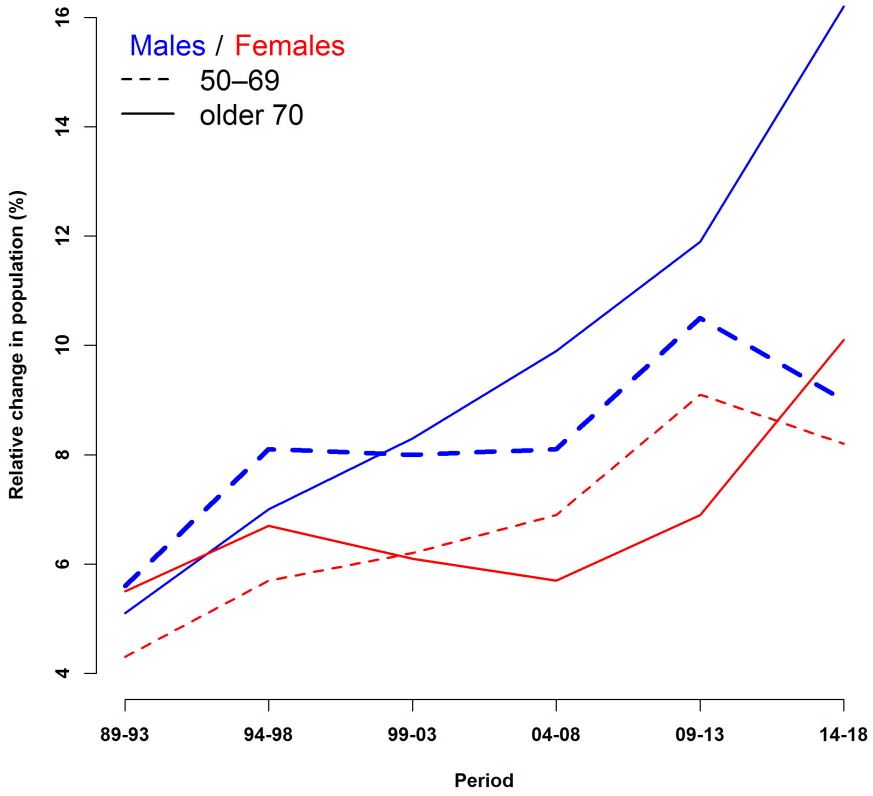


Figure 3.4: Relative change in population (in % based on preceding year) for males (red) and females (blue) and age groups 50–69 (dashed) and ≥ 70.

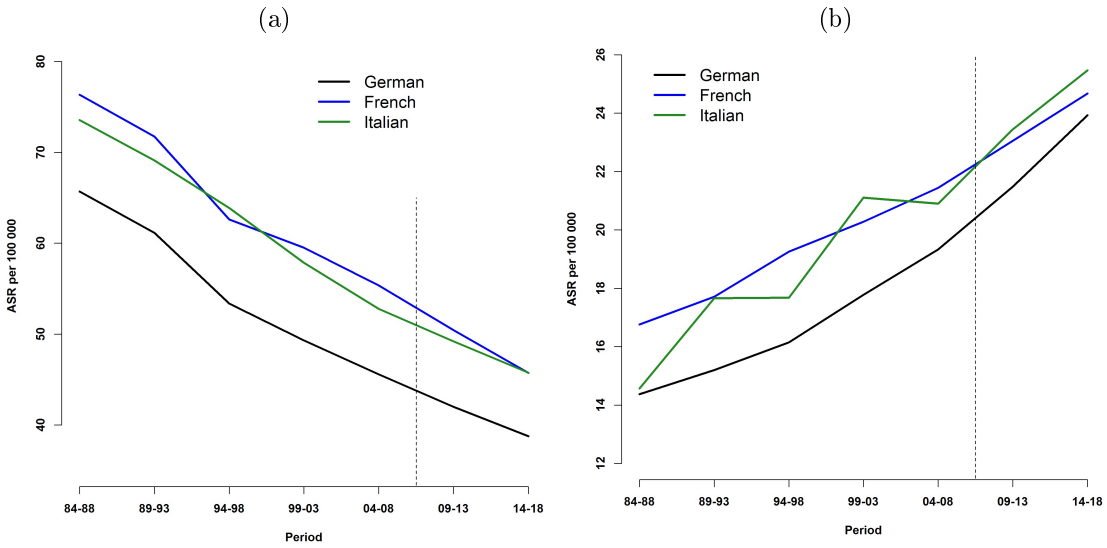


Figure 3.5: Language regions-specific age-standardized tobacco-related cancer mortality rates (per 100 000), observed (during 1984–2008) and predicted (posterior median during 2009–2013 and 2014–2018) for males (a) and females (b). The vertical line indicates the start of the prediction period.

sex. Considering cause-specific mortality rates for females, the gap between the lowest (German-speaking) and the highest (French-speaking) region-specific rates clearly decreased during the 30 study years. While the German-speaking region did show significantly lower rates up to 2008, estimates for the projected periods indicate no significant difference between the two regions. In 1984–1988 the rates for the Italian-speaking area were not significantly different from those in the German-speaking one. Subsequently, however, a significantly higher tobacco-related cancer mortality rate for females was observed in the Italian-speaking part. Again, this can not be said for the projected periods.

3.3.3 Gender-specific projections at cantonal level

Figures 3.6 and 3.7 show the observed gender-specific (male and female, respectively) tobacco-related cancer mortality (age ≥ 50) observed for period 5 (2004–2008) and the model-based projections for period 6 (2009–2013) and 7 (2014–2018) at cantonal level. Cantonal boundaries correspond to the updated structure including merged cantons. For males, an overall decreasing trend was observed, whereas rates in the French-speaking cantons still remained the highest, with the highest rates in cantons Valais and Neuchâtel.

Lowest tobacco-related cancer mortality rates were observed in the region covering the two cantons Appenzell Ausserrhoden and Appenzell Innerrhoden in Eastern Switzerland. Figure 3.9 shows less precise estimates for these regions. However, maps of the observed rates (Fig. 3.6(a), Fig. 3.7(a))

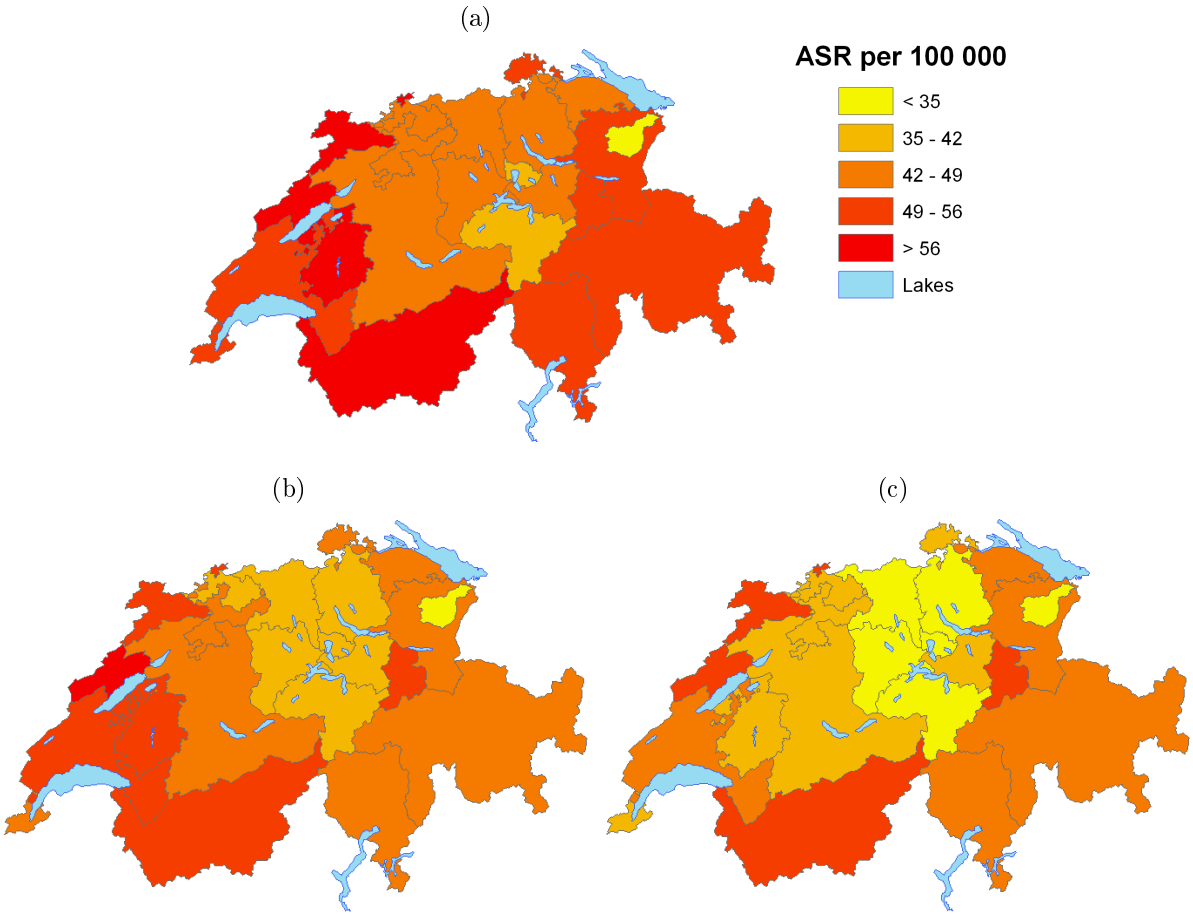


Figure 3.6: Age-standardized tobacco-related cancer mortality rates (per 100 000) for males age ≥ 50, (a) observed during 2004–2008 and predicted posterior median (b) during 2009–2013 and (c) 2014–2018.

confirm low rates.

For males, the highest decrease was observed for the two cantons Zurich and Luzern – a steady reduction of around 12–16% per period. Regarding females, tobacco-related cancer morality was projected to remain stable in the cantons Geneva and Graubünden. Furthermore, an increase of approximately 50% from 2004–2008 to 2014–2018 was estimated for the cantons Schaffhausen, Zug and Schwyz.

3.4 Discussion

This is the first study in Switzerland to project age- and gender-specific tobacco-related cancer mortality at different geographical levels during 2009–2013 and 2014–2018. Gender-specific estimates of age-standardized rates are presented for each canton and language region. In addition, projections are also made at national scale by the two age categories 50–69 and ≥ 70.

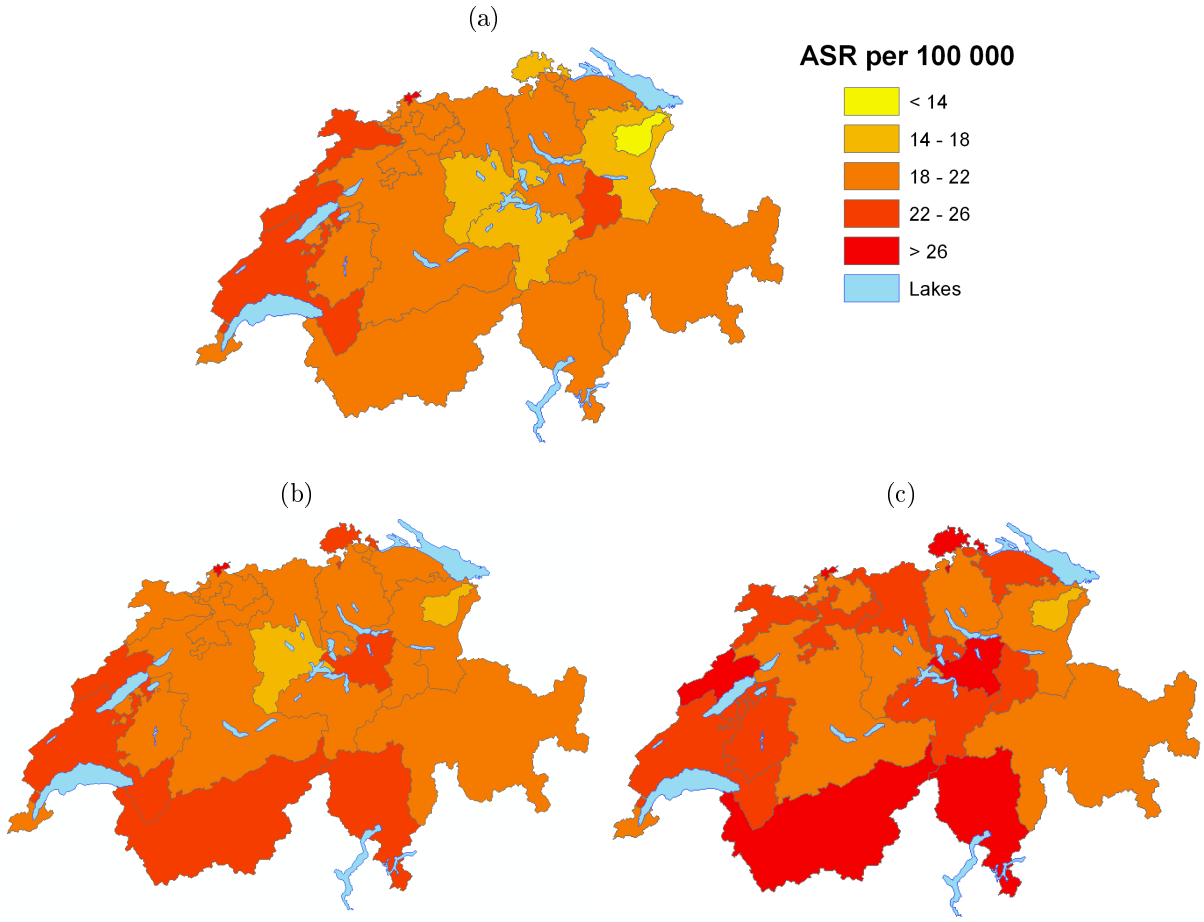


Figure 3.7: Age-standardized tobacco-related cancer mortality rates (per 100 000) for females age ≥ 50, (a) observed during 2004–2008 and predicted posterior median (b) during 2009–2013 and (c) 2014–2018.

Random effects at cantonal level have been considered to account for spatial correlation in mortality which arises from similar exposures (e. g. environmental factors , cultural characteristics, health policies) at neighbouring areas. The models have been formulated within the Bayesian framework of inference adopting CAR priors for the random effects. The Bayesian approach enables computation via the use of MCMC and smooths the rates so that estimates from areas with small populations do not dominate (Clayton and Kaldor, 1987).

Overall, tobacco-related cancer mortality rates in males are decreasing, whereas female mortality is increasing. Estimates at the national level show that mortality rates for younger (age 50–69) females increased more than those for older (age ≥ 70) ones. On the other hand, a steeper downward trend was obtained for young males compared to older ones. The less pronounced trend for males observed from 1994 onwards might be due to the change in the death certificate coding

introduced in 1995. Early Swiss cancer mortality projections at national scale estimated a plateau of male lung cancer mortality and a rising trend for females for the end of the last century (Negri et al., 1990).

Maps of the projected rates can greatly help in identifying cantons with high or low mortality rates. However, the accuracy of estimates needs to be considered carefully. In this study some cantons were merged to allow for more precise estimates. Figure 3.9 shows, that estimates for these regions have the lowest accuracy however, they seem reasonable compared to the trends of preceding periods and neighbouring cantons.

Important differences in cancer-related mortality were estimated among the language regions for the observed, but not the projected periods. The latter may be explained by the wide credible intervals of the estimates. Results for language regions could be compared with projections in Germany, France and Italy to assess the hypothesis that Swiss language regions reflect the situation in the respective neighbouring countries.

Comparing Figure 3.1 with Figures 3.6 and 3.7 it is interesting to note that the high rates occurred in the French-speaking part, which applied the strictest smoking ban regulations. Regarding males' tobacco-related cancer mortality rates in 2014–2018 (Fig. 3.6(c)), cantons with the lowest rates in the middle of the country resemble those only applying minimal smoking-ban regulations, i. e. the federal law. However, this finding may have arisen by,chance because it does not take into account the time lag of 20–30 years between smoking (exposure) and the disease/death. In this study, cancer deaths below the age of 50 were excluded, as lung cancer in this age group is rarely due to tobacco smoking.

A decreasing prevalence of smoking in Switzerland has been reported for 1992–2007 (Marques-Vidal et al., 2011). The authors reported a stronger trend, corresponding with stage 4 of the smoking epidemic model proposed by Lopez et al. (1994). This model describes different stages indicating changes in prevalence of smoking, tobacco consumption as well as mortality due to smoking. Stage 4 is characterized by an overall declining trend in smoking prevalence and a peak in male smoking-related mortality. As a consequence of rapid changes in female smoking behaviour during the mid-twentieth century, mortality for women has increased. However, a peak of female tobacco-related cancer mortality has not yet been observed. These findings confirm European trends in gender-specific lung cancer mortality. Based on observations in 2007, rates in Europe were estimated to decrease for males and increase for females in 2012, which is consistent with our results (Malvezzi et al., 2012). However studies in the United States showed a declining trend in female lung cancer mortality rates (Kohler et al., 2011).

It should be noted that we only account for cancer mortality and did not capture overall mortality due to smoking. Furthermore, we did not account for smoking behaviour and related changes over time. Model predictions could be optimized using additional information. For example, Knorr-Held and Rainer (2001) introduced in their analysis information on tobacco consumption estimated

from numbers of cigarette packages sold to project lung cancer mortality in Western Germany based on the effects of age and cohort. In Switzerland, obtaining temporal as well as spatial information on smoking patterns remains a challenge, especially for earlier years.
The current analysis cannot take into account the impact of recent interventions, such as the Swiss federal law which was introduced in 2010, because there are no mortality data to estimate the intervention effect. This may be considered as a limitation; however our analysis will allow assessing the impact of smoking ban by comparing our projections with the observed rates in the future.
Demographic changes have an impact on incidence and mortality, especially for diseases mainly occurring in elderly people, as the fraction of older people in the population is rising. Among others, improved health care is one reason for the increasing longevity.
Stable rates may be estimated despite that the absolute number of deaths is increasing because the older population is increasing. In this study, a downward trend was observed for male cause-specific cancer mortality, although raw death counts increased for these periods.
Maps of observed and projected tobacco-related cancer mortality highlight cantons with low or high rates. Estimates for language regions show lowest mortality in the German-speaking part of country but appear to become less pronounced over time especially for women. Countrywide projections indicate an ongoing declining trend for males, while for females a continuing rise is predicted, which is more pronounced in the younger age group.

Acknowledgements

This work was supported by Oncosuisse [grant number KLS-02393-02-2009]. We would like to thank the Swiss Federal Statistical Office for providing the data.

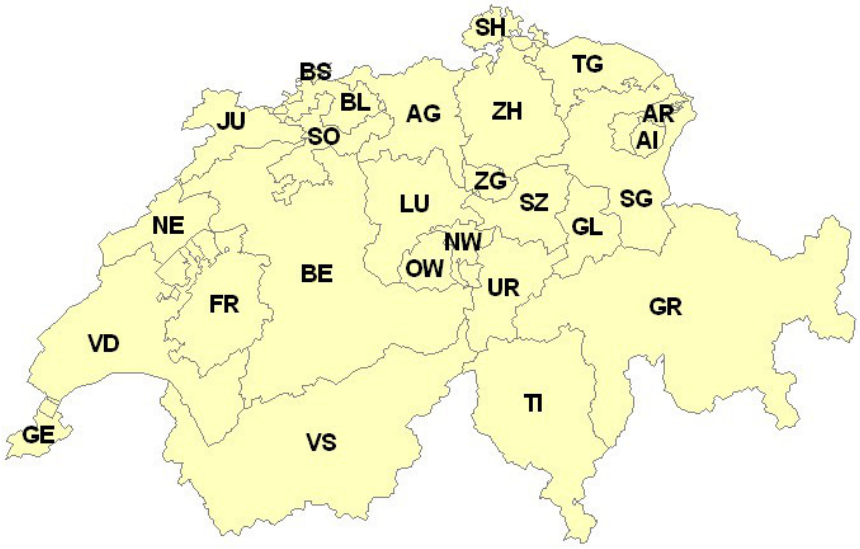


Figure 3.8: Swiss cantons labeled by their abbreviations.

3.5 Appendix

3.5.1 Details on cantons

Zurich (ZH), Bern (BE), Lucerne (LU), Uri (UR), Schwyz (SZ), Obwalden (OW), Nidwalden (NW), Glarus (GL), Zug (ZG), Fribourg (FR), Solothurn (SO), Basel-Stadt (BS), Basel-Landschaft (BL), Schaffhausen (SH), Appenzell Ausserrhoden (AR), Appenzell Innerrhoden (AI), St. Gallen (SG), Graubünden (GR), Aargau (AG), Thurgau (TG), Ticino (TI). Vaud (VD), Valais (VS), Neuchâtel (NE), Geneva (GE), Jura (JU).

3.5.2 Details on results

Accuracy of estimates

	2004–2008	2009–2013	2014–2018
Males			
German	45.6	42.0 (39.9, 45.3)	38.8 (35.1, 45.3)
French	55.4	50.4 (46.7, 55.0)	45.7 (38.9, 54.4)
Italian	52.8	49.2 (44.7, 53.8)	45.8 (38.8, 53.8)
Females			
German	19.3	21.5 (20.3, 22.7)	23.9 (21.3, 26.8)
French	21.4	23.1 (21.1, 24.9)	24.7 (21.3, 28.1)
Italian	20.9	23.5 (20.6, 26.6)	25.5 (20.7, 30.6)

Table 3.2: Age-standardized tobacco-related cancer mortality rates (per 100 000), observed (during 2004–2008) and predicted (posterior median during 2009–2013 and 2014–2018) at language region level for males and females. The numbers in the brackets correspond to 95% credible intervals.

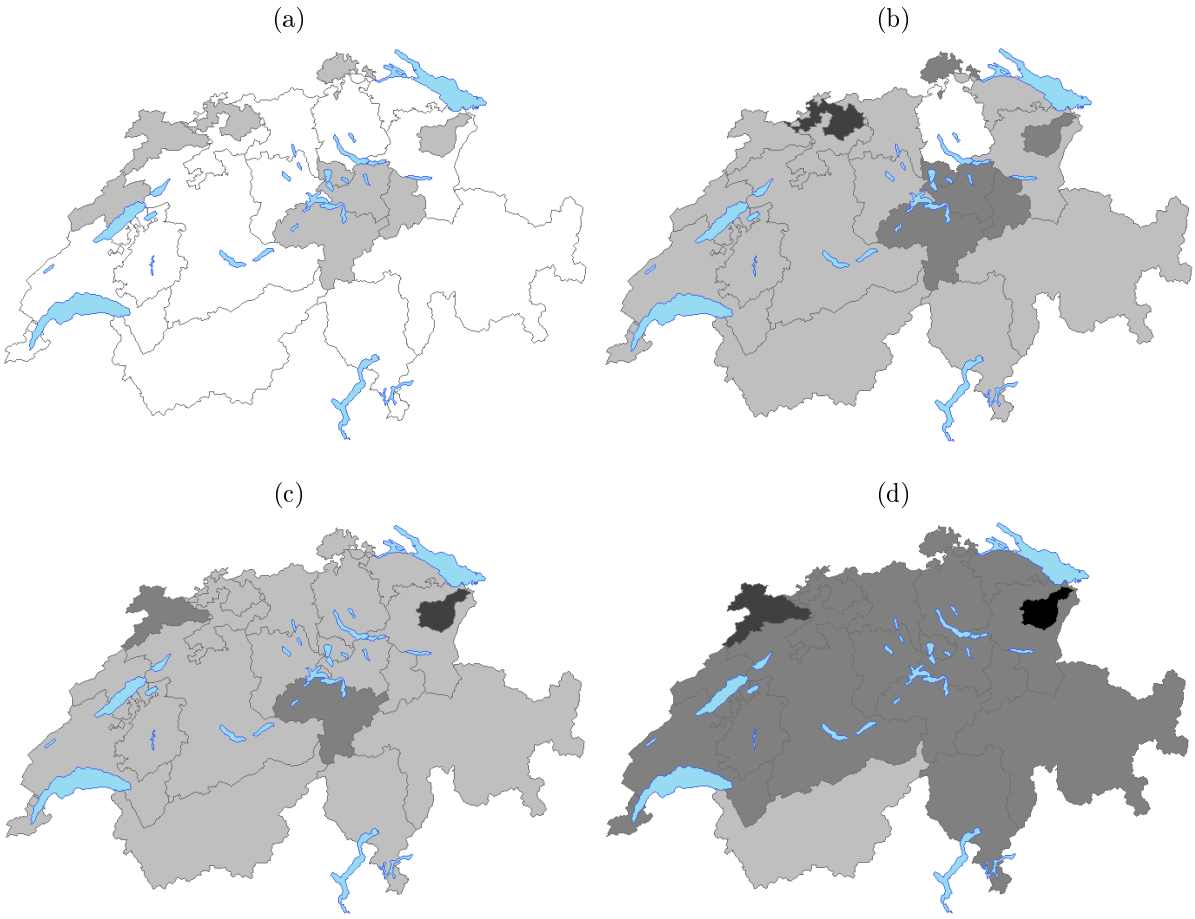


Figure 3.9: Maps showing the width of the 95% credible intervals relative (to the actual tobacco-related cancer mortality rates (shown in Fig 6–7) for males in (a) period 6 (2009–2013); (b) period 7 (2014–2018) and females in (c) period 6; (d) period 7.

	Males			Females		
Canton	2004–2008	2009–2013	2014–2018	2004–2008	2009–2013	2014–2018
ZH	43.0	38.7 (36.4, 41.3)	35.0 (31.5, 39.2)	20.6	20.6 (16.8, 24.6)	21.6 (16.6, 27.5)
BE	46.3	42.5 (39.7, 46.2)	39.2 (35.1, 45.4)	18.1	19.1 (15.6, 22.9)	20.9 (15.4, 26.8)
LU	44.6	38.2(34.3, 43.7)	33.5 (28.0, 42.6)	16.3	17.3 (14.3, 21.4)	18.8 (14.4, 24.9)
SZ	44.3	39.9 (34.0, 47.3)	35.2 (27.5, 45.9)	19.2	24.8 (19.8, 30.2)	28.6 (20.5, 37.7)
UR/OW/NW	41.3	36.9 (31.2, 44.2)	34.1 (26.5, 44.4)	17.6	20.7 (16.1, 26.6)	24.4 (17.3, 34.8)
GL	54.1	50.7 (38.0, 62.1)	49.4 (33.9, 66.1)	22.1	21.0 (16.6, 25.3)	23.8 (17.5, 31.2)
ZG	36.6	36.3 (30.2, 42.7)	33.9 (26.0, 43.1)	16.0	21.2 (16.2, 26.7)	24.2 (16.1, 33.5)
FR	58.5	50.4 (44.8, 57.0)	41.9 (34.2, 52.0)	21.3	21.5 (17.6, 25.9)	22.9 (17.6, 30.1)
SO	48.8	44.8 (40.0, 50.1)	40.6 (33.8, 48.8)	21.1	21.9 (16.9, 26.6)	24.1 (16.7, 31.3)
BS	58.1	53.3 (47.6, 60.3)	49.3 (41.3, 60.5)	27.4	27.9 (22.2, 35.0)	29.7 (21.8, 40.3)
BL	44.3	40.8 (35.1, 52.5)	38.2 (29.5, 63.5)	19.4	20.3 (16.3, 25.6)	21.4 (15.9, 29.2)
SH	50.6	45.2 (37.4, 54.2)	42.0 (31.6, 53.8)	17.4	22.6 (17.6, 27.6)	26.2 (17.9, 34.0)
AR/AI	34.0	32.1 (23.8, 38.8)	28.2 (17.8, 37.7)	13.2	15.9 (6.3, 21.1)	16.2 (2.8, 25.6)
SG	49.1	45.9 (41.9, 53.2)	43.2 (37.0, 55.8)	18.0	18.3 (15.2, 22.9)	19.4 (15.2, 26.6)
GR	51.1	47.5 (42.2, 53.9)	44.4 (36.7, 54.2)	19.9	18.8 (15.8, 24.5)	19.7 (15.4, 27.9)
AG	43.6	39.1 (35.8, 42.7)	34.6 (30.1, 40.0)	19.5	21.6 (17.9, 27.0)	23.8 (18.2, 32.3)
TG	45.3	44.5 (39.5, 49.9)	42.2 (35.5, 50.2)	18.3	20.6 (16.4, 25.1)	22.7 (16.8, 29.4)
TI	52.8	48.1 (43.9, 53.4)	44.1 (37.8, 52.9)	20.9	23.9 (19.3, 29.7)	26.1 (19.5, 35.1)
VD	54.5	49.9 (45.9, 56.2)	46.1 (39.8, 56.4)	22.6	23.6 (19.0, 28.6)	25.0 (18.6, 32.8)
VS	57.2	52.9 (47.7, 58.6)	49.1 (41.8, 57.9)	21.2	23.6 (19.9, 27.8)	26.3 (21.0, 33.4)
NE	64.3	58.1 (51.7, 68.0)	54.4 (45.6, 69.7)	24.6	25.1 (20.9, 31.5)	27.5 (21.1, 37.2)
GE	49.5	44.9 (40.8, 49.5)	40.8 (34.7, 48.0)	18.6	19.3 (15.7, 23.6)	19.1 (14.4, 25.3)
JU	57.8	55.6 (46.7, 65.6)	52.5 (40.7, 66.5)	22.0	21.5 (17.6, 28.7)	24.2 (18.3, 38.0)

Table 3.3: Age-standardized tobacco-related cancer mortality rates (per 100 000), observed (during 2004–2008) and predicted (posterior median during 2009–2013 and 2014–2018) at cantonal level for males and females age ≥ 50. The numbers in the brackets correspond to 95% credible intervals.

3.5.3 Bayesian APC model formulations

Bayesian Poisson regression models were fitted for each gender and age category. Following the APC approach, the canton-specific effects of age, period and cohort were used to model and predict tobacco-related cancer deaths μ_{ijk}, where population size n_{ijk} of each age group i and period j was considered as the exposure. As the analysis covered 26 coherent regions, models were further developed including gender- and canton-specific random effects ϕ_k $(k = 1, \ldots, 26)$ accounting for spatial correlation.

$$log(\mu_{ijk}) = log(n_{ijk}) + \alpha_{ik} + \beta_{jk} + \gamma_{j-i,k} + \phi_k \tag{3.2}$$

In addition, we extended the model to consider under-/overdispersion assuming that mortality counts follow a Negative Binomial. The Negative Binomial is an extension of the Poisson distribution and it includes an additional parameter to account for the extra-variation.
To complete Bayesian model formulation, priors are specified for the unknown parameters. For the first two groups of the time effects (i. e. age, period and cohort) a normal distribution was considered with mean 0 and vague precision depending on the overall precision of the remaining groups. For the remaining groups, a second-order autoregressive process was assigned allowing for dependence and smoothing (Bray, 2002). For example, for the age effects the above prior is formulated as follows:

$$\alpha_1 \sim \mathcal{N}(0, 1.0\text{E-}06\tau_\alpha) \tag{3.3}$$

$$\alpha_2|\alpha_1 \sim \mathcal{N}(0, 1.0\text{E-}06\tau_\alpha) \tag{3.4}$$

$$\alpha_i|\alpha_{1,\ldots,i-1} \sim \mathcal{N}(2\alpha_{i-1} - \alpha_{i-2}, \tau_\alpha), \text{for } 3 \leq i \leq I \tag{3.5}$$

where I is the total number of age groups. Uninformative priors were assigned to the precision parameter of each effect. Spatial random effects were assumed to derive from a Conditional Autoregressive (CAR) process Besag et al. (1991); Bernardinelli and Montomoli (1992). Dependence in space among the cantons was introduced by the conditional prior distribution of the ϕ_k with

$$\phi_k \mid \boldsymbol{\phi}_{-k} \sim \mathcal{N}\left(\frac{\sum_{\substack{q=1 \\ q \neq k}}^{N} c_{kq}\phi_q}{w_k}, \frac{\sigma^2}{w_k}\right) \tag{3.6}$$

	Cantonal		Language regions		National	
	Males	*Females*	*Males*	*Females*	*Males*	*Females*
Nordpred						
Poisson	293	234	13.25	3.99	1.20	0.19
Power 5	311	214	12.80	3.55	1.13	0.17
Bayesian APC model						
Poisson	113	104	3.57	1.98	0.71	0.08
Poisson spatial	109	101				
Negative Binomial	145	106	2.73	2.15	0.28	0.09
Negative Binomial spatial	175	101				

Table 3.4: Results of empirical projections of gender-specific tobacco-related cancer mortality given by the sum of squared residuals for the projected periods.

where $\boldsymbol{\phi}_{-k} = (\phi_1, \ldots, \phi_{k-1}, \phi_{k+1}, \ldots, \phi_N)$ and c_{kq} indicates the degree of spatial influence of canton k to the remaining ones, taking the value 1 if they are adjacent and 0 otherwise and w_k represents the number of neighbours of canton k.

3.5.4 Model validation

For validation purpose models were applied for empirical projection of age- and gender-specific tobacco-related cancer mortality at cantonal, national and language region level in periods with known mortality. Model performance was evaluated by comparing model-based projections with observed data in 1999–2003 and 2004–2008 using the sum of squared residuals (SSR).

$$SSR = \sum_{i,j,k} \frac{(\mathrm{R_{ijk}} - \hat{\mathrm{R}}_{\mathrm{ijk}})^2}{\mathrm{R_{ijk}}} \tag{3.7}$$

where $\mathrm{R_{ijk}}$ and $\hat{\mathrm{R}}_{\mathrm{ijk}}$ are the observed and estimated rates for canton k, age group i and period j, respectively. Table 3.4 shows the SSR for the two periods of the empirical projection of gender-specific tobacco-related cancer mortality. Regarding the projections at cantonal level the spatial Poisson model performed best for males. The spatial models showed best predictive ability for females, whereas the Negative Binomial ones gave more precise estimates. Therefore, spatial Poisson and Negative Binomial models were applied to project cantonal age- and gender-specific tobacco-related cancer mortality in Switzerland. For projections at country and language regions Negative Binomial and Poisson models were applied for males and females, respectively.

Chapter 4

A Bayesian generalized age-period-cohort power model for cancer projections

Verena Jürgens[1,2], Silvia Ess[3], Thomas Cerny[4], Penelope Vounatsou[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Cancer Registry of St. Gallen & Appenzell, St. Gallen, Switzerland
[4] Medical Oncology-Hematology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland

This paper has been submitted to *Statistics in Medicine*.

Abstract

Age-Period-Cohort (APC) models are the state of art in cancer projections, assessing past and recent trends and extrapolating mortality or incidence data into the future. Nordpred is a well established software, assuming a Poisson distribution for the counts and a log- or power-link function with fixed power, however its predictive performance is poor for sparse data. Bayesian models with log-link function have been applied but they can lead to extreme estimates. In this paper, we address criticisms of the above models by providing Bayesian formulations based on a power-link and develop a generalized APC power-link model (GPM), which assumes a random rather than fixed power parameter. In addition, a power model with a fixed power parameter of five was formulated in the Bayesian framework. The predictive performance of the new models was evaluated on Swiss lung cancer mortality data using model-based estimates of observed periods. Results indicated that the GPM provides best estimates for male and female lung cancer mortality. The gender-specific models were further applied to project lung cancer mortality in Switzerland during the periods 2009–2013 and 2014–2018.

Keywords: Bayesian inference; cancer projection; power model; Nordpred; lung cancer mortality

4.1 Introduction

Cancer projections estimate the future burden of the disease. They provide important information for health planning and evaluation of intervention effects, e. g. screening or changes in diagnostic techniques.

The most common modelling approaches for projecting cancer rates are time series and Age-Period-Cohort (APC) models (Lee et al., 2011; Qiu et al., 2010). Time series models have been applied assuming a Poisson distribution for the disease counts. The models include autoregressive error terms (Wingo et al., 1998) and/or time trends fitted by linear (Lee et al., 2011), polynomial (Wingo et al., 1998), piecewise linear (Kim et al., 2000) or spline curves. APC models typically include three components – *age*, corresponding to the age of death, *period*, representing the time period that the death was recorded and *cohort*, giving the birth cohort of the person that died. Usually these components are stratified into 5-year intervals which are identified by the relation *cohort* = *period* - *age*. The dependence between the components is well-known as the identification problem however it does not affect estimation of the projected rates. Intervals smaller than 5 years have been considered, however the smaller the interval the more likely to lead to sparse or zero death counts. Age is assumed to have the highest effect on cancer mortality and incidence rates, whereas period or cohort effects can be neglected in some settings. Short-term projections (≤ 5 years) are mainly based on age-period models, and long-term ones are based on age-period-cohort-models. However, depending on the dataset, cohort effects might also be important for short-term projections. Preliminary analyses are carried out to assess the contribution of cohort effects. Models may include a drift parameter to measure the average trend. The drift is often considered constant throughout the periods or reduced by a fixed amount (i. e. 25%) (Møller et al., 2002).

APC models are fitted within a generalized linear model framework. Different link functions can be considered to relate the predictors with the mean (Mallick and Gelfand, 1994; Pregibon, 1980). Osmond (1985) used a Poisson regression model with a log-link function to estimate age- and period-specific rates. Møller (2004) proposed the power model, using a power rather than a log-link function with a fixed power equal to 5 derived by empirical estimation. Both, the Poisson power- and log-link models are implemented in the R (R Development Core Team, 2011) software package Nordpred developed by the Cancer Registry of Norway. A Bayesian formulation of the Poisson APC model with log-link was suggested by Bray (2002), assigning autoregressive priors to the effect of age, period and cohort. Second-order autoregressive processes were appropriate in several settings. Bayesian APC models have been applied frequently to model for example lung cancer mortality in West Germany (Knorr-Held and Rainer, 2001), in France (Eilstein et al., 2008) or breast cancer mortality in Spain (Clèries et al., 2006).

Several reviews compared different models for cancer projections (Qiu et al., 2010) showing that

the Poisson power-link and Bayesian Poisson log-link model outperformed other approaches (Lee et al., 2011; Cancer Projections Network (C-Projections), 2010). The latter model assumes an exponential growth of the predictions. It has been criticized that it might give extreme predictions (Qiu et al., 2010; Bray and Møller, 2006; Møller et al., 2005). The power-link overcomes this problem however it may inaccurately estimate low or even zero counts. The Bayesian model formulation smooths age, period and cohort effects obtaining valid estimates for sparse rates or even zero counts (Qiu et al., 2010; Baker and Bray, 2005). In addition it can provide uncertainty estimates of rates and assess model performance taking into account the estimation error. To our knowledge a Bayesian formulation of the Poisson power-link model has not been developed yet. Furthermore, it is unclear whether all applications of the power-link model support a fixed power parameter equal to 5.

In this paper, we formulated the power model within the Bayesian framework and developed a generalized model considering a random rather than fixed power parameter. The above models were compared with the ones using the log-link function. All models were applied on gender-specific lung cancer mortality data from Switzerland and compared with the fixed power- and log-link model fitted in Nordpred using maximum likelihood. The sum of squared residuals (SSR) of the projected rates was calculated for all models, while the logarithmic score (LS) (Ntzoufras, 2009) and Deviance Information Criterion (DIC) was applied to the Bayesian models to assess their predictive performance. The gender-specific model with the best accuracy was employed to project lung cancer mortality rates for the periods 2009–2013 and 2014–2018.

4.2 Methodology

4.2.1 Model formulations

In this study, we developed Bayesian APC models assuming that the observed age- and period-specific death counts N_{ij} follow a Poisson $N_{ij} \sim \text{Pois}(\mu_{ij})$ distribution and using a model-specific link-function g.

The mean of the age- and period-specific death counts μ_{ij} was regressed on the effects of age, period and cohort, using the corresponding population as the offset. In particular, let α_i, β_j and γ_k $(k = I - i + j)$ be the effects of age, period and cohort and N_{ij} and M_{ij} be the age- and period-specific death counts and population, respectively.

$$g\left(\frac{\mu_{ij}}{M_{ij}}\right) = \alpha_i + \beta_j + \gamma_k \tag{4.1}$$

Under a log-link as well as power-link function the regression model relating the mortality rate to the time effects reads as follows:

$$\begin{aligned} &\textit{Log-link model:} && \log(\mu_{ij}) = \log(M_{ij}) + \alpha_i + \beta_j + \gamma_k \\ &\textit{Power-link model:} && \mu_{ij}^{\frac{1}{5}} = M_{ij}^{\frac{1}{5}}(\alpha_i + \beta_j + \gamma_k) \end{aligned} \tag{4.2}$$

In the above model the power is assumed to be fixed. We extended the model by considering the power to be a parameter estimated by the data and called it generalized APC power model (GPM). The power w was assumed to be random, as the degree of the growth of the rate might change, depending on the given dataset. The model was formulated as follows:

$$\textit{GPM:} \qquad \mu_{ij}^{\frac{1}{w}} = M_{ij}^{\frac{1}{w}}(\alpha_i + \beta_j + \gamma_k) \tag{4.3}$$

The above model overcomes estimation of extreme rates that may be obtained from the log-link. Bayesian simulation-based estimation allows flexible prior specification which smooths time effects.

4.2.2 Prior specifications

To complete Bayesian specification, prior distributions were assigned to the parameters. Following Bray (2002), time effect-specific (i. e. age, period and cohort) smoothing prior formulations were considered. The first two groups of each time effect were restricted to follow a normal distribution with mean 0 and vague precision. The remaining groups were assumed to follow a second-order autoregressive processes, depending each effect on its two predecessors, to allow for dependence and smoothing (Bray, 2002). The above prior distributions are formulated as follows (illustrated for the effect of age):

$$\alpha_1 \sim \mathcal{N}(0, 1.0E{-}06 * \tau_\alpha) \tag{4.4}$$

$$\alpha_2 | \alpha_1 \sim \mathcal{N}(0, 1.0E{-}06 * \tau_\alpha) \tag{4.5}$$

$$\alpha_i | \alpha_{1,\ldots,i-1} \sim \mathcal{N}(2\alpha_{i-1} - \alpha_{i-2}, \tau_\alpha), \text{for } 3 \leq i \leq I \tag{4.6}$$

where I is the total number of age groups. Non-informative priors were assigned to the precision parameters τ_α, τ_β and τ_γ for the effect of age α, period β and cohort γ, respectively. For comparison purposes we used the same prior distributions across models.

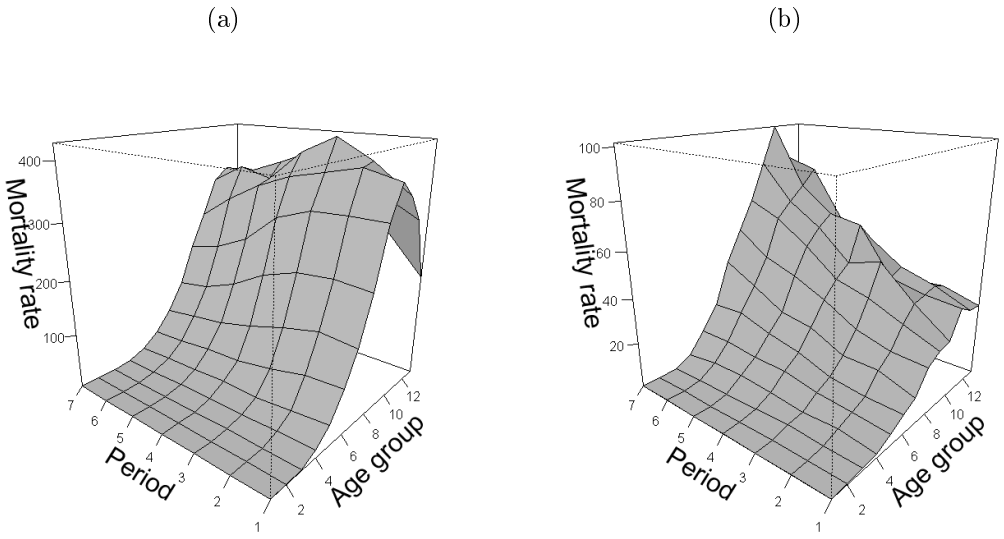


Figure 4.1: Observed crude lung cancer mortality rate (per 100 000 inhabitants) for males (a) and females (b) from 1974–1978 to 2004–2008.

A discrete prior distribution was assigned on the power parameter w, assuming that $w \sim cat(p[])$. This prior formulation specifies a set of possible values $w \in \{1, \ldots, n\}$ with probabilites $p[w = i] = p_i$ and $\sum_{i=1}^{n} p[i] = 1$. In our specification, we consider w could take any integer from 1 to 10, each with probability equal to 0.1.

4.3 Assessing model predictive performance

Empirical projections, estimating mortality for observed periods, have been done applying the different models, i. e. Bayesian log-, power-link model and the GPM, the Nordpred log- and power-link model. In addition, the performances of the log-linear model (excluding the cohort effect), which is generally used for short-term projections (Hakulinen and Dyba, 1994), and the simple averaging method (Qiu et al.) have been assessed. The latter has been recommended for the projection of low count data and is based on the assumption that the population size will change in the future, but the mortality rate remains the same as in the current period. Model-specific results have been validated by the LS, DIC, SSR as well as graphical assessment of the projected rates.

4.4 Application

In Switzerland the National Program Tobacco 2008–2012 (NPT 2008–2012) was launched to reduce tobacco-related morbidity and mortality in the country. The initiative strategic goals were set

until the end of 2012 to accomplish the mission. Projections of tobacco-related cancer mortality based on past trends, not taking into account recent interventions as the NPT 2008–2012, may increase awareness on the program. However, interpretation on the impact of the program based on the projected rates should be done carefully as the difference between estimated and observed might be due to several factors and interventions are only one possibility among others.
Count mortality data from death certificates and population size at national level were provided from the Swiss Federal Statistical Office (FSO). The time range considered for both datasets, was 1974–2008 and was splitted into seven periods (1974–1978, ..., 2004–2008). Age was classified in five-year groups (25–29,..., older than 85) and 19 ten-year overlapping cohorts were constructed following the relation cohort = period - age. Gender-specific mortality rates were age-standardized directly using Segi World population (Segi, 1960). In 1995 the rules related to coding death certificates have been changed in Switzerland – priority of certain causes has been removed from the regulations. Therefore, the number of these cause-specific deaths reported before and after this changes differed and have to be adjusted. The FSO provides a disease-related correction factor to adjust death reports before 1995 (Berrut and Junker, 2008).
Exploratory analysis of tobacco-related cancer mortality from 1974–2008 indicated different trends for each gender (see Figure 4.1). Rates for men remain stable since the last decade, while female lung cancer mortality increased steadily.
Empirical projections were carried out for the two last periods (1999–2003, 2004–2008) with known mortality rates to validate the model performance. After validation, the model with the best fit was applied to project gender-specific lung cancer mortality in Switzerland for the periods 2009–2013 and 2014–2018. In addition, predictive performance of the models was assessed for a shorter period of observed mortality, i.e. 1984–1988, ..., 2004–2008. Model formulation and fitting was based on Markov chain Monte Carlo (MCMC) (Gilks et al., 1996) and was done in WinBUGS (Imperial College and MRC, London, UK). Convergence was assessed by the Gelman and Rubin diagnostic test within the R package coda.

4.4.1 Model predictive performance

Exploratory analysis assuming different drifts indicated no effect on the fitting and the estimates. Therefore, the drift parameter was considered to be zero. Model predictive ability was evaluated by different metrics. Results were compared with the Nordpred equivalent whenever possible (see Table 4.1). Firstly, the SSR was calculated for all models and visual assessment was done by comparing model-based estimates with the observed rates for the projected periods. As a next step, Bayesian models were compared based on the LS and DIC.
The results of the empirical projections are shown in Table 4.1. SSR and visual assessment of the projected rates (Fig. 4.2) indicated that Bayesian models performed better than the equivalent implemented in Nordpred. For males, the power-link outperformed the log-link model. The latter

	Males			Females		
	LS	DIC	$SSR_{6,7}$	LS	DIC	$SSR_{6,7}$
Bayesian						
Log-link	-207.7	589.2	40.1	-207.4	494.3	8.9
Power-link	-206.5	590.3	30.9	-208.4	490.4	9.9
Log-linear	-340.5	868.4	142.6	-259.5	516.6	13.1
GPM	-205.7	597.3	21.9	-206.5	496.8	8.6
Nordpred						
Log-link			53.1			13.7
Power-link			41.6			11.4
Simple averaging			60.1			86.4

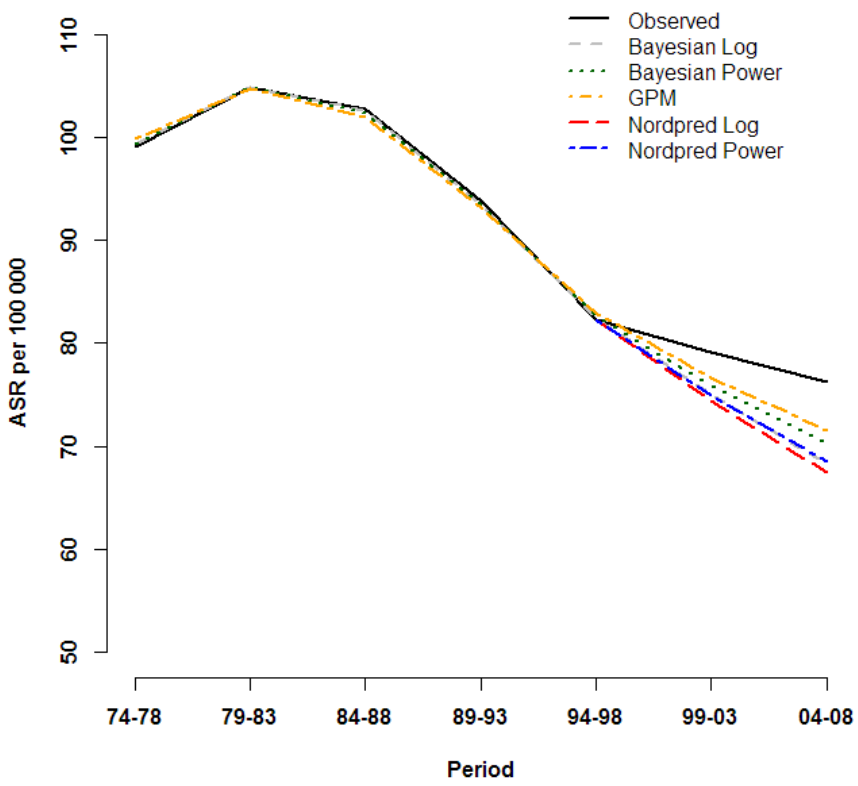
Table 4.1: SSR of empirical projections of gender-specific lung cancer mortality rates, logarithmic score (LS) and DIC for the Bayesian models.

provided extremely low estimates. For females, the log-link was better thatn the power-link, however it overestimated mortality. The Bayesian log-linear model and the simple averaging method performed worst. Best predictive ability for for both genders was given by the GPM, as indicated by the SSR, LS as well as the plotted estimates of the empirical projection. The random power parameter for males and females was estimated to be 3 and 7, respectively. For males, the mortality rates level off during the projected period. Female mortality rates continue increasing during 1999–2008 following a slightly steeper trend than during 1974–1998.
Empirical projections based on a shorter time period of observed mortality (1984–1988, ..., 2004–2008) indicated best performance of the Bayesian log-link model for males. However, the predictive performance of the GPM based on the longer observed period was better than the one based on the shorter fitting period. For females, the GPM, indicated best predictive ability for the shorter as well as for the longer fitting period, however, the latter provided better fit.

4.4.2 Projections

Gender-specific lung cancer mortality rates were projected for the periods 2009–2013 and 2014–2018, based on observed data from 1974–2008. Future rates were predicted for each gender by using the GPM and the Bayesian log-link model for males and females, respectively. Based on evaluation, different models were applied for gender-specific projections. To account for potential age-specific trends, age was splitted into four age categories 25–44, 45–64, 65–84 and above 85 years. Overall and age category-specific rates were age-standardized by Segi World population. Observed lung cancer mortality rates by gender are illustrated in Figure 4.1. Figure 4.3 and Table 4.2 indicate a steadily declining trend for male lung cancer mortality since 1979–1983. This can be

(a)



(b)

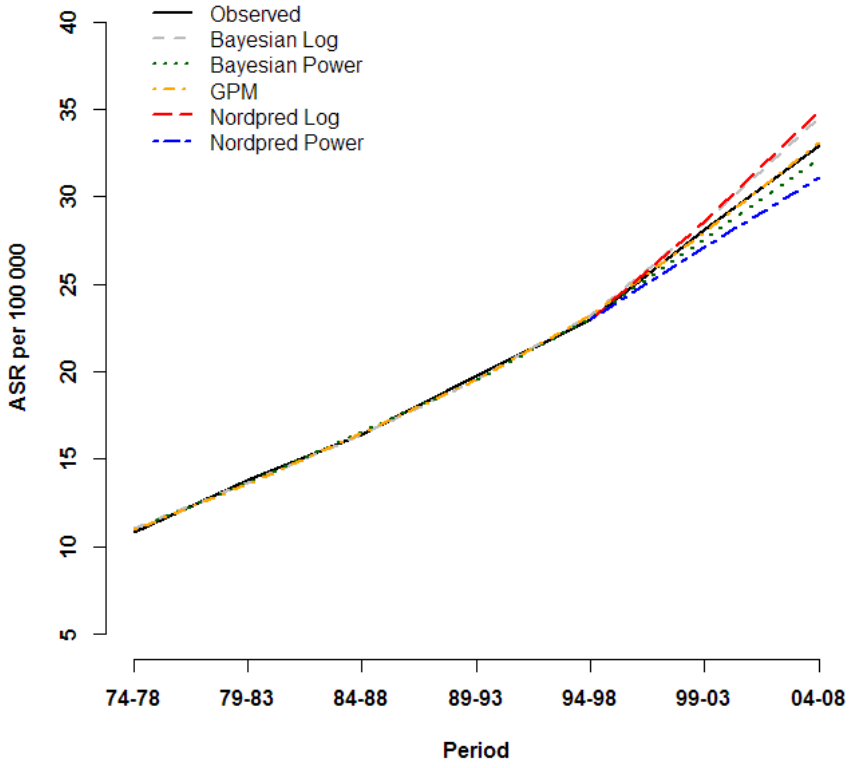


Figure 4.2: Observed and predicted Swiss lung cancer mortality rates for males (a) and females (b).

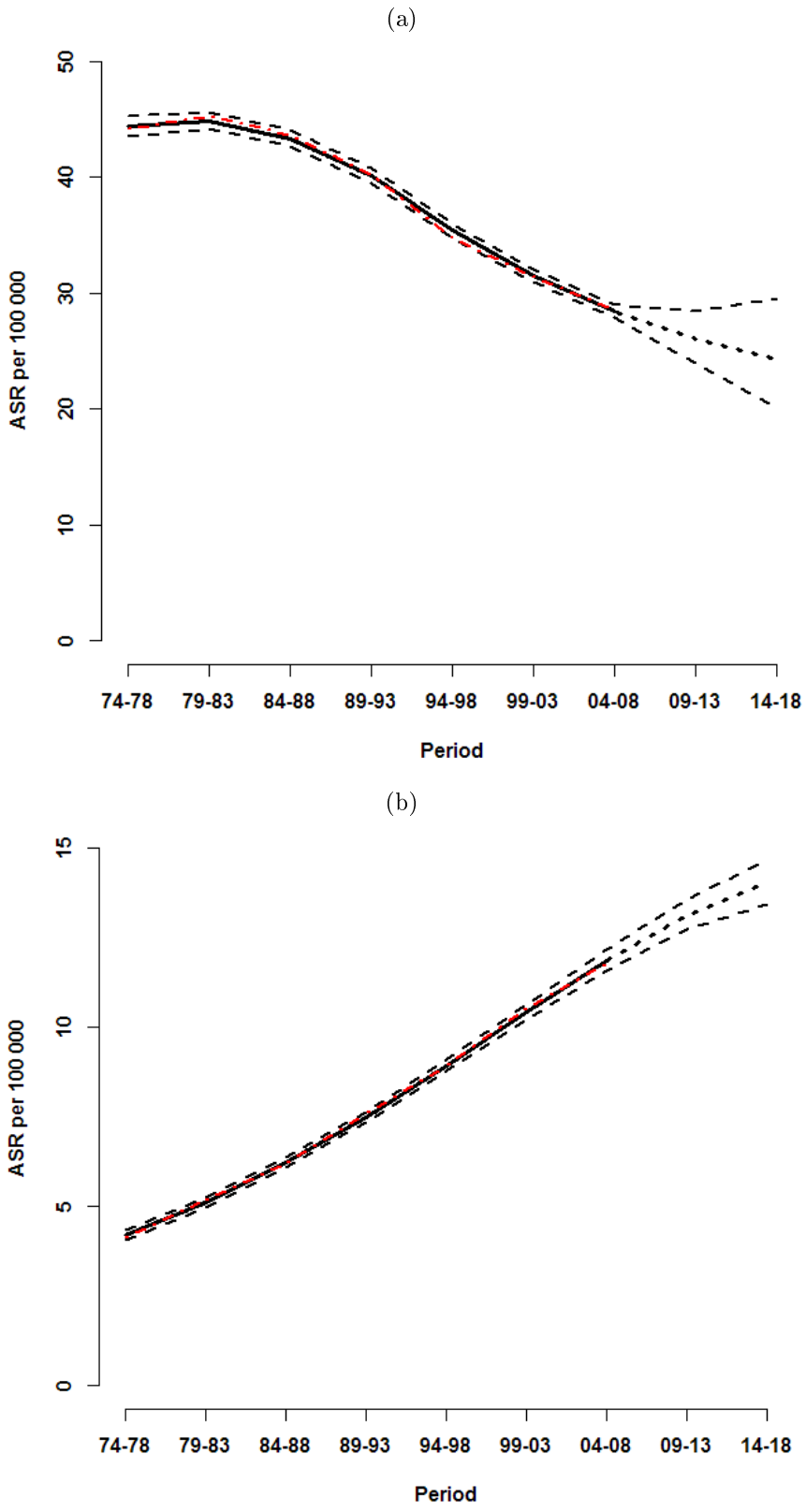


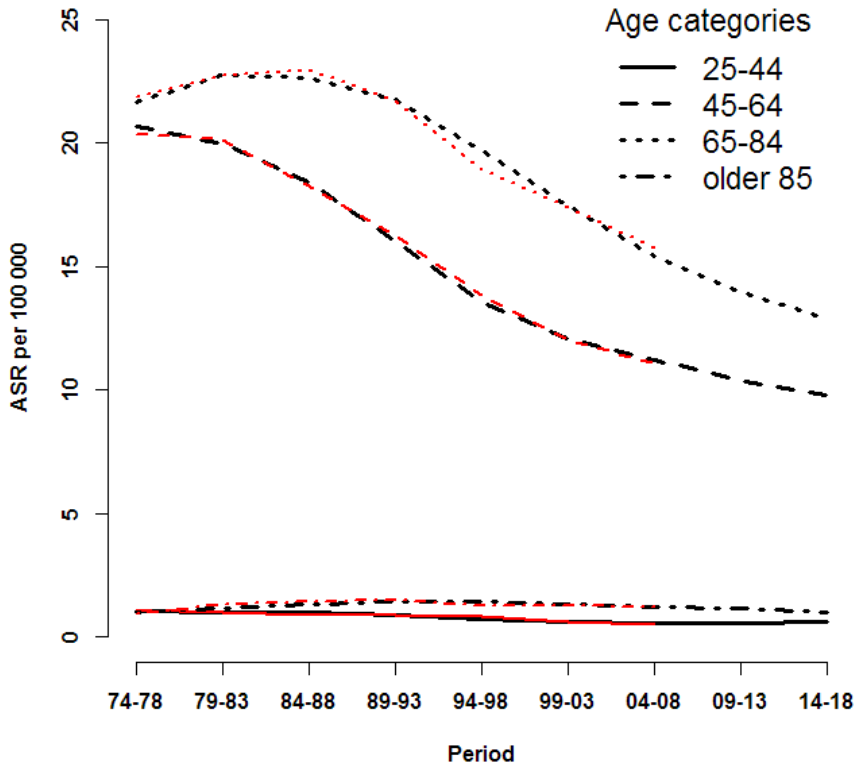
Figure 4.3: Observed (red), model-based estimates of fitted (1974–1978 to 2004–2008) and projected (2009–2013, 2014–2018) Swiss age-standardized lung cancer mortality rate for males (a) and females (b) (per 100 000 inhabitants) with 95% Credible Bounds (CB).

seen for all groups, except the ones older than 85, which shows a rising trend until 1993, followed by a decline (Fig. 4.4).

Female lung cancer mortality was estimated to be more than three times higher in 2014–2018 in comparison to 1974–1978. Age category-specific plots present a stable trend for females aged 45–64 in 2004–2008.

Final projections for males estimated the random power parameter to be 3 and 7, respectively.

(a)



(b)

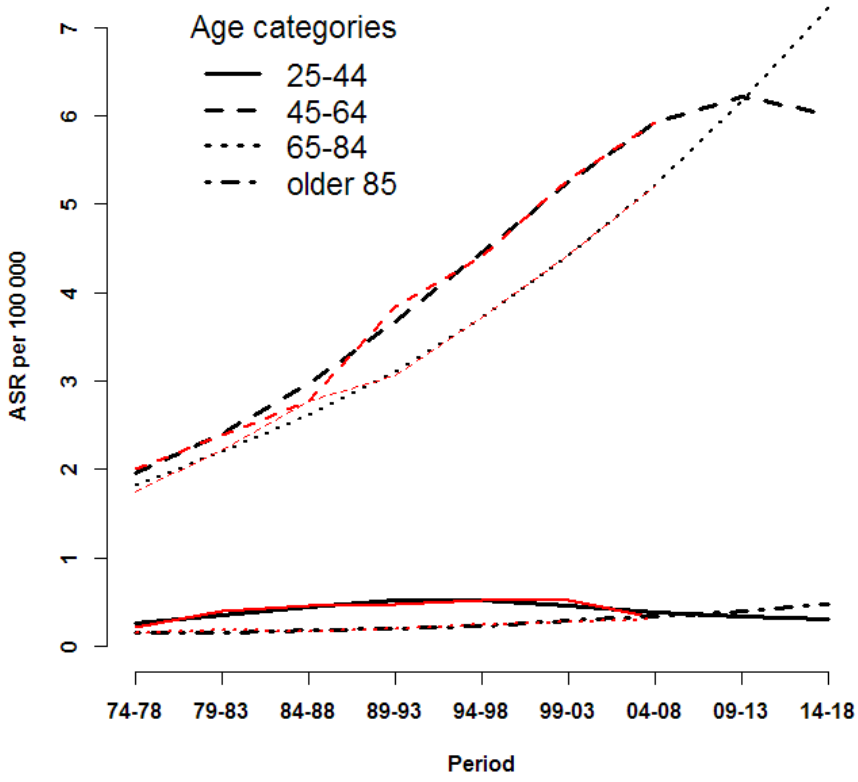


Figure 4.4: Observed (red), model-based estimates of fitted (1974–1978 to 2004–2008) and projected (2009–2013, 2014–2018) Swiss age-standardized lung cancer mortality rate for males (a) and females (b) (per 100 000 inhabitants) by age category.

	Fitted rates							Projected rates	
	74–78	79–83	84–88	89–93	94–98	99–03	04–08	09–13	14–18
Males									
Overall	44.35 (43.52;45.21)	44.84 (44.14;45.55)	43.32 (42.62;44.01)	40.09 (39.50;40.78)	35.38 (34.75;35.94)	31.49 (30.95;32.02)	28.41 (27.87;28.98)	26.08 (23.93;28.46)	24.28 (20.17;29.55)
25–44	1.07 (0.98;1.17)	1.00 (0.93;1.09)	0.98 (0.91;1.06)	0.88 (0.82;0.95)	0.72 (0.65;0.78)	0.62 (0.56;0.68)	0.57 (0.50;0.65)	0.57 (0.41;0.80)	0.59 (0.31;1.19)
45–64	20.64 (20.11;21.16)	19.94 (19.52;20.38)	18.40 (17.99;18.80)	16.02 (15.68;16.41)	13.55 (13.20;13.86)	12.09 (11.79;12.40)	11.20 (10.89;11.53)	10.41 (9.37;11.56)	9.79 (7.89;12.26)
65–84	21.65 (21.22;22.09)	22.71 (22.33;23.10)	22.62 (22.25;22.97)	21.75 (21.40;22.11)	19.69 (19.37;20.02)	17.46 (17.16;17.75)	15.41 (15.11;15.72)	13.92 (13.00;14.95)	12.89 (11.11;15.04)
older 85	1.00 (0.89;1.11)	1.18 (1.12;1.26)	1.33 (1.27;1.39	1.44 (1.38;1.50)	1.42 (1.36;1.47)	1.32 (1.26;1.37)	1.23 (1.18;1.28)	1.15 (1.06;1.24)	1.00 (0.85;1.18)
Females									
Overall	4.21 (4.07;4.36)	5.12 (4.98;5.26)	6.22 (6.08;6.36)	7.50 (7.35;7.63)	8.94 (8.78;9.09)	10.44 (10.23;10.63)	11.87 (11.59;12.17)	13.12 (12.72;13.57)	14.03 (13.42;14.64)
25–44	0.27 (0.24;0.30)	0.35 (0.32;0.39)	0.45 (0.41;0.50)	0.52 (0.48;0.57)	0.52 (0.48;0.57)	0.47 (0.42;0.51)	0.39 (0.33;0.45)	0.34 (0.26;0.44)	0.31 (0.19;0.49)
45–64	1.96 (1.87;2.06)	2.40 (2.30;2.50)	2.97 (2.85;3.08)	3.67 (3.56;3.79)	4.46 (4.32;4.58)	5.25 (5.09;5.40)	5.92 (5.71;6.14)	6.21 (5.91;6.54)	6.01 (5.60;6.44)
65–84	1.82 (1.75;1.91)	2.20 (2.13;2.28)	2.62 (2.54;2.70)	3.11 (3.03;3.19)	3.72 (3.63;3.82)	4.42 (4.31;4.55)	5.22 (5.07;5.37)	6.17 (5.96;6.40)	7.22 (6.93;7.54)
older 85	0.15 (0.13;0.18)	0.16 (0.15;0.18)	0.18 (0.16;0.19)	0.20 (0.18;0.22)	0.24 (0.23;0.26)	0.30 (0.28;0.32)	0.34 (0.32;0.36)	0.40 (0.38;0.43)	0.48 (0.45;0.51)

Table 4.2: Model-based estimates (GPM) of fitted and projected age-standardized gender-specific lung cancer mortality rate (per 100 000 inhabitants) with 95% CB.

4.5 Discussion

In this paper, we developed a Bayesian APC power model with fixed power of 5 and a generalized APC power model for cancer mortality/incidence projections which overcomes limitations of well established models. The commonly used power model was formulated within the Bayesian framework and it was further extended to allow for a random power parameter instead of the fixed one proposed by Møller (2004).

Model performance was compared to the frequently used Bayesian APC model with a log-link function. We assumed that the counts arise from a Poisson distribution. The models were compared to the commonly used ones implemented in the Nordpred software assuming the log- and the power(5)-link. The predictive performance of the new models was evaluated on Swiss lung cancer mortality data using model-based projections of observed periods. The model with the best fit was applied to project age- and gender-specific lung cancer mortality rates in Switzerland for the periods 2009–2013 and 2014–2018.

The GPM performed best for males and females. For both gender, the log-linear and simple averaging method gave poorest fit. The latter assumes continuation of the trend observed in the recent period and is sensitive to the projected population (Qiu et al.). For the empirical projection the population was known, which suggests that the assumption of a continuing trend was violated. Figure 4.2 illustrates the changing trend for both gender – a flattening decrease for males and a steepening increasing trend for females.

The unsatisfactory performance of the log-linear model can be explained by the missing cohort effect, which was found to be strongly related to lung cancer mortality rates (Lee et al., 1990). This relation can be explained by the gender-specific tobacco epidemics.

Møller et al. (2003) compared several projection methods, e. g. APC models using a power- or a log-link on several cancer incidence data from five Nordic countries. They also found that models using a power instead of a log-link function have better predictive ability. The authors assessed models with power parameters fixed to 5 and 2 and concluded that for short term predictions (≤ 5 years) the model based on the power of 2 outperformed the one with a power of 5. This result was not seen for the long term projections and they pointed out that too small values may implicate numerical problems.

The Canadian research group from the Alberta Health service evaluated long-term projections of cancer mortality and incidence data based on Nordpred, Generalized Additive and Bayesian Models (Qiu et al., 2010). The Bayesian specifications included the log-link assuming different prior distributions for the effects of age, period and cohort. They concluded that the models in Nordpred using power 5 and the Bayesian model using second-order autoregressive priors for the age, period and cohort effects performed best, however the Bayesian one performed better considering mortality projections and the Nordpred achieved better results for incidence. Mortality

is more sparse than incidence and therefore more challenging for the modelling. The latter is true for cancer sites with high survival rates. Some authors (Qiu et al., 2010) argue against the use of Bayesian APC models due to the complexity of the MCMC fitting, however over the last years the number of scientific publications, using the Bayesian framework, is increasing steadily.
The analysis of the Swiss lung cancer mortality data indicated an ongoing decline for overall male mortality from 1974 onwards. Age-specific mortality followed the overall trend with exception the group of 65–84 years old, whose decreasing trend started after 1979. Despite mortality rates are decreasing, the number of projected deaths increased after 2009. Female mortality increases since 1974 for all age groups, however the group of 45–64 years old reaches a plateau around 2004–2008.
In general, results of projections should be interpreted carefully, as they depend on different factors such as model choice. Clements et al. (2005) analysed lung cancer mortality for females in five countries. They compared different approaches, which provided different trends for the future rates. The authors pointed out that the dependence of the projections on the form of the model is an important limitation. Furthermore, cancer projections are sensitive to population projections since the proportion of elderly people influences mortality rates and crude deaths counts. In our analyses, population data for 2009–2013 and 2014–2018 were provided by the Swiss Federal Office of Statistics which extrapolates populations under different scenarios regarding migration, birth and death rates. Our population estimates are based on non-extreme demographic changes.
Projections of lung cancer mortality are based on past trends, not taking into account recent interventions. The National Program Tobacco 2008–2012 (NPT 2008–2012) was launched to reduce tobacco-related cancer morbidity and mortality in the country. The initiative strategic goals were set until the end of 2012. Our findings could help to evaluate the program, as model-based estimates can be compared with observed data in the future.
Several metrics exist to assess the predictive performance of a model (Czado et al., 2009). We used the SSR (Baker and Bray, 2005) as well as the DIC and the LS (Ntzoufras, 2009) for the for the Bayesian models. In addition, the estimates rates have been plotted and predictive ability has been assessed graphically. For males and females, the different metrics gave the same conclusion.
In this paper, different functions have been applied to link the expected value of the mortality to the predictor. Their choice depend on the dataset under consideration. While a log-link function is more suitable to model a steep trend, a power-link function allows the specification of a certain degree of increase or decrease of the trend. In this study, gender-specific trends have been observed. Males had a rather decreasing trend which flattened at the end of the study period. For females an increasing trend was observed. Models based on the log-link function provided the most extreme estimates, underestimating and overestimating the mortality rate for males and females, respectively (Fig. 4.2). Estimates of the power-link models were less extreme and performed better, while the GPM performed best in both settings. Visualization of the estimated rates by the different link-functions clearly shows the under- and overestimation by the log-link

and power-link models and indicates that a power parameter lower/greater than 5 would be more appropriate to model Swiss male/female lung cancer mortality. This was confirmed by the estimated power of 3 and 7 by the GPM for males and females, respectively.
The proposed models extend the list of APC-models and as already discussed above they had better predictive performance in modelling the lung cancer data of our application. The models have a number of strengths that make them especially suitable for modelling low counts as we have shown in the analysis of female mortality data. Lee et al. (2011) and Qiu et al. suggested that the average counts in the few years prior to projection may be a reasonable approach for sparse or low count data. Our analysis showed that the simple average method performed worse than any other approach in the case of female low count mortality data and that Bayesian formulations outperform non-Bayesian approaches especially when modelling low counts. We are currently assessing the models in projecting cause-specific mortality other than lung cancer.
The Bayesian model formulations we developed can be extended to account for extra Poisson variation, present in sparse cancer data, by considering a Negative Binomial distribution which is a mixture of a Poisson distribution with a Gamma mean parameter (Ntzoufras, 2009). However, no over- or under-dispersion was present in the dataset under consideration. In fact, the Negative Binomial model did not improve the predictive performance of the Poisson model. Choosing a distribution other than the Poisson might be reasonable for the analysis of low count data, e. g. at sub-regional level, which would allow the formulation of a subgroup-specific dispersion parameter. Furthermore, potential heterogeneity can be captured by introducing random effects, however, this specification will increase the number of parameters to be estimated.
Our analyses projected mortality rates at national level. Regional differences can be important for health planning purposes. Model formulations can be easily extended to assess regional difference and spatio-temporal pattern across the country (Lee et al., 2011).

Acknowledgements

This work was supported by Oncosuisse [grant number KLS-02393-02-2009] and the Swiss National Science Foundation [grant number 32003B_135769]. We would like to thank the Swiss Federal Statistical Office for providing the data.

Chapter 5

Filling the gaps – back-calculation of lung cancer incidence

Verena Jürgens[1,2], Silvia Ess[3], Harish C. Phuleria[1,2], Martin Früh[4], Matthias Schwenkglenks[5], Harald Frick[6], Thomas Cerny[4], Penelope Vounatsou[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Cancer Registry of St. Gallen & Appenzell, St. Gallen, Switzerland
[4] Medical Oncology-Hematology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland
[5] Institute of Pharmaceutical Medicine (ECPM), University of Basel, Basel, Switzerland
[6] Cancer Registry Grisons & Glarus, Kantonsspital Graubünden, Chur, Switzerland

This paper has been submitted to *BMC Cancer*.

Abstract

Data on lung cancer incidence provide important information on the burden of the disease. In Switzerland, incidence data are estimated by the cantonal cancer registries. However, only half of the country is covered by cancer registries. On the other hand, mortality data are available at municipality level obtained from death certificates. Back-calculation models allow for estimation of incidence by linking subgroup-specific survival with mortality.

Mortality and incidence data from the Cancer Registry of St. Gallen-Appenzell (SGA) were used to estimate lung cancer survival probabilities. Bayesian back-calculation models were developed to estimate incidence from survival distributions and lung cancer deaths. The latter was extracted from the national mortality database which is maintained by the Federal Statistical Office (FSO). The proportion of miss-reported cause of death in the FSO data was calculated from the SGA cancer registry data and considered in the analyses. Conditional autoregressive models were employed to provide gender-specific smooth maps of age standardized incidence ratios.

Validation comparing observed and estimated incidence for cantons with cancer registries indicated good model performance. Smooth maps of lung cancer incidence for females showed higher estimates in the urbanized regions, while for males a rather homogeneous distribution was observed.

The proposed models improve earlier methodology and are important not only in mapping the spatial distribution of the disease but also in assessing temporal trends of lung cancer incidence in cantons without registries.

Keywords: Bayesian inference; lung cancer; incidence; survival

5.1 Introduction

Cancer registries provide essential information to assess trends of cancer survival, incidence and mortality in order to control and intervene. In Switzerland, 11 cancer registries cover 15 of the 26 cantons. Recently, attempts have been undertaken to establish cancer registries for the remaining ones, however in 2005–2009 cancer registries covered 97.7% of Western Switzerland and Ticino, while only 45.2% of the German-speaking part was covered (Federal Statistical Office, 2013a).
On the other hand, reliable information on cause-specific mortality at national level is given by the death certificates provided by the Federal Statistical Offices. Therefore, the lack of availability of cancer incidence data complicates the analysis space-time patterns of the cancer-specific burden. The Federal Office of Statistics (FSO) extrapolated observed incidence to regions without registration and provides gender-specific incidence estimates aggregated for two language-regions (German-speaking Switzerland, French- and Italian-speaking Switzerland combined) (Federal Statistical Office, 2013b). This approach assumes homogeneity of the cancer incidence between those regions covered and uncovered by cancer registries (Bouchardy et al., 2011).
Statistical approaches have been developed, which approximate incidence by observed mortality data (Mitton et al., 2011; Ong and Soo, 2006). An established method is the back-calculation model (Brookmeyer and Gail, 1988), which allows the estimation of incidence data by linking survival functions and mortality data via the equation:

$$M(t) = \int_0^t I(t-i) f(t-i, i) di \tag{5.1}$$

where $M(t)$ is the death count at time point t, $I(t-i)$ the incident case at time $(t-i)$ and $f(t-i, i)$ the probability density of dying at time t when diagnosed at $(t-i)$.
Model formulation requires an appropriate survival function to model the time to death since diagnosis. Early diagnosis and treatment options after detection influence survival. Some cancers, like lung cancer, are difficult to be detected early and therefore it is less likely that there is regional variation in patient survival. Other cancers, like breast cancer, can be diagnosed at an early stage. Therefore screening policies have an impact on survival. For the latter, it is expected to differ not only in individuals but also among the regions.
Lung cancer survival is rather low. In Switzerland, the 5-year survival for males is 10.0% and 13.4% for females in 1995–1999, while the majority dies within the first year (survival of 35.6% and 36.8% for males and females, respectively) (Bordoni et al., 2012).
A number of studies have tried to estimate the survival distribution applying different functions, either parametric or non-parametric (Collett, 1994). Mezzetti and Robertson (1999) modelled lung cancer survival using a mixed distribution, considering a fixed probability for failure during the first year after diagnosis. The survival for the years afterwards was assumed to follow an exponential distribution, which assumes a constant hazard $h(t) = \lambda$ over time t. De Angelis

et al. (1999) analysed colon cancer survival in Finland, considering a Weibull density to model failure time, which assumes the hazard to increase in time. Furthermore, a logistic function was considered to account for the proportion of cured cases. Recently, Armero et al. (2012) used a multi-state Bayesian analysis to examine lung cancer survival time and the progress of the disease. Sánchez et al. (2010) estimated cancer incidence in Spain, modelling incidence as a polynomial function of age, period and cohort and back-calculating its model parameters by the means of a maximum likelihood regression based on the observed death counts using the software MIAMOD (mortality-incidence analysis model) (Angelis et al., 1994). Mezzetti and Robertson (1999) estimated lung cancer incidence in France by developing Bayesian back-calculation models. Survival probabilities were considered to be known, based on preliminary analysis. The authors assumed the observed mortality counts to follow a weighted linear regression with a Normal distribution. The assumption of normality is not always justifiable, especially if the mean count is low. Furthermore, they smooth incidence over time by adopting an autoregressive normal prior distribution for the mean incidence parameter. This assumption might also be violated, especially in the case of analysis at a lower geographical level or a sparse disease, where the number of cases is low and a rather skewed distribution is expected.

In this study, we extend existing Bayesian back-calculation models by using a count distribution instead of a normal one to model observed mortality counts. Furthermore, we adopt a gamma autoregressive process for the incidence parameter. Models were applied to assess space-time patterns of lung cancer incidence from 1981–2009 in Switzerland. The incidence data from the SGA cancer registry have been used to estimate age- and gender-specific lung cancer survival rates. Model goodness of fit was assessed by comparing observed and model-based incidence of the SGA region. Furthermore, smooth maps of age standardized lung cancer incidence ratio (SIR) for both genders in 2005–2009 are presented at sub-district level.

5.2 Materials and methods

5.2.1 Data

Information on age- and gender-specific lung cancer incidence was provided from the cancer registry St. Gallen-Appenzell (SGA). The dataset was stratified by the year of incidence, municipality of residence as well as days to event and status at end of follow-up (dead, alive or lost to follow-up). Data covered the period from 1980 up to 2010. Some incidence records were reported with a follow-up time of zero days. These are captured via the death certificates and are known as *death certificates only* (DCO) (Parkin and Hakulinen, 1991). Our dataset covered 6% of DCO cases for males and 5% for females. These counts were excluded from the survival analysis, but used as death counts for the back-calculation model.

Mortality data for the SGA region were aggregated by year of death, age and gender. Countrywide

analysis at cantonal level considered lung cancer mortality provided from the FSO.
Population data from the Swiss censuses were used for spatial modelling SIR during the period 2005–2009. The unit of analysis was ms-region (ms = mobilité spatiale). Switzerland is divided in 106 such regions which are often used for scientific analyses (Federal Statistical Office, 2013c). Digital information files (shapefiles) at municipality level for the year 2009 have been obtained from the FSO and modified within the software ArcGIS to map the results at ms-region level.

5.2.2 Methods

Subgroup-specific differences in survival were assessed in the preliminary analysis. Age was categorized in two groups (45–64, 65–84) and the study period was separated in 5-year periods (1980–1984, 1985–1989, 1990–1994, 1995–1999, 2000–2004 and 2005–2010). Bayesian gender-specific survival models were adjusted for the two age categories and the time periods. Models were fitted in WinBUGS (Imperial College and MRC, London, UK) for each gender separately. We considered different survival distribution, such as exponential, Weibull and mixed following Mezzetti and Robertson (1999), to account for high lung cancer mortality in the first year, which was up to 70% in the SGA region. Based on model performance, a mixed distribution was considered assuming to be fixed for the first year and was modelled based on a Weibull distribution for the years 2–5.
A single model was used for the estimation of survival distributions and back-calculation formulations. Observed age- and gender-specific death counts x_t in year t were assumed to follow a Poisson or negative binomial distribution with parameter $\mu_{x,t}$.

$$x_t \sim Pois(\mu_{x,t}) \qquad \text{or} \qquad x_t \sim NegBin(p_{x,t}, r_x) \tag{5.2}$$

where $p_{x,t} = r_x/(r_x + \tilde{\mu}_{x,t})$. Mortality was linked to survival probabilities β_i via the following relation:

$$\mu_{x,t} = \sum_{i=1}^{5} \beta_i \mu_{I,t-i} \tag{5.3}$$

The survival function β_i for year i is estimated by the difference of the failure probability in year i and $i-1$ as follows:

$$\beta_i = exp(-\lambda(i-1)^a) - exp(-\lambda i^a) \tag{5.4}$$

where λ and a are the scale and shape parameters of the Weibull distribution defined by the survival function as follows:

$$S(i) = exp(-\lambda i^a) \tag{5.5}$$

A gamma prior distribution was considered for the shape parameter with shape parameter equal to zero and a scale parameter equal to 0.0001.

The autoregressive prior formulation of the incidence allows the estimation of a smooth curve over time. In addition, unstructured random effects were included in the regression formulation to capture temporal variation.

$$\mu_{I,t} = \mu_{I,t-1} * \epsilon_t \tag{5.6}$$

A gamma process was assumed for the random effects, to ensure estimation of positive incidence and to allow for skewness.

As lung cancer survival is known to be very low, the maximal time of survival was restricted to be 5 years following Mezzetti and Robertson (1999).

The model goodness of fit was assessed by comparing model-based estimates and observed incidence of the SGA region using the mean square errors. The model subsequently was used to obtain age- and gender-specific incidence estimates at cantonal level from mortality data extracted from the FSO database, assuming the same survival among all cantons.

Difference in cancer death records from certificates and those reported from cancer registries have been discussed in literature (Burnett and Dosemeci, 1994; Freedman et al., 2006). Miss- reported cause-of death on the death certificates sent to FSO are the main sources of this discrepancy. We calculated the proportion of under-reporting by comparing mortality data by FSO with those from SGA cancer registry. We then applied the back-calculation to relate lung cancer survival distributions with FSO mortality data, adjusted for the proportion of under-reporting.

The above model was employed to produce country-wide smooth maps of lung cancer SIR at the level of ms-region for the period 2005–2009. In particular, conditional autoregressive (CAR) models (Besag et al., 1991; Bernardinelli and Montomoli, 1992) were applied on incidence estimated from the back-calculation model using population data of 2010 to convert to SIR. This quantity relates the estimated incidence rate in a region to the one expected based on the national average. An SIR of 1.0 indicates same incidence as observed for the whole country. CAR models took into account spatial variation by introducing a random effect ϕ_j at ms-region level j, which was assumed to follow a conditional autoregressive prior distribution.

$$\phi_j \mid \boldsymbol{\phi}_{-j} \sim \mathcal{N}\left(\frac{\sum_{\substack{q=1 \\ q \neq j}}^{N} c_{jq}\phi_q}{w_j}, \frac{\sigma^2}{w_j}\right) \tag{5.7}$$

where $\boldsymbol{\phi}_{-j} = (\phi_1, \ldots, \phi_{j-1}, \phi_{j+1}, \ldots, \phi_N)$ and c_{jq} is the degree of spatial influence of ms-region j to the remaining ones. c_{jq} is equal to 1 if they are adjacent and 0 otherwise. The w_j indicates the number of neighbours of ms-region j. Final estimated rates were mapped within the software ArcGIS.

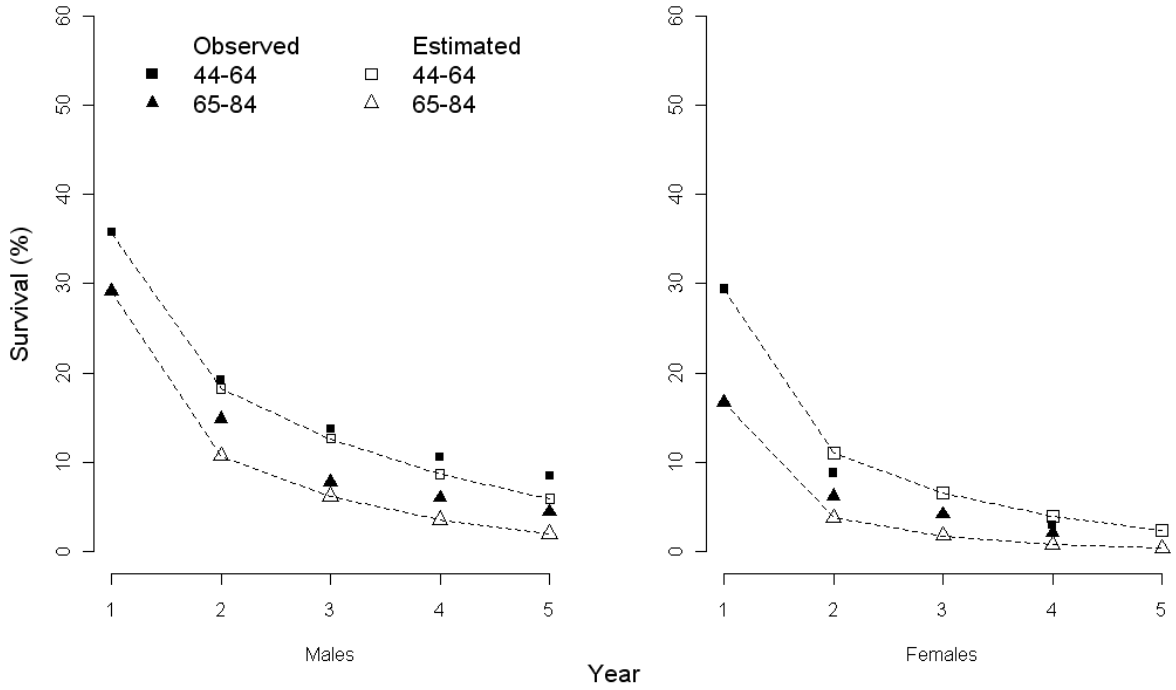


Figure 5.1: Observed and fitted age- and gender-specific survival in the SGA region in 1980–1985.

5.3 Results

Exploratory analysis using the log-rank test indicated significant differences between groups defined by age category, gender and period. Fitted and observed curves suggest a good fitting based on the mixed Weibull distribution for all periods (results shown only for the first period). Figure 5.1 indicates lower lung cancer survival of the older age group for both genders. The curves are dominated by the very low 1- and 2-year survival.

Figure 5.2 displays yearly observed and model-based gender-specific incidence for the SGA region during 1980–2010. Estimated incidence resembles the observed trends and indicates good model predictive performance. Lung cancer incidence for males remains stable within the range around 120–180 cases per year. Female lung cancer incidence data increased from around 20 cases per year in 1980 up to more than 80 counts in 2009.

Table 5.1 shows our model-based and reported (FSO) incidence estimates aggregated to German- and French/Italian-speaking region for the period 2005–2009. Estimates from the FSO are based on extrapolation of observed incidence to those cantons that are not covered by cancer registries. Incidence estimates based on mortality non-adjusted for the proportion of miss-specified cause of death are around 10–20% lower than the incidence reported by the FSO by language region.

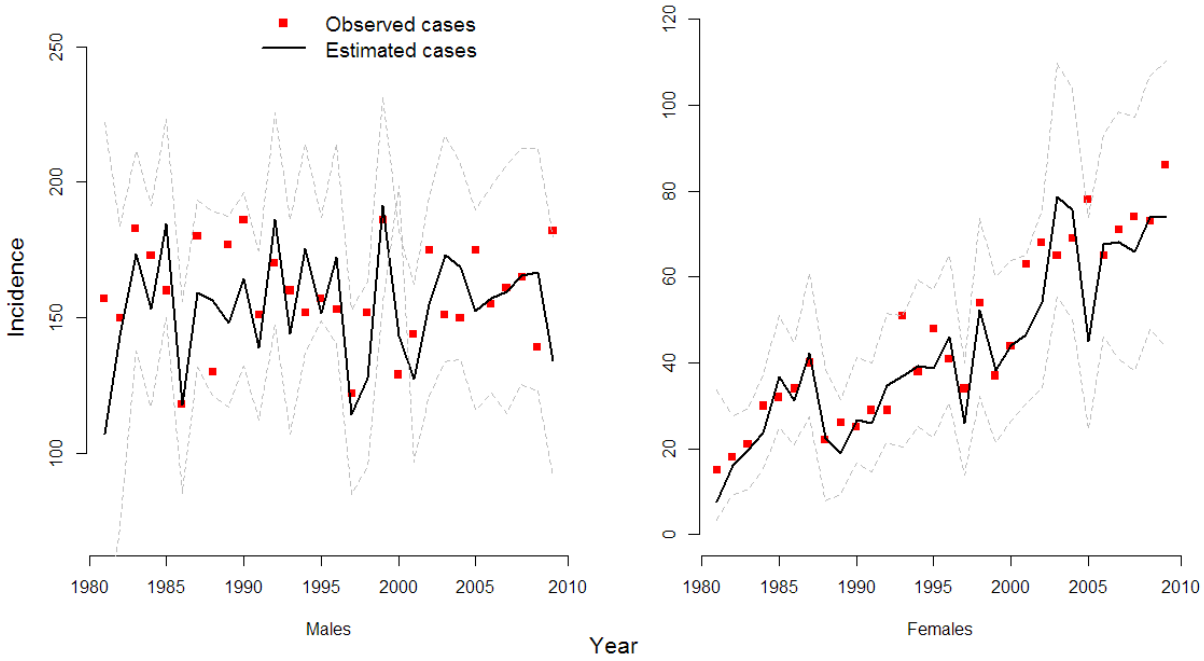


Figure 5.2: Observed and estimated incidence (with 95% Bayesian credible interval (BCI), dotted lines) for males (left) and females (right) in the SGA region from 1980–2010.

		FSO	Model-based	Model-based adjusted
German	M	1 666	1 503 (1 461;1 543)	1 648 (1 606;1 691)
	F	937	758 (725;790)	838 (802;873)
French/Italian	M	794	690 (663;720)	756 (723;783)
	F	443	372 (351;398)	413 (390;438)

Table 5.1: Lung cancer incidence estimates based on extrapolation reported by the FSO and our model-based median (with 95% BCI in brackets) for males (M) and females (F) in the German- and French/Italian-speaking in 2005–2009.

Estimates based on adjusted mortality indicated good predictive performance, with exception the estimates for females for the German-speaking part, which was estimated to be around 10% lower than the reported one based on extrapolation by the FSO. Evaluation of the model-based incidence estimates at cantonal level is limited due to missing data. Estimates published by the National Institute for Cancer Epidemiology and Registration (NICER) are compared with estimated incidence. Table 5.2 shows published lung cancer incidence and our model-based estimates (based on adjusted mortality) for each gender by canton.

GE		**1985–1989**	**1990–1994**	**1995–1999**	**2000–2004**	**2005–2009**
	Males	151	128	148	142	150
		148.4 (136.9;161.9)	*128.0 (117.9;138.6)*	*122.5 (111.9;134.3)*	*117.8 (107.6;128.8)*	*130.5 (118.9;141.9)*
	Females	53	56	67	78	86
		51.3 (44.3;57.7)	*52.9 (46.2;60.7)*	*54.3 (47.4;62.8)*	*70.5 (62.1;79.7)*	*96.4 (85.7;107.6)*
FR	Males					71
						92.0 (82.3;101.9)
	Females					34
						54.0 (46.4;62.5)
ZH		**1989–1992**	**1993–1996**	**1997–2000**	**2001–2003**	**2004–2005**
	Males	403	389	389	388	387
		418.3 (394.2;441.3)	*386.0 (363.8;410.0)*	*374.8 (351.5;398.7)*	*363.5 (338.0;390.5)*	*402.9 (366.0;444.6)*
	Females	120	138	173	196	218
		116.2 (105.0;128.1)	*129.0 (116.3;143.4)*	*143.0 (129.3;157.8)*	*146.8 (128.6;166.5)*	*196.8 (169.6;224.4)*
VS	Males	80	90	102		
		83.7 (73.4;94.5)	*90.9 (80.6;102.0)*	*97.7 (86.6;109.2)*		
	Females	18	26	37		
		17.3 (12.8;22.7)	*25.3 (20.0;31.8)*	*37.0 (30.8;44.3)*		
GR			**1989–1993**	**1994–1998**	**1999–2003**	**2004–2008**
	Males		63	65	64	72
			62.4 (55.3;70.1)	*60.7 (53.0;68.4)*	*58.5 (51.3;66.3)*	*69.1 (61.0;77.8)*
	Females		14	17	22	29
			16.2 (12.6;20.3)	*13.8 (10.8;17.7)*	*20.4 (16.3;25.4)*	*29.0 (23.5;35.3)*
TI	Males			75	129	135
				121.7 (110.6;132.9)	*119.9 (108.0;132.2)*	*120.0 (108.0;131.9)*
	Females			24	55	65
				37.7 (31.8;45.0)	*51.9 (44.9;59.9)*	*67.6 (59.2;76.3)*
GL	Males			12	17	14
				9.6 (6.9;12.8)	*15.9 (12.3;20.0)*	*14.2 (10.8;18.0)*
	Females			3	5	5
				2.9 (1.5;4.6)	*3.6 (2.1;5.7)*	*5.9 (3.9;8.3)*

Table 5.2: Published[1] (upper line) and model-based adjusted (italic in bottom line with 95% BCI in brackets) estimates of gender-specific lung cancer incidence for Switzerland, different cantons and yearly averages over different time periods.

Results confirm a good model predictive performance capturing well the temporal trends. The only exceptions however were estimates for Fribourg during 2005–2009 and Ticino during 1994–1998 which were higher than the actual number of reported cases.
Tables 5.3 and 5.4 show model-based incidence at cantonal level for the periods 1990–1994, 1995–1999, 2000–2004 and 2005–2009. Age group-specific estimates are presented in the Appendix.

[1]Estimates were retrieved from the corresponding publication of Statistics of Cancer Incidence from NICER (National Institute for Cancer Epidemiology and Registration & Federal Statistical Office, 2012; National Institute for Cancer Epidemiology and Registration, 2011a; Foundation National Institute for Cancer Epidemiology and Registration, 2012; Ceschi et al., 2009; Observatoire valaisan de la santé, 2005; National Institute for Cancer Epidemiology and Registration, 2011b,c,d)

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	413.7 (393.3;435.6)	409.1 (389.8;431.2)	379.6 (358.6;400.0)	362.9 (344.7;382.4)	371.8 (351.3;391.6)
BE	410.4 (391.4;431.3)	368.2 (350.9;389.6)	336.9 (319.8;354.8)	323.6 (305.1;342.4)	334.4 (315.9;353.3)
LU	119.0 (108.1;130.2)	125.3 (114.8;136.6)	101.1 (90.7;111.1)	105.4 (95.6;116.8)	103.9 (93.0;114.2)
UR	10.6 (7.8;14.0)	8.8 (6.1;12.0)	7.4 (5.1;10.3)	10.3 (7.4;13.7)	8.4 (6.0;11.5)
SZ	30.5 (25.2;36.5)	38.1 (32.3;44.8)	38.9 (32.5;45.5)	33.1 (27.5;39.6)	38.6 (32.8;45.3)
OW	7.5 (5.3;10.6)	7.9 (5.5;10.9)	6.6 (4.2;9.6)	5.5 (3.3;7.9)	7.9 (5.6;11.2)
NW	10.5 (7.9;13.9)	8.5 (6.1;11.6)	10.2 (7.3;13.7)	10.4 (7.3;13.7)	11.9 (8.8;15.3)
GL	10.7 (7.8;13.7)	17.0 (13.4;21.2)	10.4 (7.4;14.0)	18.1 (14.3;22.6)	11.2 (8.2;15.0)
ZG	17.1 (13.6;21.3)	15.5 (12.2;19.2)	24.3 (19.9;29.5)	23.6 (19.5;28.8)	24.2 (19.6;29.7)
FR	85.8 (77.4;95.4)	96.6 (86.6;106.8)	82.6 (74.2;93.0)	84.7 (76.0;94.9)	92.0 (82.3;101.9)
SO	106.1 (96.0;116.5)	101.2 (91.3;111.0)	84.2 (75.1;94.4)	97.3 (87.3;107.5)	85.9 (76.8;96.1)
BS	120.0 (110.0;130.5)	114.0 (102.3;125.4)	88.0 (78.7;98.4)	88.8 (79.7;97.7)	81.1 (72.0;90.1)
BL	95.1 (86.4;105.7)	83.6 (75.3;93.1)	75.0 (66.5;84.9)	90.0 (80.7;100.8)	86.4 (77.0;95.9)
SH	26.8 (21.6;32.6)	30.8 (25.6;36.6)	25.5 (20.8;31.5)	23.9 (19.3;28.9)	25.6 (21.0;31.0)
AR	13.3 (10.3;17.2)	20.1 (16.4;24.4)	17.0 (13.4;21.0)	13.5 (10.3;17.7)	9.8 (7.2;13.3)
AI	3.9 (2.3;5.8)	3.1 (1.7;4.9)	3.1 (1.8;5.0)	3.3 (1.9;5.2)	2.8 (1.5;4.7)
SG	137.8 (125.5;149.7)	138.7 (127.5;149.9)	122.9 (112.1;134.1)	137.9 (126.9;149.8)	138.0 (126.3;150.9)
GR	59.9 (52.2;67.8)	60.0 (52.8;67.8)	59.8 (52.1;68.1)	64.2 (56.2;72.1)	66.0 (58.0;75.1)
AG	179.4 (165.9;192.8)	173.1 (161.0;187.2)	167.9 (155.9;182.6)	165.5 (152.3;178.9)	162.3 (150.2;175.0)
TG	73.8 (66.2;82.2)	59.4 (51.4;66.6)	67.6 (58.7;76.0)	59.5 (51.6;67.4)	75.8 (67.1;84.1)
TI	124.9 (114.3;136.3)	120.2 (109.7;132.9)	121.4 (110.7;132.6)	119.7 (109.1;131.3)	124.7 (113.8;135.8)
VD	248.6 (114.3;136.3)	247.8 (232.3;262.9)	204.0 (189.2;218.6)	208.2 (193.8;223.5)	211.8 (197.6;225.4)
VS	78.9 (70.1;87.7)	87.6 (78.6;96.6)	88.3 (79.6;97.6)	102.3 (92.4;112.4)	105.4 (96.4;115.8)
NE	80.3 (72.0;88.7)	76.6 (68.6;85.5)	70.4 (62.4;79.7)	69.3 (61.4;78.6)	65.9 (58.2;74.0)
GE	148.4 (136.9;161.9)	128.0 (117.9;138.6)	122.5 (111.9;134.3)	117.8 (107.6;128.8)	130.5 (118.9;141.9)
JU	28.0 (23.6;33.3)	28.7 (23.9;34.5)	26.3 (21.3;31.6)	28.1 (23.2;33.6)	23.8 (19.1;29.4)

Table 5.3: Model-based median lung cancer incidence for males at cantonal level (with 95% BCI in brackets).

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	96.0 (86.3;105.5)	128.2 (116.9;139.6)	132.8 (121.7;145.6)	167.2 (154.0;181.6)	294.8 (276.8;315.2)
BE	59.9 (52.6;67.6)	76.6 (68.1;85.7)	84.2 (74.3;94.1)	117.6 (106.4;129.3)	199.6 (185.2;215.5)
LU	12.4 (9.1;16.4)	20.8 (16.7;25.5)	27.2 (22.8;33.0)	33.5 (27.8;40.0)	57.0 (49.3;65.1)
UR	1.3 (0.5;2.6)	0.7 (0.2;1.7)	0.6 (0.2;1.4)	2.0 (1.0;3.6)	4.2 (2.6;6.5)
SZ	4.5 (2.8;6.7)	9.1 (6.6;12.2)	9.2 (6.7;12.3)	16.0 (12.5;20.0)	25.1 (20.3;30.7)
OW	1.6 (0.8;2.9)	0.8 (0.3;1.7)	2.4 (1.3;4.0)	2.0 (0.9;3.7)	4.5 (2.8;6.9)
NW	1.0 (0.4;2.2)	2.1 (1.0;3.7)	3.1 (1.8;5.2)	3.5 (2.1;5.7)	9.8 (6.8;13.5)
GL	1.8 (0.8;3.2)	2.9 (1.6;4.7)	2.6 (1.4;4.3)	4.1 (2.6;6.5)	6.6 (4.5;9.3)
ZG	3.9 (2.3;6.1)	4.8 (3.1;7.0)	7.2 (5.0;10.2)	11.4 (8.6;15.4)	12.8 (9.5;16.6)
FR	9.8 (7.3;13.0)	14.8 (11.7;18.7)	22.6 (18.3;27.5)	26.5 (21.8;32.1)	54.0 (46.4;62.5)
SO	12.7 (9.7;16.1)	20.8 (16.1;25.4)	22.1 (17.8;27.4)	33.6 (27.8;39.7)	44.7 (37.3;52.4)
BS	31.3 (25.9;37.3)	40.6 (34.5;47.2)	42.4 (36.6;49.9)	46.2 (39.9;54.1)	66.3 (58.1;75.2)
BL	15.5 (12.1;19.9)	21.9 (18.0;26.8)	25.6 (20.8;30.6)	39.1 (33.5;45.6)	66.4 (57.5;75.8)
SH	3.7 (2.2;5.8)	4.7 (3.0;7.1)	9.1 (6.5;12.5)	9.7 (7.0;13.4)	16.1 (12.3;20.5)
AR	4.2 (2.6;6.2)	4.5 (2.8;6.7)	4.3 (2.7;6.5)	7.7 (5.5;10.7)	10.1 (7.2;13.6)
AI	0.5 (0.1;1.3)	0.7 (0.3;1.6)	2.0 (1.0;3.5)	0.8 (0.2;1.9)	1.7 (0.8;3.2)
SG	22.4 (17.9;27.0)	31.6 (26.3;37.2)	35.2 (29.5;41.5)	53.4 (46.0;60.4)	75.9 (67.7;85.2)
GR	13.1 (9.7;16.9)	14.9 (11.2;18.7)	13.4 (10.4;17.1)	23.2 (18.9;28.1)	38.2 (32.1;44.8)
AG	25.3 (20.7;30.8)	32.0 (27.0;38.1)	53.5 (46.3;61.5)	53.4 (45.9;61.0)	115.2 (104.2;126.6)
TG	10.4 (7.4;13.9)	13.5 (10.0;17.5)	15.3 (11.4;19.2)	22.5 (18.0;28.0)	37.5 (30.9;44.2)
TI	27.3 (22.4;32.4)	36.6 (31.0;42.3)	41.3 (34.9;48.5)	52.7 (45.4;60.6)	89.1 (78.7;99.5)
VD	53.1 (45.9;59.9)	72.8 (65.0;82.4)	83.9 (74.4;93.1)	108.7 (98.2;119.5)	159.2 (146.5;172.8)
VS	14.6 (11.3;18.9)	21.3 (16.9;26.2)	31.1 (26.0;37.5)	34.1 (28.4;40.1)	68.7 (60.4;78.2)
NE	14.6 (11.5;18.7)	22.5 (18.4;27.6)	25.5 (20.7;30.8)	27.8 (22.4;33.4)	55.1 (47.5;63.8)
GE	51.3 (44.3;57.7)	52.9 (46.2;60.7)	54.3 (47.4;62.8)	70.5 (62.1;79.7)	96.4 (85.7;107.6)
JU	2.5 (1.4;4.2)	5.6 (3.8;8.1)	5.6 (3.5;8.1)	8.0 (5.5;11.3)	19.0 (14.6;23.6)

Table 5.4: Model-based median lung cancer incidence for females at cantonal level (with 95% BCI in brackets).

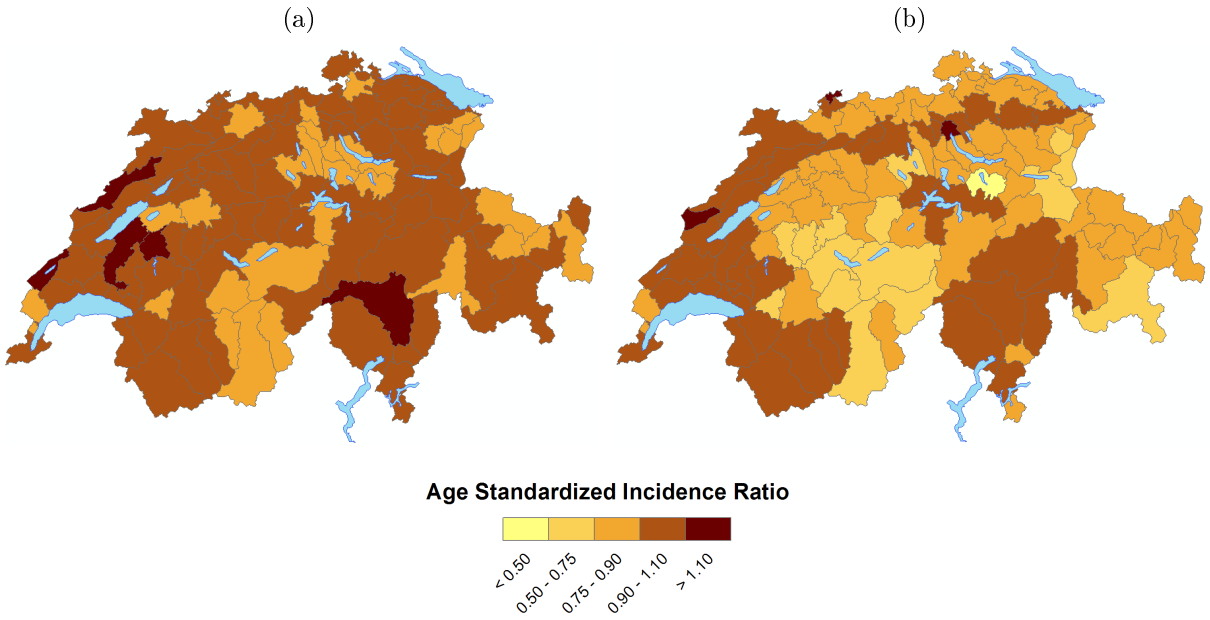


Figure 5.3: Model-based estimates of age standardized incidence ratio for males (a) and females (b) in Switzerland during 2005–2009.

Figure 5.3 shows estimated lung cancer SIR for both gender at ms-region level during the period 2005–2009. For males, a rather homogenous distribution was obtained. Highest incidence was shown in the French-(west) and Italian-(south) speaking part of Switzerland. Female lung cancer incidence estimates were higher in urbanized regions, i. e. Basel and Zurich.

5.4 Discussion

In this study, Bayesian back-calculation models were developed to estimate age- and gender-specific lung cancer incidence at different geographical levels. Observed mortality data was linked with estimated survival functions to approximate the incidence counts. Model goodness of fit was assessed by comparing observed with estimated incidence for the SGA region, an area administrated by a cancer registry covering the three cantons St. Gallen, Appenzell Innerrhoden and Appenzell Ausserroden.

Back-calculation models were applied to estimate age- and gender-specific incidence for all cantons from 1985–2009, assuming a constant survival across the country. Mortality data from the FSO were adjusted according to the missed proportion of lung cancer deaths in the FSO mortality database. Bayesian spatial count regression models were formulated on the incidence estimates to create smooth maps of SIR covering the whole country at ms-region level.

Comparison of the observed with the estimated incidence counts for the SGA region indicated a good predictive performance of the formulated back-calculation models. As shown in Figure 5.2,

variation of observed incidence for both, males and females, for the SGA region was covered by the estimated cases.

Preliminary analysis highlighted differences in survival by gender, age as well as calendar time. This finding coincides with the report of Bordoni et al. (2012) who analysed Swiss lung cancer survival in 1995–1999 and 2005–2009. The authors reported nationwide period- and gender-specific survival estimates, which correspond to our results. Therefore, assuming a constant survival for the whole country based on those estimated in the SGA area appears to be valid.

Mezzetti and Robertson (1999) followed a similar methodology to obtain model-based nationwide lung cancer incidence estimates in France. Data was aggregated by age at national level. They did account for age-specific survival distributions, but did not assess gender-differences. Results of our preliminary analysis indicated a significant difference in gender-specific lung cancer survival in Switzerland. Furthermore, the authors reported constant survival over time, which was not confirmed in our analysis. However, their analysis covered the period from 1950 up to 1990, while our study time was 1980–2010. The authors pointed out that individual-based survival analysis might improve the estimation and that uncertainty should be taken into account. In our study, survival time was assessed at individual level and a single model formulation, covering estimation of survival and incidence, allowed to taking into account the uncertainty driven by the estimation of the survival function within the back-calculation process.

The model formulation of Mezzetti and Robertson (1999) considered i) the Normal approximation of the Poisson distribution to model mortality counts, ii) known survival probabilities in the back-calculation estimated via maximum likelihood in a preliminary analysis and iii) positively restricted Normal prior distribution on the incidence parameter. We were not able to apply the above approach because our analysis was carried out at a lower geographical level, reporting small numbers of deaths. In our setting, Normal prior distribution on the incidence parameters was not appropriate because of the low number of incidence cases, resulting in a skewed distribution.

Colonna et al. (1999) considered the incidence-mortality ratio (IMR) to estimate incidence from mortality.This approach assumes a constant IMR, estimated from regions with both sources of data. In France, only 10% of the population is covered by cancer registries. The authors applied the above method to study the geographical trend and variation of breast and colorectal cancer. Validation of the estimates based on adjusted mortality data at national and cantonal level demonstrated good model performance. However, a few discrepancies have been obtained as for example for the canton Fribourg and Ticino during the last and first period, respectively.

Limitations of death certificates, such as incompleteness and lack of accuracy, are known and have been discussed (Burnett and Dosemeci, 1994; Freedman et al., 2006). In 2009, Bedford (2009) assessed the utility of death certificate data in predicting cancer incidence in Iowa. Lung cancer was concluded to be useful for this approach, as 75% or more of the incidence cases recorded at the cancer registry were captured by the underlying cause of death.

Our analysis assumed that i) the proportion of miss-specified lung cancer deaths in the FSO mortality database and ii) the survival distribution were the same for all cantons. Both have been estimated using data from the SGA cancer registry. Data from other cantonal cancer registries would allow us to assess the above assumptions and potentially improve the estimation of incidence for cantons with limited data.
In conclusion, the proposed back-calculation models improved earlier methodology and provided age- and gender-specific estimates of lung cancer incidence at cantonal level. Besides estimates for cantons without cancer registries, our results cover years for which cancer registries did not exist. In addition, smooth, gender-specific SIR maps highlight spatial patterns of the disease during 2005–2009.

Acknowledgements

This work was supported by Oncosuisse [grant number KLS-02393-02-2009] and the Swiss National Science Foundation [grant number 32003B_135769]. We would like to thank the Swiss Federal Statistical Office for providing the data.

5.5 Appendix

Cantonal abbreviations

Zurich (ZH), Bern (BE), Lucerne (LU), Uri (UR), Schwyz (SZ), Obwalden (OW), Nidwalden (NW), Glarus (GL), Zug (ZG), Fribourg (FR), Solothurn (SO), Basel-Stadt (BS), Basel-Landschaft (BL), Schaffhausen (SH), Appenzell Ausserrhoden (AR), Appenzell Innerrhoden (AI), St. Gallen (SG), Graubünden (GR), Aargau (AG), Thurgau (TG), Ticino (TI). Vaud (VD), Valais (VS), Neuchâtel (NE), Geneva (GE), Jura (JU).

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	146.9 (134.9;160.0)	144.0 (132.2;157.9)	136.9 (124.2;150.9)	122.0 (110.5;134.2)	118.6 (107.2;132.2)
BE	124.2 (112.1;136.0)	111.6 (100.4;122.6)	111 (99.8;122.9)	95.1 (84.6;106.5)	99.3 (89.2;109.9)
LU	44.0 (37.2;51.1)	39.2 (33.3;46.2)	32.0 (25.6;38.5)	30.0 (24.7;35.9)	29.3 (23.9;35.3)
UR	3.3 (1.8;5.5)	1.7 (0.8;3.4)	2.5 (1.3;4.5)	2.6 (1.3;4.5)	3.3 (1.8;5.3)
SZ	11.4 (8.4;15.1)	14.1 (10.6;18.2)	14.4 (11.0;18.7)	10.9 (7.8;14.7)	14.4 (11.0;18.6)
OW	3.2 (1.7;5.2)	2.5 (1.3;4.3)	2.5 (1.3;4.3)	2.2 (1.1;3.8)	1.9 (0.8;3.2)
NW	4.6 (2.7;7.1)	3.3 (1.8;5.4)	3.1 (1.7;5.3)	3.6 (2.0;5.7)	3.8 (2.2;6.1)
GL	3.2 (1.8;5.3)	5.9 (3.9;8.6)	2.4 (1.1;4.4)	4.8 (3.1;7.4)	3.8 (2.2;6.1)
ZG	6.7 (4.4;9.6)	5.4 (3.2;8.0)	8.3 (5.7;11.5)	8.7 (6.0;11.8)	8.9 (6.2;12.3)
FR	35.2 (29.7;41.7)	35.4 (29.3;42.3)	31.1 (25.4;37.1)	29.3 (23.4;35.3)	38.9 (32.8;45.8)
SO	36.8 (31.2;43.4)	32.8 (27.7;39.4)	21.8 (17.2;27.5)	31.7 (26.1;38.2)	27.9 (22.2;33.7)
BS	34.1 (28.2;40.8)	34.0 (28.2;40.5)	28.6 (22.9;34.4)	28.3 (23.1;34.8)	25.4 (20.6;30.7)
BL	38.4 (32.1;45.2)	27.4 (22.3;32.7)	27.0 (21.8;32.9)	28.8 (23.2;35.1)	28.6 (23.2;34.8)
SH	8.2 (5.4;11.4)	10.5 (7.6;14.0)	8.1 (5.6;11.2)	8.6 (5.8;11.9)	5.6 (3.5;8.4)
AR	4.7 (2.7;7.2)	4.6 (2.8;6.9)	4.2 (2.4;6.5)	4.9 (3.1;7.3)	2.8 (1.5;5.0)
AI	0.5 (0.1;1.5)	0.9 (0.3;1.9)	1.4 (0.6;3.1)	1.2 (0.4;2.6)	0.8 (0.2;2.0)
SG	49.7 (42.3;57.5)	52.4 (44.6;60.6)	44.4 (37.5;51.7)	49.2 (42.4;56.5)	42.9 (36.3;50.1)
GR	25.5 (20.6;31.0)	23.6 (18.8;29.0)	21.6 (17.2;26.8)	19.8 (15.5;24.8)	21.0 (16.4;26.2)
AG	66.2 (57.6;75.3)	60.2 (51.5;68.1)	60.3 (52.2;68.8)	58.1 (50.8;67.0)	51.8 (44.1;59.8)
TG	21.7 (16.9;26.7)	20.8 (16.6;25.7)	25.5 (20.7;31.6)	20.3 (15.9;25.3)	27.2 (22.0;33.0)
TI	47.6 (40.9;54.9)	45.1 (38.3;51.4)	37.3 (31.0;44.3)	40.1 (34.0;47.7)	31.3 (25.4;37.5)
VD	93.3 (83.9;104.2)	91.7 (82.7;102.0)	71.0 (62.1;81.0)	72.2 (62.6;81.6)	81.6 (73.0;91.6)
VS	31.4 (25.7;37.4)	32.6 (27.0;38.2)	29.7 (24.2;35.6)	33.6 (28.0;39.9)	38.5 (32.7;46.3)
NE	29.2 (24.3;34.7)	28.1 (23.1;34.4)	23.5 (18.6;29.5)	23.0 (18.3;28.6)	20.9 (16.3;26.1)
GE	55.2 (47.7;63.4)	48.5 (41.7;55.8)	47.6 (41.0;55.8)	40.7 (33.9;47.2)	45.8 (39.0;53.8)
JU	10.7 (7.5;14.2)	9.5 (6.8;13.0)	8.5 (6.0;11.6)	9.6 (6.5;13.1)	8.6 (5.8;11.8)

Table 5.5: Model-based median lung cancer incidence for males aged 45–64 at cantonal level (with 95% BCI in brackets).

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	258.4 (242.1;274.5)	252.9 (235.7;268.8)	222.5 (208.3;237.5)	223.1 (209.4;238.5)	232.7 (217.4;248.5)
BE	281.4 (264.5;298.9)	246.8 (231.3;263.4)	213.3 (198.6;229.1)	214.0 (200.3;229.8)	215.9 (201.5;231.3)
LU	73.5 (65.4;82.9)	84.4 (75.6;95.5)	66.9 (58.7;75.5)	69.9 (61.7;78.7)	70.9 (62.1;80.3)
UR	6.7 (4.4;9.6)	7.1 (4.8;10.0)	4.2 (2.5;6.4)	6.8 (4.4;9.6)	4.7 (2.8;7.2)
SZ	19.8 (15.5;24.5)	21.1 (16.8;26.1)	21.2 (16.9;25.7)	21.1 (17.0;26.0)	22.0 (17.4;27.0)
OW	4.2 (2.4;6.5)	4.9 (3.0;7.4)	4.5 (2.7;6.8)	3.0 (1.6;4.9)	5.7 (3.8;8.3)
NW	6.1 (4.2;8.8)	5.1 (3.2;7.5)	6.8 (4.5;9.6)	6.8 (4.5;9.6)	7.6 (5.3;10.6)
GL	7.5 (5.0;10.6)	9.7 (6.8;13.1)	7.8 (5.4;10.8)	11.8 (8.7;15.2)	7.3 (4.8;10.2)
ZG	9.9 (7.0;13.4)	9.8 (7.1;13.1)	15.2 (11.8;19.2)	15.4 (11.9;20.0)	14.3 (10.8;18.2)
FR	48.8 (42.0;56.2)	61.3 (53.7;69.2)	48.5 (41.9;56.2)	56.8 (49.8;65.1)	50.9 (44.0;58.1)
SO	67.5 (59.7;75.4)	66.8 (59.2;75.9)	57.3 (49.8;65.0)	60.9 (53.4;69.1)	54.4 (47.0;61.7)
BS	83.6 (74.7;93.0)	73.9 (65.6;83.1)	55.3 (47.6;63.5)	56.9 (49.4;64.9)	51.5 (44.8;58.9)
BL	55.3 (47.8;63.7)	56.1 (49.2;63.9)	47.3 (39.9;54.8)	59.5 (52.4;67.4)	56.5 (49.2;65.4)
SH	18.9 (14.9;23.4)	19.1 (14.8;23.6)	17.0 (13.2;21.8)	13.1 (9.8;17.2)	18.7 (14.6;23.4)
AR	8.7 (6.2;11.9)	15.7 (11.9;19.9)	11.1 (8.0;14.8)	8.1 (5.6;11.1)	6.8 (4.5;9.8)
AI	2.7 (1.5;4.6)	1.7 (0.8;3.3)	1.8 (0.9;3.2)	1.9 (0.9;3.4)	1.8 (0.8;3.3)
SG	88.6 (79.4;98.0)	85.2 (75.5;95)	72.8 (64.3;82.2)	86.4 (77.4;96.3)	91.4 (81.9;101.7)
GR	34.2 (28.7;40.8)	35.6 (29.8;42.0)	38.4 (32.3;44.6)	42.4 (35.7;49.2)	44.6 (38.3;52.2)
AG	110.4 (99.8;121.7)	110.0 (99.5;120.8)	101.9 (91.3;113.0)	99.6 (89.6;110.6)	101.7 (91.1;112.8)
TG	49.6 (42.7;56.9)	39.0 (32.5;46.1)	38.4 (32.2;45.2)	37.5 (31.3;44.4)	46.9 (40.1;53.9)
TI	77.2 (69.3;86.6)	73.8 (65.5;82.9)	80.6 (72.1;90.4)	73.8 (64.7;83.4)	84.5 (75.7;94.3)
VD	148.7 (135.6;161.8)	153.6 (141.4;166.8)	122.0 (110.4;132.9)	128.2 (117.1;141.1)	119.9 (109.1;131.0)
VS	46.9 (40.5;53.6)	54.2 (47.8;61.7)	56.7 (48.5;64.5)	65.5 (57.4;74.3)	67.8 (59.7;76.5)
NE	48.5 (42.3;56.6)	47.3 (39.9;54.7)	44.9 (38.5;51.5)	44.7 (38.0;51.5)	44.0 (37.6;51.1)
GE	85.3 (76.3;95.1)	71.1 (62.0;80.6)	71.5 (63.3;81.0)	72.9 (64.4;82.7)	77.9 (69.9;86.9)
JU	17.7 (13.6;22.3)	18.4 (14.4;23.0)	17.3 (13.3;21.8)	17.2 (13.6;21.8)	14.0 (10.5;18.3)

Table 5.6: Model-based median lung cancer incidence for males aged 65–84 at cantonal level (with 95% BCI in brackets).

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	35.6 (29.5;41.9)	51.1 (44.2;59.0)	49.0 (41.8;57.4)	67.4 (58.5;76.6)	80.2 (69.7;92.5)
BE	18.3 (14.1;22.9)	27.3 (22.0;32.9)	29.0 (23.2;35.1)	42.5 (35.7;49.7)	54.9 (46.8;64.6)
LU	5.7 (3.8;8.4)	7.0 (4.6;10.0)	9.6 (6.6;13.4)	13.4 (9.9;17.7)	16.6 (12.5;20.7)
UR	0.1 (0.01;0.5)	0.1 (0.01;0.7)	0.2 (0.01;0.7)	0.7 (0.2;1.6)	1.5 (0.6;3.0)
SZ	1.5 (0.5;2.7)	4.3 (2.6;6.7)	4.0 (2.2;6.4)	8.0 (5.6;11.0)	7.1 (4.8;10.4)
OW	0.8 (0.3;1.7)	0.3 (0.1;1.1)	1.4 (0.6;2.9)	0.9 (0.3;2.1)	1.5 (0.7;3.0)
NW	0.4 (0.1;1.3)	0.9 (0.3;2.1)	1.2 (0.4;2.5)	1.9 (0.9;3.3)	3.4 (1.9;5.7)
GL	0.3 (0.1;0.9)	1.6 (0.7;3.0)	1.4 (0.5;2.9)	2.3 (1.2;4.1)	2.8 (1.6;4.9)
ZG	0.8 (0.3;1.7)	1.3 (0.6;2.6)	2.4 (1.2;4.2)	6.6 (4.5;9.4)	3.7 (2.1;6.1)
FR	4.8 (3.0;7.1)	7.5 (5.2;10.6)	11.3 (8.2;15.2)	9.9 (6.8;13.8)	21.7 (16.9;27.1)
SO	4.1 (2.5;6.1)	6.8 (4.5;9.4)	7.9 (5.4;11.3)	11.9 (8.8;16.1)	15.3 (11.5;20.0)
BS	12.1 (9.0;15.9)	11.7 (8.5;15.6)	13.1 (9.6;17.1)	13.1 (9.5;16.9)	14.1 (10.4;18.6)
BL	7.1 (4.8;10.1)	9.0 (6.2;12.4)	13.3 (10.0;17.2)	14.0 (10.4;18.1)	21.7 (17.3;27.0)
SH	1.1 (0.4;2.4)	1.4 (0.6;2.8)	3.4 (1.8;5.6)	4.9 (2.9;7.5)	5.1 (2.9;7.8)
AR	0.9 (0.3;1.9)	1.2 (0.5;2.6)	2.9 (1.5;4.8)	4.0 (2.3;6.2)	2.1 (1.0;3.8)
AI	0.2 (0.01;0.7)	0.3 (0.01;0.9)	1.3 (0.5;2.5)	0.5 (0.1;1.4)	0.1 (0.01;0.4)
SG	7.1 (4.9;10.2)	13.9 (10.5;17.9)	13.8 (10.3;18.0)	20.8 (16.1;25.9)	21.6 (16.6;27.1)
GR	5.0 (3.1;7.5)	6.6 (4.3;9.7)	3.4 (1.9;5.6)	8.4 (5.9;11.7)	12.0 (8.8;16.1)
AG	8.3 (5.9;11.4)	15.7 (12.0;19.7)	26.6 (21.3;32.4)	23.8 (18.8;29.3)	36.6 (30.3;43.8)
TG	3.6 (2.2;5.6)	7.8 (5.4;11.0)	8.5 (5.8;12.1)	9.2 (6.5;12.8)	13.9 (10.1;18.0)
TI	9.0 (6.3;12.4)	10.9 (8.1;14.7)	14.7 (10.8;19.2)	17.3 (13.6;22.2)	15.9 (12.1;20.4)
VD	20.9 (16.7;25.5)	24.9 (20.1;30.2)	29.3 (24.2;35.6)	44.6 (37.8;51.8)	45.8 (38.6;53.1)
VS	5.6 (3.7;8.3)	10.6 (7.6;13.8)	11.9 (8.4;15.8)	13.5 (9.9;18.2)	20.8 (15.9;26.4)
NE	5.3 (3.6;7.8)	7.1 (4.8;9.9)	7.3 (4.7;10.2)	10.9 (7.9;14.7)	13.9 (10.2;18.2)
GE	12.9 (9.6;17.0)	14.4 (10.7;18.9)	19.5 (15.6;24.3)	23.9 (18.8;29.2)	21.3 (16.2;27.2)
JU	0.9 (0.4;2.0)	1.7 (0.8;3.4)	1.7 (0.7;3.2)	3.1 (1.7;5.1)	5.1 (3.1;7.9)

Table 5.7: Model-based lung cancer incidence for females aged 45–64 at cantonal level.

Canton	1985–1989	1990–1994	1995–1999	2000–2004	2005–2009
ZU	54.8 (48.3;62.1)	69.4 (61.7;78.1)	70.0 (62.3;78.7)	87.4 (79.1;96.7)	111.9 (102.1;122.2)
BE	35.0 (29.8;40.8)	41.5 (35.4;47.6)	52.7 (46.1;60.3)	60.4 (53.1;68.4)	73.2 (65.8;82.1)
LU	6.0 (3.8;8.4)	10.5 (7.8;13.7)	14.8 (11.4;18.4)	15.7 (12.1;19.9)	21.2 (16.9;26.1)
UR	0.9 (0.4;1.9)	0.4 (0.1;1.1)	0.4 (0.1;1.1)	1.0 (0.4;2.1)	1.2 (0.4;2.5)
SZ	2.3 (1.1;3.9)	3.0 (1.6;4.8)	3.9 (2.3;6.0)	6.2 (4.1;9.2)	9.1 (6.4;12.3)
OW	0.5 (0.1;1.3)	0.4 (0.1;1.2)	1.0 (0.3;2.2)	0.8 (0.3;1.9)	2.2 (1.0;3.8)
NW	0.5 (0.1;1.3)	1.0 (0.4;2.0)	1.1 (0.4;2.3)	1.5 (0.7;2.9)	2.0 (1.0;3.6)
GL	1.2 (0.5;2.4)	1.1 (0.4;2.4)	1.1 (0.4;2.2)	1.9 (0.9;3.5)	2.5 (1.4;4.3)
ZG	2.3 (1.2;3.9)	2.5 (1.4;4.3)	3.9 (2.4;6.1)	3.9 (2.3;6.3)	6.7 (4.6;9.3)
FR	3.6 (2.2;5.7)	5.5 (3.5;7.9)	7.0 (4.8;9.8)	12.7 (9.7;16.1)	15.0 (11.3;19.1)
SO	7.0 (4.9;10.1)	11.4 (8.5;14.8)	13.1 (9.8;16.8)	17.0 (13.3;21.1)	18.5 (14.6;23.0)
BS	17.0 (13.3;21.7)	25.2 (20.9;30.2)	26.5 (21.8;31.8)	26.4 (22.0;31.4)	30.9 (25.9;36.6)
BL	7.4 (5.1;10.2)	12.3 (9.4;16.0)	11.6 (8.5;15.0)	20.5 (16.4;25.3)	24.3 (19.8;29.3)
SH	1.7 (0.8;3.1)	3.0 (1.7;5.0)	4.2 (2.5;6.5)	4.1 (2.3;6.2)	5.8 (3.8;8.4)
AR	2.5 (1.4;4.1)	2.7 (1.5;4.4)	1.8 (0.9;3.4)	3.2 (1.7;5.1)	2.4 (1.3;4.2)
AI	0.4 (0.1;1)	0.5 (0.1;1.3)	0.5 (0.1;1.3)	0.2 (0.01;0.8)	0.9 (0.3;2.1)
SG	12.5 (9.3;16.3)	15.5 (12.0;19.7)	17.4 (13.9;22)	25.6 (21.1;30.5)	30.6 (25.3;36.5)
GR	6.7 (4.6;9.4)	7.2 (4.8;9.9)	6.9 (4.7;9.9)	11.5 (8.7;14.8)	15.4 (12.1;19.3)
AG	15.0 (11.5;19.0)	15.4 (11.9;19.6)	20.1 (15.8;24.7)	24.7 (20.2;29.4)	39.1 (33.1;45.4)
TG	6.2 (4.0;8.5)	4.7 (2.9;7.1)	5.9 (3.9;8.5)	10.9 (8.2;14.4)	15.6 (12.0;19.5)
TI	15.8 (12.5;19.9)	18.8 (14.8;22.9)	21.0 (16.9;25.5)	28.9 (24.0;34.6)	38.9 (33.4;45.3)
VD	24.0 (19.6;29.2)	39.7 (33.7;46.0)	47.2 (40.8;53.9)	52.9 (45.7;59.7)	59.8 (52.7;68.4)
VS	6.3 (4.4;8.8)	9.0 (6.4;11.9)	16.3 (12.9;20.5)	17.5 (13.9;21.7)	27.0 (22.1;32.2)
NE	7.7 (5.4;10.7)	11.7 (8.8;15.2)	15.8 (12.3;19.9)	14.7 (11.3;18.6)	20.5 (16.3;25.3)
GE	32.2 (27.3;37.8)	32.5 (27.1;38.4)	27.9 (23.2;33.2)	37.0 (31.2;43.8)	38.6 (33.2;45.1)
JU	1.4 (0.6;2.8)	2.5 (1.4;4.1)	3.1 (1.8;4.9)	3.9 (2.4;5.8)	7.8 (5.5;10.9)

Table 5.8: Model-based lung cancer incidence for females aged 65–84 at cantonal level.

Chapter 6

Indirect approximation of smoking patterns in space

Verena Jürgens[1,2], Silvia Ess[3], Marco Storni[4], Matthias Schwenkglenks[5], Thomas Cerny[6], Penelope Vounatsou[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland
[2] University of Basel, Basel, Switzerland
[3] Federal Statistical Office, Neuchâtel, Switzerland
[4] Cancer Registry of St. Gallen & Appenzell, St. Gallen, Switzerland
[5] Institute of Pharmaceutical Medicine (ECPM), University of Basel, Basel, Switzerland
[6] Medical Oncology-Hematology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland

This paper is prepared for submission to *Tobacco Control*.

Abstract

Background Smoking is the leading cause of lung cancer. Besides the consumption of tobacco, other (non-smoking) factors, such as radon and air pollution, have been associated with lung cancer mortality. Information on tobacco-use is sparse in Switzerland and existing survey data are not sufficient to assess space-time patterns of smoking behaviour. Knowledge about the distribution of tobacco use and the attributable number of deaths is essential in order to estimate its contribution to lung cancer burden.

The study aims to assess whether lung cancer mortality adjusted for non-smoking risk factors can be used as a proxy of smoking patterns in Switzerland.

Bayesian conditional autoregressive (CAR) logistic regression models have been applied on smoking data from a national survey of 1992 to estimate spatial patterns of gender-specific frequency of smokers. Negative binomial regression models with CAR spatial random effects (SREs) have been employed on lung cancer mortality data of 2008–2010 to obtain smoking proxies based on smooth mortality rates and SREs adjusted for radon and air pollution exposures (PM_{10} and NO_2). Association between the spatial patterns of smoking prevalence with the estimated smoking proxies was assessed graphically as well as by quantifying the magnitude of association via the Kendall's τ_b and logistic regression. Population attributable fractions due to smoking were estimated by age and gender to assess the burden of tobacco-use on lung cancer mortality in Switzerland.

There is a moderate correlation between observed smoking prevalence with smoking proxies. The correlation is stronger in females. This study showed that smooth lung cancer mortality rates, unadjusted for non-smoking risk factors, is a better proxy than the adjusted SREs.

In the absence of sufficient smoking survey data, the proposed indirect proxy can be used to assess spatio-temporal smoking patterns in Switzerland.

Keywords: Bayesian inference; smoking prevalence; lung cancer mortality; Switzerland

6.1 Introduction

Smoking is a well-known predominant risk factor of lung cancer. In the middle of the 20*th* century a causal relation was confirmed by the findings of case-control and cohort studies (Armero et al., 2012). Several subsequent studies assessed the influence of tobacco-consumption on several diseases such as various cancer sites, cardiovascular and heart diseases. In 1964, the Surgeon General's committee of the United States published its first report on the effects of smoking on health (United States Public Health Service, 1964), giving definitive evidence about the risk caused by smoking, followed by counterattacks from the tobacco industry. In the 20*th* century the number of deaths due to tobacco use worldwide was estimated to be around 100 million and predicted to increase up to 1 billion deaths during the following century (WHO, 2008). Besides direct tobacco-consumption, passive smoking has been confirmed to be a risk factor of the disease. First investigations on this relation have been considered in the early '80s (Trichopoulos et al., 1981).
The leading non-smoking, environmental lung cancer risk factor is exposure to radon (Federal Office of Public Health, 2012b), which is a radioactive gas, occurring in nature. In Europe, around 9% of lung cancer mortality and 2% of all cancer deaths are caused by radon exposure in the home. The risk for smokers or recent ex-smokers has been estimated to be notably greater than for lifetime non-smokers (Darby et al., 2005). The spatial distribution of radon depends on the geology of the country, as it can be mainly found in rocks and soil. The level of exposure is related to the source of material, where the gas is released. The Swiss Federal Office of Public Health has carried out a survey on indoor radon measurements over the last 20 years. The study covered more than 70 000 units and aimed to identify houses with high indoor radon concentrations. In Switzerland the average indoor radon concentration is 78 Bq/m^3. Based on these figures, Menzler et al. (2008) determined the population attributable fraction (PAF) for lung cancer due to residential radon. They reported a PAF of 8.3% for the Swiss population. One hundred and sixty-nine male and 62 female lung cancer deaths per year can be attributed to residential radon. Overall, more than 90% of these cases are in smokers. The authors emphasized the large spatial variation in the PAF, determined by cantonal discrepancies of radon concentrations and population structure.
Another risk factor for lung cancer is air pollution. While in the United States evidence of positive association of air pollution and lung cancer based on cohort studies has already been published in 1993 (Dockery et al., 1993), results of European long-term exposure cohort studies have been reported in 2000. The authors of a Dutch study concluded air pollution to be a potential risk factor (Hoek et al., 2002). Results of a Norwegian study in 2003 indicated a significant increase risk of lung cancer in NO_x exposure (Nafstad et al., 2003).
Nevertheless, smoking is the main risk factor of cancer related to the lung. Identifying high risk regions and analysing site-specific risk factors to improve intervention planning are the elementary steps for the reduction of tobacco-related diseases.

Data on smoking patterns in Switzerland are sparse. Only few surveys have been carried out to collect data on smoking behaviour among the Swiss population. There is no study about trends of tobacco-consumption covering the whole country. Smoking survey data are available for the last two decades, but they cannot be used for a long-term analysis, due to the time lag between consumption/exposure to tobacco and the outbreak of lung cancer which is 20–30 years.
Tobacco Monitoring Switzerland, a survey of tobacco consumption among those aged 14–65, was initiated at the University of Zurich in 2001. These are telephone surveys, collecting smoking data four times a year and covering around 2 500 participants each. The data however cannot provide useful information for assessing the effect of smoking behaviour 20–30 years ago on current lung cancer mortality data. Another source is the Swiss Health Survey conducted every five years since 1992, providing different information on smoking, e. g. number of cigarettes per day, duration of smoking in years and the number of smokers out of those surveyed.
Patterns of tobacco consumption, the leading cause of lung cancer mortality, are driven by numerous indicators. One factor is the cultural behaviour of a population. In Switzerland, a possible cultural classification is the language region. The country is splitted into four linguistic areas – following the order of the cardinal directions starting in the northern part, one can describe the partition as German, Romansh, Italian and French. However, the German and French parts cover the main part of the Swiss population. Another characteristic is the level of urbanization of a region, which is driven by the population density.
In this study, we studied spatial patterns of gender-specific frequency of smokers in Switzerland. Moreover, we assessed whether lung cancer mortality unadjusted and adjusted for non-smoking risk factors can be used as a proxy of smoking patterns in Switzerland. Population attributable fractions due to smoking were estimated by age and gender to quantify the burden of tobacco-use on lung cancer mortality in the country.

6.2 Materials and methods

6.2.1 Data

Mortality data at municipality level were acquired from the Swiss Federal Statistical Office (FSO). Data were aggregated by ms-region (ms = mobilité spatiale) for the three most recent available years (2008–2010). This geographical unit divides Switzerland into 106 regions which are often used for scientific analyses (Federal Statistical Office, 2013c). Population data of the census 2010 were retrieved from the FSO. Digital maps of municipalities were provided by the FSO for the year 2009. They were modified within the software ArcGIS to obtain the boarder structure of the ms-regions. The population density of each ms-region was calculated and used to classify

it as urban or rural, following the definition of the OECD[1] (Schrader, 1997), i. e. an ms-region was classified as rural if the density was less than 150 people per square kilometre. Language information at the ms-region level was obtained from the FSO.
Data on radon, categorising the exposure risk in four groups, were provided from the Swiss Federal Office of Public Health. The data are available during 2000–2011, however data after 2008 were ensured to be error-free and used for this analysis. Yearly gridded surfaces of Particulate Matter of 10 (PM_{10}) and Nitrogen Dioxide (NO_2) at 200 m by 200 m spatial resolution were provided by Meteotest and used as indicators of air pollution. The data are available since 1990 for NO_2 and 1998 for PM_{10} and allow for a 20–30 years time lag between the exposure and the disease. Data on radon, PM_{10} and NO_2 were aggregated at ms-region level by their mean exposure.
Data on smoking behaviour were provided from the Swiss Health Survey conducted in 1992. This is a nationwide population-based telephone survey carried out by the FSO and covered 14 521 participants. Several information were recorded such as the frequency of smokers (number of smokers and surveyed) and frequency of the number of cigarettes and years smoked (both by categories and summarized by median). However, only the number of smokers and those surveyed has been considered for this analysis based on preliminary analysis.

6.3 Methods

Estimating spatial patterns of smoking prevalence

Bayesian spatial logistic regression models were applied to analyse gender-specific smoking prevalence p_i $(i = 1, \ldots, m)$ in 1992, adjusted for the covariates $\boldsymbol{X_i}$ to assess differences of smoking behaviour by language region and urbanization level.

$$logit(p_i) = \alpha + \boldsymbol{X_i^T}\boldsymbol{\beta} + \phi_i \tag{6.1}$$

where α is the intercept and $\boldsymbol{\beta}$ is the vector of the regression coefficients. Spatial random effects (SREs) ϕ_i were included in the model to account for spatial correlation at ms-region level, which were assumed to follow a Conditional Autoregressive Gaussian (CAR) distribution (Besag et al., 1991; Bernardinelli and Montomoli, 1992) as follows:

$$\phi_i \mid \boldsymbol{\phi}_{-i} \sim \mathcal{N}\left(\frac{\sum_{\substack{q=1 \\ q \neq i}}^{N} c_{iq}\phi_q}{w_i}, \frac{\sigma^2}{w_i}\right) \tag{6.2}$$

where $\boldsymbol{\phi}_{-i} = (\phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_N)$. The degree of spatial dependency of ms-region i to the remaining ones is indicated by c_{iq}, taking the value 1 if they are adjacent and 0 otherwise. w_i

[1] Organisation for economic co-operation and development

represents the number of neighbours of ms-region i. Following a Bayesian model specification, prior distributions were assigned to the unknown parameters. Vague normal prior distributions were considered for the coefficients.

Estimating proxies of spatial smoking patterns based on lung cancer mortality

Bayesian negative binomial regression models with spatial CAR random effects were fitted to model gender-specific lung cancer mortality unadjusted and adjusted for non-smoking risk factors (exposure to radon, PM_{10} and NO_2). The spatial distribution of radon was assumed to be constant over time and therefore data of 2008 were considered representative of radon exposure over the last decades. This appears to be a valid assumption, as its occurrence is in rocks and soil where it is stored for long periods. In addition, random effects ϕ_i were included in the model to account for spatial correlation at ms-region level. Covariate effects (radon, $PM_{10}$and NO_2) $\boldsymbol{\beta}$ were modelled on the log mean of lung cancer death counts μ_i, using the gender-specific population n_i of the corresponding ms-region i as the offset.

$$log(\mu_i) = log(n_i) + \alpha + \boldsymbol{X_i^T}\boldsymbol{\beta} + \phi_i \tag{6.3}$$

We assessed whether lung cancer mortality unadjusted for non-smoking risk factors fit the observed smoking behaviour data. Furthermore, we investigated the hypothesis that the SRE from the adjusted regression model can serve as an indirect proxy for smoking patterns. The SRE captures spatial patterns unexplained by the risk factors implemented as covariates in the regression model(Law and Ping, 2011), therefore, the spatial distribution of the SRE is expected to be similar to the one of the smoking-related factors which were not included in the regression model.

Validating smoking proxies

The goodness of fit between the mortality rates and SREs adjusted for non-smoking risk factors with the smoking prevalence was determined: (i) graphically, by comparing the smooth maps of model-based estimates of smoking prevalence with estimates of adjusted lung cancer mortality rates and SREs; (ii) by estimating Kendall's τ_b correlation coefficients and applying a logistic regression model on the smoking data from the Swiss Health Survey to assess the magnitude of association between the different proxies and the smoking prevalence.

Estimating PAF of tobacco use on lung cancer mortality

The number of lung cancer cases due to smoking was quantified following the approach of Parkin (2011). The author estimated the tobacco-attributable cancer burden in the UK in 2010. Regarding lung cancer, estimation of incidence in never-smokers was based on the American Cancer

Society's second Cancer Prevention Study (CPS II, (Thun et al., 2006)) and the attributable fraction of lung cancer was given by

$$\text{PAF} = (O - E)/O \tag{6.4}$$

where O is the number of overall lung cancer deaths and E is the expected number of lung cancer deaths in never-smokers. This approach is referring to the method developed by Peto et al. (1992) and assumes that the lung cancer deaths exceeding the number expected in never-smokers, is due to smoking. Age- and gender-specific lung cancer rates in never-smokers were applied to the Swiss population under consideration to calculate those lung cancer cases expected in a non-smoking population. The number of lung cancer cases due to smoking was directly estimated by subtracting the expected cases from those observed in Switzerland in 2010. High PAFs would support the hypothesis that lung cancer mortality can be used as a proxy for smoking.
Preliminary analyses were carried out in Stata/IC 12 (Stata Corporation; College Station TX, USA) and Bayesian models were fitted in using WinBUGS (Imperial College and MRC, London, UK).

6.4 Results

Estimating spatial patterns of smoker frequency

Logistic regression estimates on gender-specific frequency of smokers are provided in Table 6.1. The odds of smoking were significantly higher in the French-speaking region compared to the German-speaking one for both genders (higher by 24% for males and 39% females). No significant differences were observed in the odds of smoking between the Italian- and Romansh-speaking compared to the German-speaking one. Results indicated a significant increase of 27% for female smoking odds in urbanized ms-regions. Figure 6.1 illustrates the spatial distribution of the environmental covariates, i. e. radon risk, PM_{10} and NO_2 (annual means). Higher risk of radon exposure was found in the south-eastern and western part of the country, defined by the Alps and the Jura Mountains. Spatial patterns of PM_{10} and NO_2 are rather similar, characterized by hotspots in the urbanized regions Basel, Zurich, Geneva and Lugano. In particular, the Swiss midland is dominated by high levels of PM_{10}, due to the main roads in this region, as road traffic is one of the major sources of particulates.

Figure 6.2 shows model-based smooth gender-specific smoking prevalence in 1992 adjusted for language region and urbanization level. The maps show higher male smoking prevalence in the French- and Italian-speaking regions. For females, the map is dominated by high estimates in the urban regions as well the French-speaking part. For both genders, the estimated smoking prevalence did not exceed 42%.

	OR (95% BCI)	
	Males	*Females*
Language		
German	1.00	1.00
French	1.24 (1.05;1.47)	1.39 (1.17;1.65)
Italian	1.26 (0.97;1.67)	1.18 (0.88;1.57)
Romansh	1.12 (0.63;1.96)	1.35 (0.73;2.28)
Urbanization level		
Rural	1.00	1.00
Urban	1.08 (0.92;1.27)	1.27 (1.08;1.49)

Table 6.1: Logistic regression parameter estimates for the gender-specific frequency of smokers in Switzerland in 1992 (with 95% Bayesian credible interval (BCI)).

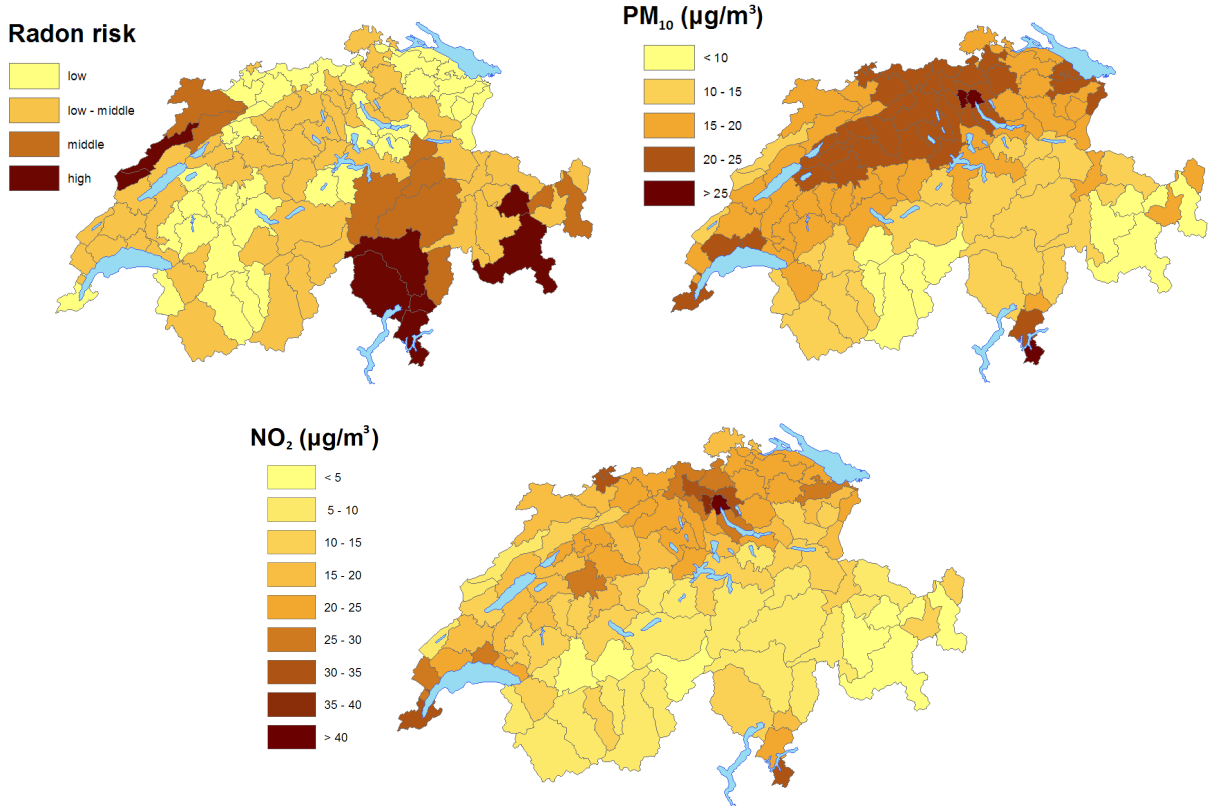


Figure 6.1: Spatial distribution of the environmental risk factors radon (in 2008), PM_{10} (in 1998) and NO_2 (in 1990).

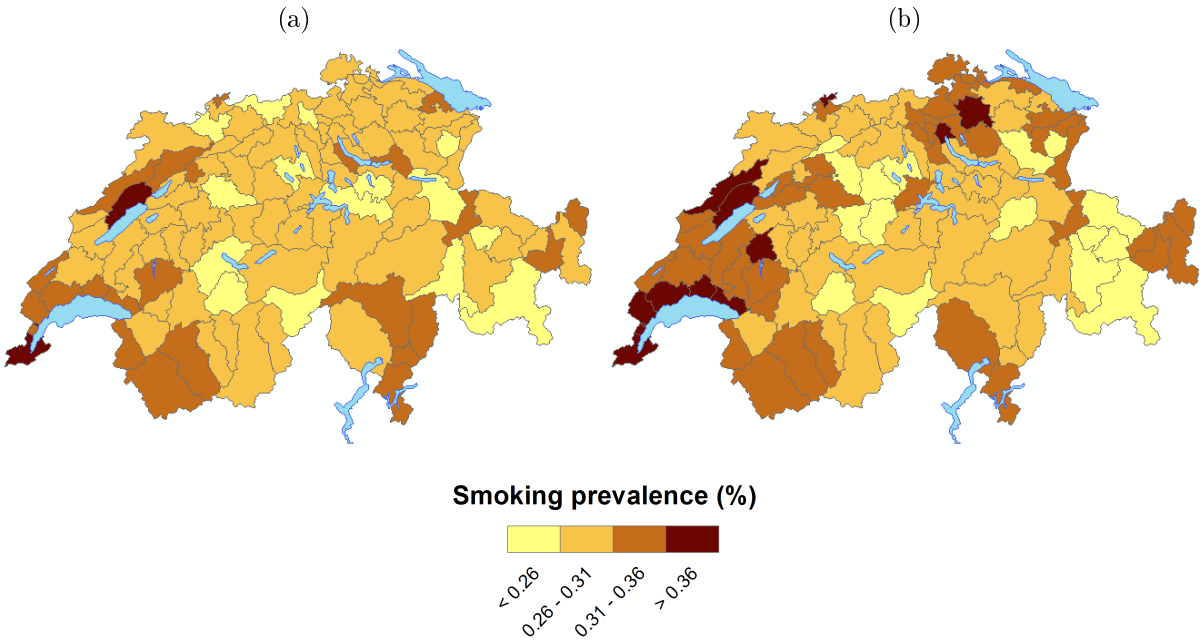


Figure 6.2: Model-based smooth smoking prevalence (%) for males (a) and females (b) in Switzerland in 1992.

Validating smoking proxies

Gender-specific maps of unadjusted smooth mortality rates and SREs (derived from the adjusted negative binomial regression) are given in Figures 6.3 and 6.4. For validation purposes, the spatial patterns of gender-specific smoking prevalence are presented and estimates are standardized (by subtracting the mean and dividing by the standard deviation) to achieve a common scale for all maps.

Resulting maps for males indicate a good approximation of smoking prevalence by the SRE and the unadjusted mortality rate (see Fig. 6.3). Clustered regions with high frequency of smokers are captured by both proxies, however, the mortality rate overestimated smoking frequency in the alpine region. On the other hand, maps of the SRE indicate an overestimation around Zurich and underestimation in the south-western part of Switzerland. Logistic regression results provided in Table 6.2 confirm a better association of the smoking prevalence with the unadjusted smooth mortality rates than the SREs, however, estimated OR are important for both proxies. Kendall's τ_b indicates a rather moderate association (0.43 and 0.55 for males and females, respectively).

Similarly to males, maps of indirect smoking proxies for females suggest that the unadjusted smooth mortality rate is a better proxy (see Fig. 6.4). Logistic regression estimates provide a higher OR for the mortality rates, compared to the SREs, while the Kendall's τ_b indicate a

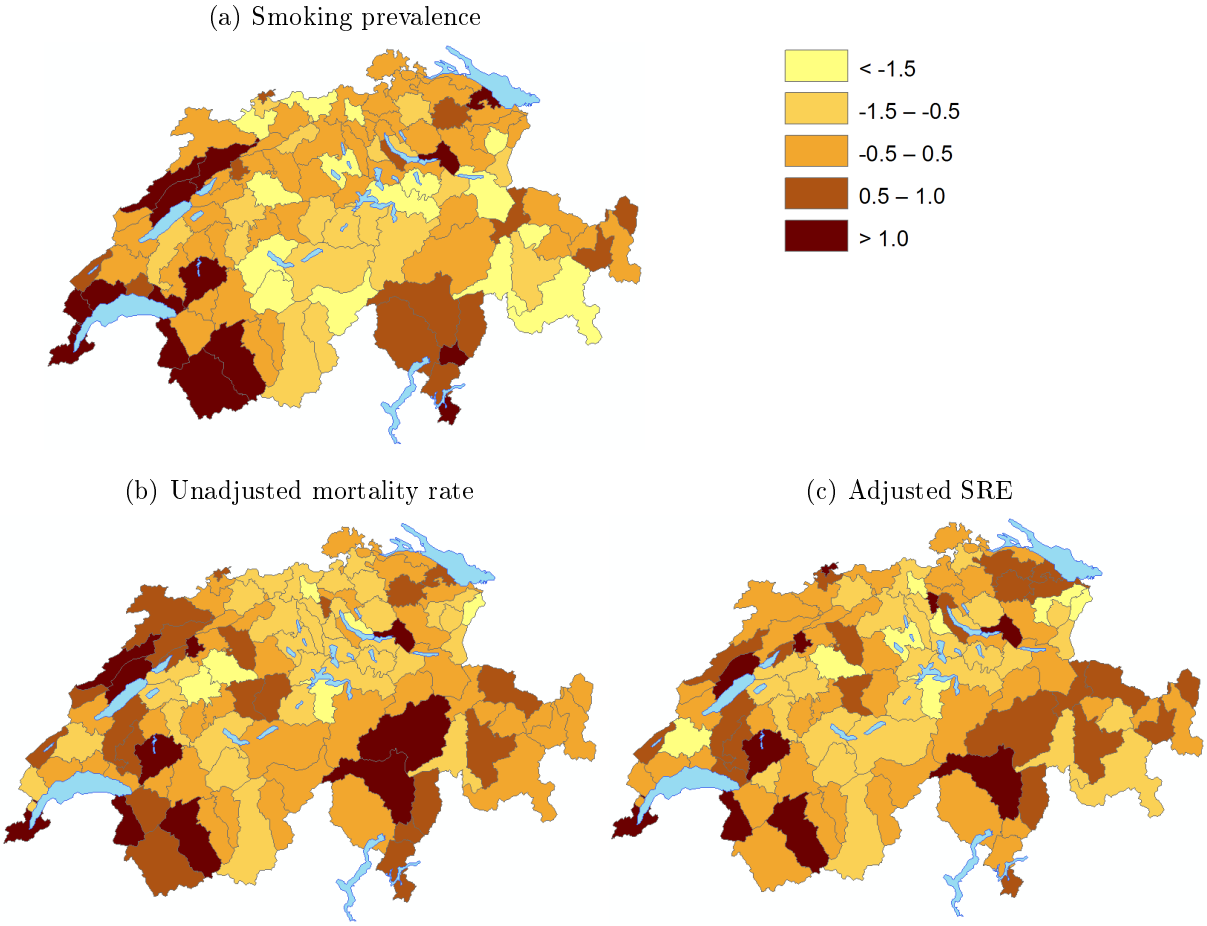


Figure 6.3: Model-based estimates of smoking prevalence, lung cancer mortality rates and SREs for males.

moderate positive correlation for both proxies.

Estimating PAF of tobacco use on lung cancer mortality

Excess lung cancer mortality attributable to smoking by age group together with observed and expected cases is presented in Table 6.3. For males, the highest PAF (88%) was observed for those aged 55–69, while a peak of 80% was obtained for females aged 55–59. The overall PAF was 85% and 69% for males and females, respectively. The estimates are based on the assumption that the lung cancer mortality rates among non-smokers in Switzerland in 2010 are the same as those of the US CPS II(Thun et al., 1997) in the '80s.

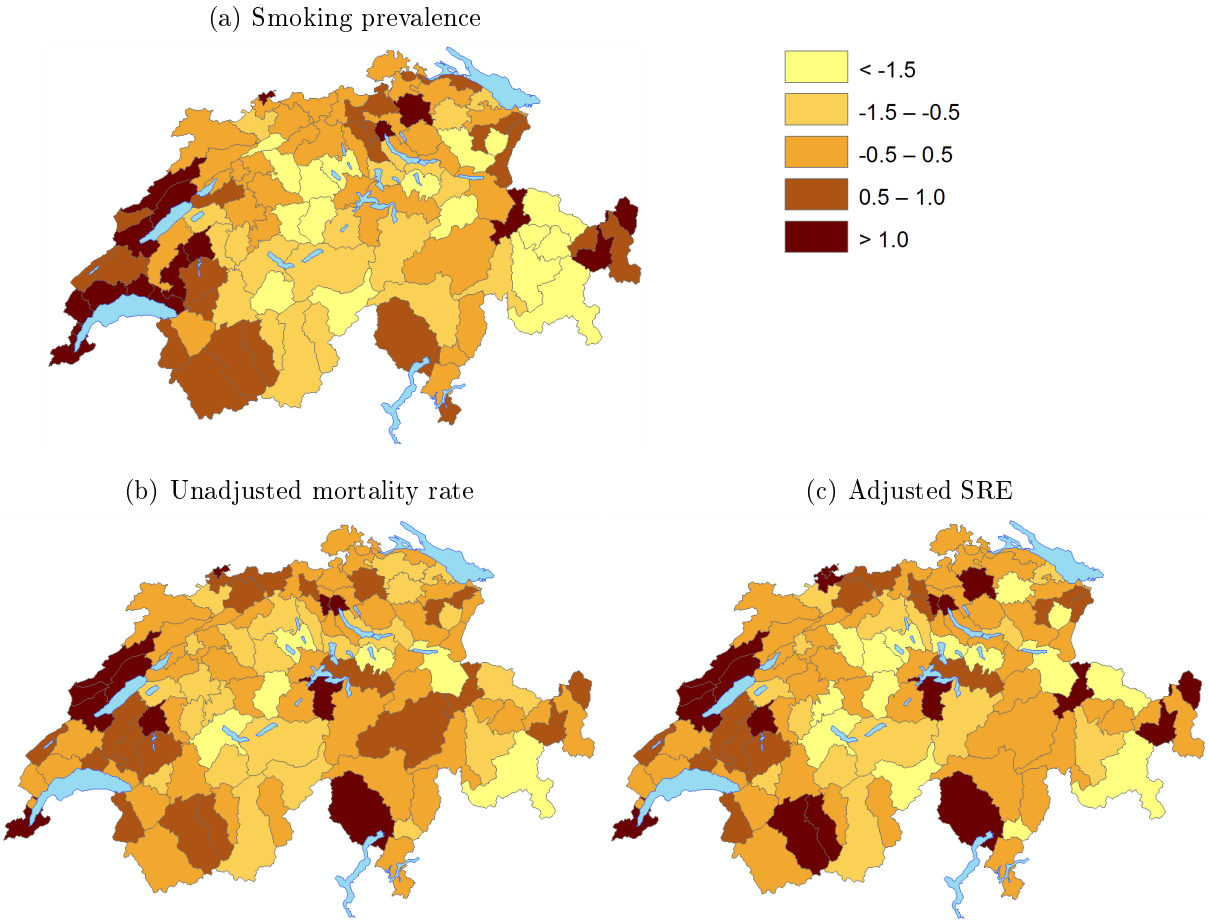


Figure 6.4: Model-based estimates of smoking prevalence, lung cancer mortality rates and SREs for females.

	OR (95% BCI)		DIC		τ_b	
Mortality rate (unadj)	2.8 (1.4;6.6)	16.6 (5.9;74.9)	1362.4	1248.5	0.43*	0.55*
SRE	2.1 (1.4;3.7)	2.8 (2.0;4.6)	1352.3	1248.1	0.45*	0.56*

Table 6.2: Logistic regression parameter estimates, DIC and Kendall's τ_b for adjusted frequency of smokers in Switzerland in 1992 for males and females (grey).
*p<0.001

Age (years)	Observed cases		Expected cases		Excess attributable cases		PAF (%)	
35–39	13	7	13	17	0	0	0	0
40–44	59	45	22	28	37	18	63	39
45–49	140	102	34	40	105	63	75	61
50–54	288	180	47	50	241	131	84	73
55–59	496	314	62	63	434	251	88	80
60–64	765	420	90	87	675	333	88	79
65–69	1 017	460	118	110	899	350	88	76
70–74	970	431	132	127	837	304	86	71
75–79	1 031	469	163	163	868	307	84	65
80–84	791	378	171	190	620	188	78	50
Total	5 564	2 812	854	872	4 716	1 943	85	69

Table 6.3: Age-specific observed/expected lung cancer deaths and its population attributable cases/fraction due to smoking in 2008–2010 for males and females (grey).

6.5 Discussion

In this work, we aimed to assess whether lung cancer mortality rates unadjusted and adjusted for environmental risk factors (i. e. radon and air pollution) can be used to indirectly approximate gender-specific smoking patterns in space. Two proxies were estimated from negative binomial models, a smooth mortality rate unadjusted for non-smoking risk factors and regression-adjusted SREs. Moreover, we applied Bayesian spatial logistic regression models on survey data and provided nationwide model-based maps of smoking prevalence in 1992 for males and females. To quantify the burden of tobacco use on lung cancer, population attributable fractions due to smoking were estimated during the period 2008–2010.

Based on graphical and numerical validation of the two smoking proxies, we obtained moderate association between smoking prevalence with mortality rate and SRE, which was stronger for females. Mortality rates performed better in approximating smoking patterns than SREs. Residuals of the gender-specific lung cancer adjusted mortality have been also estimated and tested regarding their ability to approximate smoking prevalence patterns. Results indicated no association (Kendall's τ_b of 0.16 and 0.15 for males and females, respectively).

Chellini et al. (2006) used female lung cancer mortality trends in central Italy from 1987–2002 as an indicator of female smoking trends. They reported higher female lung cancer mortality in urbanized regions, which is consistent with findings in Switzerland. The authors estimated yearly per cent change in overall age-standardized lung cancer mortality rates of 24.5% in the urban and 17.2% in the rural Italian areas. Their findings also suggested that female smoking prevalence increased during the last decades.

Frequency of smoking by language region and urbanization level coincides with findings of gender-specific lung cancer mortality, which was found to be higher in urban regions for females. For males a rather homogeneous distribution of the disease was seen.

The smoking data used in this study were provided from the Swiss Health Survey of 1992. These data cover different information on smoking, e. g. number of cigarettes per day, duration of smoking in years and the number of smokers out of those surveyed. Preliminary analysis of the data indicated only association of the latter with lung cancer mortality, which was then used for the analysis described in this work.

The geographical coverage of the Swiss Health Survey of 1992 is incomplete. This might explain some associations, which were expected to be stronger, as for example the moderate correlation between lung cancer mortality rate and smoking prevalence.

The estimation of the PAF is based on the assumption that the lung cancer mortality rates among non-smokers in Switzerland in 2010 are the same as those of the US CPS II (Thun et al., 1997) in the '80s. This assumption is based on previous studies showing that lung cancer mortality in never smokers did not change over time (Thun et al., 2008; Samet et al., 2009) and that the figures observed within the US CPS II study are representative of those in other developed countries (Peto et al., 1992).

Availability of smoking data in the country is rather limited. Data on non-smoking factors are also sparse for early years. For radon, information of good quality was obtained for the year 2008. We assumed that the spatial distribution and exposure to radon is constant over the last decades and we related mortality in 2008–2010 with radon exposure in the same year. Countrywide surfaces for air pollution indicators, such as NO_2, are already available yearly from 1990 onwards. Therefore, a time gap of around 20 years can be considered between mortality and exposure. However, lack of data at high-resolution introduces limitation to the study. For example information on PM_{10} exists only since 1998.

Smoking ranks the first cause of lung cancer followed by radon exposure (Federal Office of Public Health, 2012b). The lung cancer risk associated with radon exposure is higher for smokers than for non-smokers (Federal Office of Public Health, 2011). However, preliminary analysis did not indicate any interaction between radon and smoking prevalence, which might be due to the incompleteness of the smoking data.

In this study, we developed and assessed different smoking proxies with high spatial and temporal resolution. Our approach enables to indirectly estimate smoking patterns from observed lung cancer mortality. These estimates allow studying gender- and age-specific geographical patterns and time trends of smoking at a sub-regional scale. In addition, they can be used to assess the contribution of smoking on the burden of other cancer types or diseases such as cardiovascular or respiratory ones.

Acknowledgements

This work was supported by Oncosuisse [grant number KLS-02393-02-2009] and the Swiss National Science Foundation [grant number 32003B_135769]. We would like to thank the Swiss Federal Statistical Office for providing the data.

Chapter 7

Discussion

This PhD thesis contributes to the field of tobacco-related cancer epidemiology with statistical methodology and epidemiological knowledge on spatio-temporal patterns of cancer mortality and incidence in Switzerland. The main contributions of the research are (i) improved statistical models for projecting cancer mortality. These models allow forecasting of sparse count data; (ii) improved statistical models for estimating tobacco-related incidence from mortality data. These models enable estimation of incidence rates of the disease in the parts of the population which are not covered by cancer registries; (iii) smooth maps of gender-, age- and site-specific patterns of lung and other tobacco-related cancer mortality over time. These maps identify discrepancies of disease burden and assist implementation and evaluation of the National Programme Tobacco; (iv) a better insight into the differences in tobacco-related cancer mortality rates between linguistic regions and urbanisation by gender; (v) smooth maps of gender-specific patterns of lung cancer incidence for a recent period; (vi) estimates of the geographical patterns of tobacco-related cancer mortality for the next 10 years: this information is useful for planning purposes such as resource allocation and costing of medical supplies for diagnosis and treatment; (vii) smooth maps of lung cancer mortality adjusted for non-smoking risk factors which allow to find a good proxy of smoking behaviour in order to study gender-specific geographical patterns of smoking in Switzerland from limited smoking surveys.
The methods, modelling approaches and results are discussed and presented in five manuscripts included as chapters in this thesis. Each chapter of the thesis provides a detailed discussion. This section gives an overview about the main findings of each piece of work, its limitations, impact and an outlook covering open issues and potential future directions.

7.1 Significance

7.1.1 Spatio-temporal modelling of cancer data

Age-period-cohort (APC) analysis is an established method for predicting disease future burden. The most commonly used models for cancer projections are the Power model (implemented within the R software package Nordpred) (Møller et al., 2003) and the Bayesian APC-model which is based on a log-link function (Bray, 2002). In Chapter 4 we address limitations of the above models by developing the Bayesian power model. This is the first model applying a link-function other than the log-link within the Bayesian framework and therefore expands the options for modelling cancer data. The model overcomes extreme estimation of rates arising from the log-link function and benefits from the Bayesian framework, which smooths among the subgroups of age, period and cohort. Furthermore, our model allows the inclusion of covariates into the regression model (e. g. information on tobacco-consumption), which can improve estimation of future cases of lung cancer and incorporates random effects to take into account spatial variation. We further developed the Bayesian power model to a generalized analogue with a random power parameter

estimated by the data under consideration. The improvement of our models over the existing ones in forecasting cancer mortality rates are shown in Chapter 4, where the different modelling approaches are compared on Swiss lung cancer mortality data.
The model developed in Chapter 5 contributes to the methodology of back-calculating cancer incidence from observed mortality data. A single-model formulation estimates age-, gender- and period-specific survival distributions and incidence by combining survival parameters with observed subgroup-specific mortality via back-calculation. This combined approach allows the uncertainty of the survival curve estimation to be taken into account in the back-calculation. Furthermore, model formulations can be easily adjusted for risk factors by including covariates. Our methodology overcomes limitations of the widely used MIAMOD (Mortality-Incidence Analysis Model) software (Angelis et al., 1994), which models cancer incidence based on a maximum likelihood regression and cannot consider analysis-specific covariates. Moreover, the latter assumes survival to be known input data.
Our back-calculation model assumes gamma prior distributions for the incidence parameters. This is a more appropriate choice than the previously used normal distributions (Mezzetti and Robertson, 1999), as it allows for a skewed distribution of the incidence, which is common for rare cancer sites or analyses at a subregional level.

7.1.2 Tobacco-related cancer spatial epidemiology in Switzerland

It is well known that tobacco accounts for the majority of global morbidity and mortality. Among others, it is the main risk factor for several cancer sites, in particular for cancer of the lung. Although consequences of smoking have been observed and defined since decades, yet a high proportion of the population consumes tobacco. While in the first half of the last century its consumers were mainly males, female smoking has developed in the second half, mainly as a side-effect of the emancipation. A further trend arose when the tobacco industry targeted another audience, which was teenagers.
During the last decade, cancer mapping has become the state-of-art in assessing spatial patterns of the disease burden. A map, illustrating the spatial distribution of the cancer, gives a clear overview of the disease and is rather intuitive, as endemic areas are highlighted and can be easily identified, while numbers might not be that easily understood. Cancer maps of crude rates can be misleading and non-informative since estimates from areas with small population may dominate the map resulting in high variability. Estimates based on Bayesian spatial models highlight spatial patterns by smoothing the rates, taking into account the neighbouring structure of the areas. Furthermore, these models can adjust for risk factors and assess their contribution taking into account geographical correlation. Consideration of the latter is crucial, as spatial association might be present, violating the assumption of independence. Maps at high geographical resolution are particularly useful, as they give insights into local spatial patterns. Maps of the residuals of

spatial regressions can indicate geographical patterns that are not explained by the risk factors considered in the model. Cancer maps over time illustrate potential temporal trends of the disease. Cancer atlases are available for several countries, e. g. U. S., United Kingdom or Germany, providing a collection of cancer site-specific nationwide maps.
In Switzerland, data on cancer mortality and partially cancer incidence are available. However, no countrywide model-based cancer maps, adjusted for risk factors, exist. Schüler and Bopp (1997) published an atlas of cancer mortality for the period 1970–1990, based on maps of raw data and unadjusted analysis. Recently, Oberaigner and Vittadello (2010) presented cancer maps in the alpine regions from 2001–2005. However, the maps only cover the eastern part of the country and do not provide analysis of cancer-specific risk factors.
In this thesis, first countrywide smooth estimates of age- and gender-specific tobacco-related cancer mortality are provided at a high geographical resolution, introducing spatial correlation via Bayesian Conditional Autoregressive Models (CAR) (Besag et al., 1991; Bernardinelli and Montomoli, 1992). Maps of gender-specific model-based estimates showed a contradictory temporal trend of tobacco-related cancer mortality, decreasing for males and increasing for females. Furthermore, risk factor analysis presented significant differences in gender-specific tobacco-related cancer mortality among the Swiss language regions and level of urbanization. The latter could easily be identified by the smooth nationwide maps, as female mortality shows high estimates in the cities, i. e. Zurich and Basel.
In addition to estimation of tobacco-related cancer mortality this thesis provides spatio-temporal estimates of lung cancer mortality. However, both analyses showed similar patterns.

7.1.3 Intervention planning

In Switzerland, the National Programme Tobacco 2008–2012 (NPT 2008–2012) has been launched by the Federal Office of Public Health. The programme has been extended up to 2016. The overall aim of the NPT is to reduce the number of tobacco-related morbidity and mortality in the country. Among others, one strategic goal of the programme is the ratification (practical implementation) of the WHO Framework Convention on Tobacco Control, which was signed in 2004 and defines policies regarding tobacco control at a global scale (WHO, 2013a). Results of our risk factor analyses and the estimated maps are important for implementation and evaluation of the NPT, as they can be used to target and intensify anti-tobacco campaigns in high priority areas. In particular, the maps can serve as baseline ones to assess the success of the programme implementation in the future.
Furthermore, the NPT 2008–2012 aims to sensitize the population and specific target groups about the consequences of smoking (Federal Office of Public Health, 2010). Smoothed nationwide maps of tobacco-related cancer mortality can serve as powerful tools for supporting the above strategy of the NPT 2008–2012. Maps illustrating the spatial as well as temporal distribution and extent of

the disease might increase the awareness of its burden, which in turn can increase the acceptance of tobacco interventions in the country. Furthermore, awareness of smoking consequences may be raised at an individual level as maps allow people to identify the area of their residence and get information of the disease mortality in their region.
Estimates on the future trends of gender-specific tobacco-related cancer mortality might also increase the awareness of the population. Especially for a disease like lung cancer with a time lag of 20–30 years, future predictions might be linked to recent personal behaviour. In the case of female lung cancer, strictly increasing mortality rates could sensitize women who are currently smoking. Those estimates are presented at national, language-region and cantonal level in Chapter 3. In addition, the latter is illustrated via maps and might support the acceptance of tobacco control planning, in particular at cantonal level.

7.1.4 Public health

The global burden due to tobacco has been estimated to be 4.9 million deaths (8.8% of total) worldwide and 59.1 disability life years (DALYs) (4.1% of total) in 2000 (WHO, 2013b). According to WHO, a continuation of the current trend will lead to an expected increase up to 10 million deaths per year in the 2020s or early 2030s, while 70% of those deaths will be in developing countries. Future trends of a disease are important to be able to project the burden.
In Chapter 3 and 4 we quantified the future trends of age- and gender-specific cancer mortality due to tobacco use. Trends of all tobacco-related and lung cancer mortality are presented for each gender at different regional scale – nationwide, language-regional and cantonal. The latter is an important administrative level in Switzerland, as many laws are not decided by the federal institutions and therefore can be refined at cantonal level, as for example smoking ban regulations.
Estimation of the burden of a disease requires consideration of different data sources. Beside the knowledge of how many people die from a disease, it is essential to know how many people are actually having the disease. Temporal trends of disease-specific survival distributions, derived by incidence and mortality data, reflect developments in treatment. Overall incidence or mortality provide a general picture, while gender- and age-specific trends allow for detailed assessment of interventions. In Switzerland, cancer incidence data are recorded at cancer registries, which only covered 60.7% of the country in 2005–2009, mainly due to cantons in the German-speaking part. While those cantons are establishing new registries, data are not available in the past for long-term analyses. In Chapter 5 we presented nationwide estimates of age- and gender-specific lung cancer incidence at cantonal level. Our results do not only give information on lung cancer incidence for those cantons without registries, but also contribute to the data completeness of all cantons as they cover the time period starting in 1981. At that time less than 50% of the Swiss population was covered by registries.
Smoking is a factor describing the public health situation of a country. In Switzerland, tobacco

use accounts for more than 9 000 deaths, which is 14% of all deaths (in 2007). Information on the smoking behaviour of the Swiss population is sparse. The Swiss Health Survey (SHS) has been conducted in 1992, 1997, 2002 and 2007. Smoking data of those telephone-surveys have been analysed regarding its temporal trend (Federal Office of Public Health, 2010). However, to our knowledge, spatial and spatio-temporal analyses of smoking behaviour have not been conducted so far. In Chapter 6 we studied gender-specific spatial smoking patterns in Switzerland based on survey data in 1992. Analysis of the smoking data highlighted gender-specific differences by language regions and urbanization level, resembling gender-specific patterns of lung cancer mortality presented in Chapter 2. In addition, the age- and gender-specific proportional attributable fraction of lung cancer due to smoking in 2010 has been estimated, which quantifies the proportional reduction in lung cancer mortality if the risk factor smoking would be eliminated. These estimates quantify the impact of smoking in Switzerland.
We identified the spatial pattern of those risk factors not considered in the analysis and assessed the hypothesis that unadjusted lung cancer mortality rates are more appropriate to serve as a proxy of smoking patterns in space than the adjusted spatial random effects.
The tobacco epidemic in Switzerland accounts for CHF 5–10 billion every year. On the contrary, a yearly amount of CHF 100 million is spent by the tobacco industry to advertise smoking (Lee and Glantz, 2001). Meanwhile, many European countries introduced a law against tobacco advertisement, except Germany, Switzerland and Greece. Nowadays Philip Morris and British-American Tobacco are the leading companies controlling the Swiss tobacco market. In Switzerland, the tobacco lobby is influential and might control interventions. It has been criticized that smoking interventions do not take into account the actions and influence of the local tobacco lobby (Lee and Glantz, 2001). Future control planning should consider this important factor and research should not be driven by the interests of the tobacco industry.

7.2 Limitations

Maps of age- and gender-specific tobacco-related cancer mortality at municipality level have been obtained for the years around the censuses, i. e. 1969–1972, 1979–1982, 1989–1992 and 1999–2002. We did not consider mortality for the intercensual years, as population data were only provided for the census years. By assuming no big changes in the size of the population for the years close to the census, we pooled those years. Analysis was done for subgroups at a high geographical scale. As the country is separated into nearly 2 700 municipalities, this resulted in many zero counts in municipalities characterized by the subgroups. Therefore, simple linear interpolation between two censuses was inappropriate, as several factors should be taken into consideration to achieve realistic estimates. For example, migration plays an important role for the estimation of population development and differs depending on the location. Municipalities close to boarders

and big cities as Zurich might be more affected than rural areas. However, Proportional Mortality Ratios, relating cause-specific deaths to all-cause mortality, have been estimated over the whole time period and confirmed that no trend was missed by our period-specific analysis.
Although mortality data were available for later years, we did not consider them, as the corresponding population size was missing. Also, as the analysis was done before the publication of the census in 2010, we did not consider recent years.
The APC-models developed and applied in this thesis do not take into account environmental and behavioural factors, i. e. smoking habits. Lung cancer morbidity and mortality are strongly influenced by smoking trends and further information included in the model would clearly improve the estimation of future trends of the disease burden.
Information on tobacco-consumption is rather sparse in Switzerland. The SHS has been conducted in 1992, 1997, 2002 and 2007, which partially requested information about the smoking pattern of the person. This information is only given for a certain part of the country.
In Chapter 5 countrywide incidence estimates were obtained at cantonal level. However, these estimates are based on the assumption of a constant gender- and period-specific survival distribution and a fixed proportion of misclassification in the FSO death certificate database across cantons. Incidence from other cancer registries would allow us to investigate on these assumptions, further analyse the spatial distribution of survival and improve incidence estimation.
Data availability of smoking behaviour in Switzerland was limited. Data used in Chapter 6 helped identifying risk factors and highlighted spatial patterns, however, due to limited coverage an analysis at municipality level was not feasible.

7.3 Extension of the work

Our work cannot only be improved by addressing the above limitations, but also by a number of potential extensions. An immediate next step will be to extend this work to other cancer types with the aim of obtaining an up-to-date cancer atlas providing a complete overview of the burden of all cancer sites in Switzerland indicated by nationwide maps, future predictions of the number of cases and contribution of risk-factors. In fact, built on this work, methods are currently applied within a new project at the Swiss Tropical and Public Health Institute looking at gender-specific cancer in Switzerland. Furthermore, we plan to apply the developed models to other chronic diseases in Switzerland in order to analyse spatio-temporal patterns and project disease future trends.
To overcome the lack of intercensual population counts, Bayesian migration models will be developed to allow for estimation of the number of men and women at a certain age living in each municipality. Among others, these models will take into account important factors as migration

as well as demographic developments.
The nationwide lung cancer incidence maps are based on survival parameters estimated from data from the St. Gallen-Appenzell cancer registry. Incidence estimates can be improved when data from other cancer registries are used to take into account geographical variation in site-specific cancer survival distributions as well as in the disagreement of cause-specific death classification between cancer registries and the national mortality database built up on death certificates.
In this thesis, mortality due to cancer has been analysed. However, several other diseases, e. g. cardiovascular diseases, exist, which might give further information to the health impact of smoking. In particular those diseases, which are not characterized by a time lag of some decades between exposure/smoking and the disease, might be helpful for assessing the impact of control interventions and different smoking regulations. As discussed in Chapter 2, the spatial heterogeneity of approaches regarding protection against passive smoking in Switzerland offers the possibility to easily assess their impacts.

Bibliography

Angelis, GD, Angelis, RD, Frova, L, Verdecchia, A, 1994. MIAMOD: A computer package to estimate chronic disease morbidity using mortality and survival data. *Computer Methods and Programs in Biomedicine*, 44(2):99–107.

Armero, C, Cabras, S, Castellanos, M, Perra, S, Quirós, A, Oruezábal, M, Sánchez-Rubio, J, 2012. Bayesian analysis of a disability model for lung cancer survival. *Statistical Methods in Medical Research*.

Bailey, T, Gatrell, A, 1995. *Interactive spatial data analysis*. Longman Scientific & Technical.

Baker, A, Bray, I, 2005. Bayesian projections: What are the effects of excluding data from younger age groups? *American Journal of Epidemiology*, 162(8):798–805.

Bedford, RL, 2009. Utility of death certificate data in predicting cancer incidence. University of Iowa. PhD thesis.

Bernardinelli, L, Montomoli, C, 1992. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine*, 11(8):983–1007.

Berrut, S, Junker, C, 2008. Von Generation zu Generation: Entwicklung der Todesursachen 1970 bis 2004, Bundesamt für Statistik, Neuchâtel, Schweiz.

Besag, J, York, J, Mollié, A, 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1):1–20.

Best, N, Richardson, S, Thomson, A, 2005. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, 14(1):35–59.

Bordoni, A, Lorez, M, Bouchardy, C, Camey, B, Betticher, D, Dehler, S, Gautschi, O, Konzelmann, I, Clough-Gorr, K, the NICER Working Group, 2012. Trends in lung cancer survival in Switzerland. *Schweizer Krebsbulletin*, 2:107–122.

Bouchardy, C, Lutz, JM, Kühni, C, 2011. Cancer in Switzerland. Situation and development from 1983 to 2007. Neuchâtel: Federal Statistical Office, National Institute for Cancer Epidemiology and Registration, Swiss Childhood Cancer Registry.

Boyle, P, 1997. Cancer, cigarette smoking and premature death in Europe: A review including the recommendations of European Cancer Experts Consensus Meeting, Helsinki, October 1996. *Lung Cancer*, 17(1):1–60.

Bray, F, Møller, B, 2006. Predicting the future burden of cancer. *Nature Reviews Cancer*, 6(1):63–74.

Bray, I, 2002. Application of Markov chain Monte Carlo methods to projecting incidence and mortality. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 51(2):151–164.

Breslow, NE, Day, NE, 1987. Statistical methods in cancer research. Volume II – The design and analysis of cohort studies. *IARC Scientific Publications*, 82:1–406.

Brookmeyer, R, Gail, MH, 1988. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *Journal of the American Statistical Association*, 83(402):301–308.

Burnett, CA, Dosemeci, M, 1994. Using occupational mortality data for surveillance of work-related diseases of women. *Journal of Occupational Medicine*, 36(11):1199–1203.

Cancer Projections Network (C-Projections), 2010. Comparison of short-term projection methods: Proposing the hybrid approach and Bayesian models. Alberta Health Services.

Ceschi, M, Bopp, M, Dick, A, Probst-Hensch, N, 2009. Krebs im Kanton Zürich – Ein Bericht des Krebsregisters. Institut für Sozial- und Präventivmedizin der Universität Zürich.

Chaturvedi, P, 2004. Tobacco is a "weapon of mass destruction". Should Western countries be invaded for that? *Postgraduate Medical Journal*, 80(946):477.

Chellini, E, Gorini, G, Martini, A, Giovannetti, L, Costantini, AS, 2006. Lung cancer mortality patterns in women resident in different urbanization areas in central Italy from 1987–2002. *Tumori*, 92(4):271–275.

Clayton, D, Kaldor, J, 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43(3):671–681.

Clements, MS, Armstrong, BK, Moolgavkar, SH, 2005. Lung cancer rate predictions using generalized additive models. *Biostatistics*, 6(4):576–589.

Clèries, R, Ribes, J, Esteban, L, Martinez, JM, Borràs, JM, 2006. Time trends of breast cancer mortality in Spain during the period 1977–2001 and Bayesian approach for projections during 2002–2016. *Annals of Oncology*, 17(12):1783–1791.

Coleman, MP, Gatta, G, Verdecchia, A, Estève, J, Sant, M, Storm, H, Allemani, C, Ciccolallo, L, Santaquilani, M, Berrino, F, the EUROCARE Working Group, 2003. EUROCARE-3 summary: Cancer survival in Europe at the end of the 20th century. *Annals of Oncology*, 14(suppl 5):v128–v149.

Collett, D, 1994. *Modelling survival data in medical research.* Texts in statistical science. Chapman & Hall.

Colonna, M, Grosclaude, P, Faivre, J, Revzani, A, Arveux, P, Chaplain, G, Tretarre, B, Launoy, G, Lesec'h, JM, Raverdy, N, Schaffer, P, Buémi, A, Ménéegoz, F, Black, RJ, 1999. Cancer registry data based estimation of regional cancer incidence: Application to breast and colorectal cancer in French administrative regions. *Journal of Epidemiology & Community Health*, 53(9):558–564.

Costanza, MC, Salamun, J, Lopez, AD, Morabia, A, 2006. Gender differentials in the evolution of cigarette smoking habits in a general European adult population from 1993–2003. *BMC Public Health*, 6:130.

Couraud, S, Zalcman, G, Milleron, B, Morin, F, Souquet, P, 2012. Lung cancer in never smokers – A review. *European Journal of Cancer*, 48(9):1299–1311.

Curtin, F, Morabia, A, Bernstein, M, 1997. Smoking behavior in a Swiss urban population: The role of gender and education. *Preventive Medicine*, 26(5):658–663.

Czado, C, Gneiting, T, Held, L, 2009. Predictive model assessment for count data. *Biometrics*, 65(4):1254–1261.

Darby, S, Hill, D, Doll, R, 2001. Radon: A likely carcinogen at all exposures. *Annals of Oncology*, 12(10):1341–1351.

Darby, S, Hill, D, Auvinen, A, Barros-Dios, JM, Baysson, H, Bochicchio, F, Deo, H, Falk, R, Forastiere, F, Hakama, M, Heid, I, Kreienbrock, L, Kreuzer, M, Lagarde, F, Mäkeläinen, I, Muirhead, C, Oberaigner, W, Pershagen, G, Ruano-Ravina, A, Ruosteenoja, E, Schaffrath Rosario, A, Timarche, M, Tomáscaron, L, Whitley, E, Wichmann, HE, Doll, R, 2005. Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. *BMJ*, 330(7485):223.

De Angelis, R, Capocaccia, R, Hakulinen, T, Soderman, B, Verdecchia, A, 1999. Mixture models for cancer survival analysis: Application to population-based data with covariates. *Statistics in Medicine*, 18(4):441–454.

Dockery, DW, Pope, CA, Xu, X, Spengler, JD, Ware, JH, Fay, ME, Ferris, Jr., BG, Speizer, FE, 1993. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329(24):1753–1759. PMID: 8179653.

Doll, R, Cook, P, 1967. Summarizing indices for comparison of cancer incidence data. *International Journal of Cancer*, 2(3):269–279.

Doll, R, Peto, R, Doll, R, Peto, R, 1976. Mortality in relation to smoking: 20 years' observations on male British doctors. *The British Medical Journal*, 2(6051):1525–1536.

Doll, R, Peto, R, Boreham, J, Sutherland, I, 2004. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ*, 328(7455):1519.

Downing, A, Forman, D, Gilthorpe, MS, Edwards, KL, Manda, SOM, 2008. Joint disease mapping using six cancers in the Yorkshire region of England. *International Journal of Health Geographics*, 7(1):41.

Eberly, LE, Carlin, BP, 2000. Identifiability and convergence issues for Markov chain Monte Carlo fitting of spatial models. *Statistics in Medicine*, 19(17-18):2279–2294.

Eilstein, D, Zoé, U, Lim, TA, Bloch, J, 2008. Lung cancer mortality in France: Trend analysis and projection between 1975 and 2012, using a Bayesian age-period-cohort model. *Lung Cancer*, 59(3):282–290.

Estève, J, Benhamou, E, Raymond, L, 1994. Statistical methods in cancer research. Volume IV. Descriptive epidemiology. *IARC Scientific Publications*, 128:1–302.

Ezzati, M, Henley, SJ, Thun, MJ, Lopez, AD, 2005. Role of smoking in global and regional cardiovascular mortality. *Circulation*, 112(4):489–497.

Faeh, D, Minder, C, Gutzwiller, F, Bopp, M, 2009. Culture, risk factors and mortality: Can Switzerland add missing pieces to European puzzle? *Journal of Epidemiology & Community Health*, 63(8):639–645.

Fano, V, Michelozzi, P, Ancona, C, Capon, A, Forastiere, F, Perucci, CA, 2004. Occupational and environmental exposures and lung cancer in an industrialised area in Italy. *Occupational & Environmental Medicine*, 61(9):757–763.

Federal Office of Public Health, 2010. Nationales Programm Tabak. NPT 2008–2012.

Federal Office of Public Health, 2011. Nationaler Radonaktionsplan 2012–2020.

Federal Office of Public Health, 2012a. Rauchen verursacht tödlichen Lungenkrebs. `http://www.bag.admin.ch/themen/drogen/00041/00612/03652/03659/index.html?lang=de`. Accessed 29 May 2012.

Federal Office of Public Health, 2012b. Radon. `http://www.bag.admin.ch/themen/strahlung/00046/index.html?lang=de`. Accessed 29 May 2012.

Federal Statistical Office, 2009. Smoking-attributable mortality in Switzerland. Estimation for the years 1995 to 2007. neuchâtel.

Federal Statistical Office, 2012a. Sprachgebiete der Schweiz 2000. `http://www.bfs.admin.ch/bfs/portal/de/index/themen/01/05/blank/key/sprachen.html`. Accessed 29 May 2012.

Federal Statistical Office, 2012b. Generalisierte Gemeindegrenzen. `http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/geostat/datenbeschreibung/generalisierte_gemeindegrenzen.html`. Accessed 29 May 2012.

Federal Statistical Office, 2013a. Surveys, sources – Cancer epidemiology. `http://www.bfs.admin.ch/bfs/portal/en/index/infothek/erhebungen__quellen/blank/blank/kbs/02.html`. Accessed 29 January 2013.

Federal Statistical Office, 2013b. Cancer – Data, indicators. `http://www.bfs.admin.ch/bfs/portal/de/index/themen/14/02/05/key/02/03.html`. Accessed 31 January 2013.

Federal Statistical Office, 2013c. MS-Regionen, Arbeitsmarktregionen. `http://www.bfs.admin.ch/bfs/portal/de/index/regionen/11/geo/analyse_regionen/03.html`. Accessed 14 March 2013.

Flandorfer, P, Wegner, C, Buber, I, 2010. Gender Roles and Smoking Behaviour (Working Papers). Vienna Institute of Demographics.

Foundation National Institute for Cancer Epidemiology and Registration, 2012. Geneva Statistics of Cancer Incidence 1985–2009.

Freedman, DM, Sigurdson, AJ, Doody, MM, Love-Schnur, S, Linet, MS, 2006. Comparison between cancers identified by state cancer registry, self-report, and death certificate in a prospective cohort study of US radiologic technologists. *International Journal of Epidemiology*, 35(2):495–497.

Gelfand, A, Vounatsou, P, 2003. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics*, 4(1):11–25.

Gelfand, AE, Smith, AFM, 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.

Gilks, W, Richardson, S, Spiegelhalter, D, 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC Interdisciplinary Statistics Series. Chapman and Hall.

Gmel, G, 2000. Prevalence of tobacco use in Switzerland in the 1990's – Estimation of consumption trends based on 2 methods. *Sozial- und Präventivmedizin*, 45(2):64–72.

Gómez-Rubio, V, Ferrándiz-Ferragud, J, López-Quílez, A, 2005. Detecting clusters of disease with R. *Journal of Geographical Systems*, 7(2):189–206.

Graham, H, 1996. Smoking prevalence among women in the European community 1950–1990. *Social Science & Medicine*, 43(2):243–254.

Hakulinen, T, Dyba, T, 1994. Precision of incidence predictions based on Poisson distributed observations. *Statistics in Medicine*, 13:1513–1523.

Heishman, SJ, 1999. Behavioral and cognitive effects of smoking: Relationship to nicotine addiction. *Nicotine & Tobacco Research*, 1(Suppl 2):S143–S147.

Hirrle, B, I/2005. Lungenkrebs bei Frauen – Tendenzen und Risikofaktoren. *Onkologie*.

Hoek, G, Brunekreef, B, Goldbohm, S, Fischer, P, van den Brandt, PA, 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *The Lancet*, 360(9341):1203–1209.

International Agency for Research on Cancer, 1988. IARC monographs on the evaluation of carcinogenic risks to humans. Volume 43. Man-made mineral fibres and radon. Lyon.

International Agency for Research on Cancer, 2004. IARC monographs on the evaluation of carcinogenic risks to humans. Volume 83. Tobacco smoke and involuntary smoking. Lyon.

Johnson, GD, 2004. Small area mapping of prostate cancer incidence in New York State (USA) using fully Bayesian hierarchical modelling. *International Journal of Health Geographics*, 3(1):29.

Keller, R, Radtke, T, Krebs, H, Hornung, R, 2011. Der Tabakkonsum der Schweizer Wohnbevölkerung in den Jahren 2001 bis 2010 – Tabakmonitoring – Schweizerische Umfrage zum Tabakkonsum.

Keshamouni, VG, Arenberg, DA, Kalemkerian, GP (eds.), 2009. *Lung cancer metastasis*. Springer.

Kim, HJ, Fay, MP, Feuer, EJ, Midthune, DN, 2000. Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine*, 19(3):335–351.

Knorr-Held, L, Best, NG, 2001. A shared component model for detecting joint and selective clustering of two diseases. *Journal of the Royal Statistical Society: Series A*, 164(1):73–85.

Knorr-Held, L, Rainer, E, 2001. Projections of lung cancer mortality in West Germany: A case study in Bayesian prediction. *Biostatistics*, 2(1):109–129.

Kohler, B, Ward, E, McCarthy, B, Schymura, M, Ries, L, Eheman, C, Jemal, A, Anderson, R, Ajani, U, Edwards, B, 2011. Annual report to the nation on the status of cancer, 1975–2007, featuring tumors of the brain and other nervous system. *Journal of the National Cancer Institute*, 103(9):714–736.

Krebs, H, Keller, R, Radtke, T, Hornung, R, 2008. Werbe- und Verkaufseinschränkungen für Tabakwaren, höhere Zigarettenpreise und Rauchverbote: Einstellungen der Schweizer Bevölkerung 2006/2007.

Krebsliga Schweiz, 2012. Lungenkrebs – Bronchialkarzinom. Eine Information der Krebsliga für Betroffene und Angehörige. Krebsliga Schweiz, Bern.

Kulldorff, M, Nagarwalla, N, 1995. Spatial disease clusters: Detection and inference. *Statistics in medicine*, 14(8):799–810.

La Vecchia, C, Levi, F, Decarli, A, Wietlisbach, V, Negri, E, Gutzwiller, F, 1988. Trends in smoking and lung cancer mortality in Switzerland. *Preventive Medicine*, 17(6):712–724.

Law, J, Ping, WC, 2011. Monitoring residual spatial patterns using Bayesian hierarchical spatial modelling for exploring unknown risk factors. *Transactions in GIS*, 15(4):521–540.

Lee, C, Glantz, S, 2001. The tobacco industry's successful efforts to control tobacco policy making in Switzerland. University of California, San Francisco.

Lee, PN, Fry, JS, Forey, BA, 1990. Trends in lung cancer, chronic obstructive lung disease, and emphysema death rates for England and Wales 1941–85 and their relation to trends in cigarette smoking. *Thorax*, 45(9):657–665.

Lee, TC, Dean, C, Semenciw, R, 2011. Short-term cancer mortality projections: A comparative study of prediction methods. *Statistics in Medicine*, 30(29):3387–3402.

Lee, YC, Boffetta, P, Sturgis, EM, Wei, Q, Zhang, ZF, Muscat, J, Lazarus, P, Matos, E, Hayes, RB, Winn, DM, Zaridze, D, Wünsch-Filho, V, Eluf-Neto, J, Koifman, S, Mates, D, Curado, MP, Menezes, A, Fernandez, L, Daudt, AW, Szeszenia-Dabrowska, N, Fabianova, E, Rudnai, P, Ferro, G, Berthiller, J, Brennan, P, Hashibe, M, 2008. Involuntary smoking and head and neck cancer risk: Pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiology, Biomarkers & Prevention*, 17(8):1974–1981.

Levi, F, Lucchini, F, Negri, E, La Vecchia, C, 2004. Trends in mortality from major cancers in the European Union, including acceding countries, in 2004. *Cancer*, 101(12):2843–2850.

Levi, F, Lucchini, F, La Vecchia, C, 2006. Trends in cancer mortality in Switzerland, 1980–2001. *European Journal of Cancer Prevention*, 15(1):1–9.

Lienert, T, Serke, M, Schönfeld, N, Loddenkemper, R, 2000. Lung cancer in young females. *European Respiratory Journal*, 16(5):986–990.

Lopez, AD, Collishaw, NE, Piha, T, 1994. A descriptive model of the cigarette epidemic in developed countries. *Tobacco Control*, 3(3):242–247.

Luthi, JC, Clere-Berod, A, Favre, F, de Weck, D, 2005. Distribution géographique des cancers en Valais, 1989–2000. Registre des Tumeurs. Observatoire valaisan de la santé 2005.

MacKay, J, Eriksen, M, 2004. *The tobacco atlas*. World Health Organisation.

Mallick, BK, Gelfand, AE, 1994. Generalized linear models with unknown link functions. *Biometrika*, 81(2):pp. 237–245.

Malvezzi, M, Bertuccio, P, Levi, F, La Vecchia, C, Negri, E, 2012. European cancer mortality predictions for the year 2012. *Annals of Oncology*, 23(4):1044–1052.

Marques-Vidal, P, Cerveira, J, Paccaud, F, Cornuz, J, 2011. Smoking trends in Switzerland, 1992–2007: A time for optimism? *Journal of Epidemiology & Community Health*, 65(3):281–286.

Marshall, RJ, 1991. A review of methods for the statistical analysis of spatial patterns of disease. *Journal of Royal Statistical Society*, 154(3):421–441.

Menzler, S, Piller, G, Gruson, M, Rosario, AS, Wichmann, HE, Kreienbrock, L, 2008. Population attributable fraction for lung cancer due to residential radon in Switzerland and Germany. *Health Physics*, 95(2):179–189.

Mezzetti, M, Robertson, C, 1999. A hierarchical Bayesian approach to age-specific back-calculation of cancer incidence rates. *Statistics in Medicine*, 18(8):919–933.

Mitton, N, Colonna, M, Trombert, B, Olive, F, Gomez, F, Iwaz, J, Polazzi, S, Schott-Petelaz, A, Uhry, Z, Bossard, N, Remontet, L, 2011. A suitable approach to estimate cancer incidence in area without cancer registry. *Journal of Cancer Epidemiology*, 2011 (2011):418968.

Møller, B, 2004. Prediction of cancer incidence – Methodological considerations and trends in the Nordic countries 1958–2022. Faculty of Medicine, University of Oslo, Unipuc AS, Oslo.

Møller, B, Fekjaer, H, Hakulinen, T, Tryggvadóttir, L, Storm, HH, Talbäck, M, Haldorsen, T, 2002. Prediction of cancer incidence in the Nordic countries up to the year 2020. *European Journal of Cancer Prevention*, 11 Suppl 1:S1–96.

Møller, B, Fekjaer, H, Hakulinen, T, Sigvaldason, H, Storm, HH, Talbäck, M, Haldorsen, T, 2003. Prediction of cancer incidence in the Nordic countries: Empirical comparison of different approaches. *Statistics in Medicine*, 22(17):2751–2766.

Møller, B, Weedon-Fekjaer, H, Haldorsen, T, 2005. Empirical evaluation of prediction intervals for cancer incidence. *BMC Medical Research Methodology*, 5 (21).

Nafstad, P, Håheim, LL, Oftedal, B, Gram, F, Holme, I, Hjermann, I, Leren, P, 2003. Lung cancer and air pollution: A 27 year follow up of 16 209 Norwegian men. *Thorax*, 58(12):1071–1076.

National Institute for Cancer Epidemiology and Registration, 2011a. Fribourg Statistics of Cancer Incidence 1985–2009.

National Institute for Cancer Epidemiology and Registration, 2011b. Graubünden Statistics of Cancer Incidence 1984–2008.

National Institute for Cancer Epidemiology and Registration, 2011c. Ticino Statistics of Cancer Incidence 1984–2008.

National Institute for Cancer Epidemiology and Registration, 2011d. Glarus Statistics of Cancer Incidence 1984–2008.

National Institute for Cancer Epidemiology and Registration & Federal Statistical Office, 2012. Switzerland Statistics of Cancer Incidence 1985–2009.

Negri, E, La Vecchia, C, Levi, F, Randriamiharisoa, A, Decarli, A, Boyle, P, 1990. The application of age, period and cohort models to predict Swiss cancer mortality. *Journal of Cancer Research and Clinical Oncology*, 116(2):207–214.

Ntzoufras, I, 2009. *Bayesian modeling using WinBUGS*. John Wiley & Sons, Inc.

Oberaigner, W, Vittadello, F, 2010. *Cancer mapping in Alpine regions 2001–2005.*

Observatoire valaisan de la santé, 2005. Le cancer en Valais 1989–2000.

Ong, HC, Soo, KL, 2006. Backcalculation of HIV infection rates in Malaysia. *Medical Journal of Malaysia*, 61(5):616–620.

Osmond, C, 1985. Using age, period and cohort models to estimate future mortality rates. *International Journal of Epidemiology*, 14(1):124–129.

PAHO, 2002. Epidemiological Bulletin - Pan American Health Organization: Celebrating 100 years of health (September 2002), Standardization: A classic epidemiological method for the comparison of rates. 23:3.

Parkin, DM, 2011. Tobacco-attributable cancer burden in the UK in 2010. *British Journal of Cancer*, 105:S6–S13.

Parkin, DM, Hakulinen, T, 1991. *Analysis of survival.*

Peto, R, Lopez, AD, Boreham, J, Thun, M, Heath, C, 1992. Mortality from tobacco in developed countries: Indirect estimation from national vital statistics. *Lancet*, 339(8804):1268–1278.

Pregibon, D, 1980. Goodness of link tests for Generalized Linear Models. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 29(1):pp. 15–23, 14.

Pury, P, Lutz, J, Junker, C, Gantenbein, E, 2007a. Statistics of cancer mortality 1986–2005. Swiss Association of cancer registries 2007.

Pury, P, Lutz, JM, Jundt, G, et al, 2007b. Statistics of cancer incidence 1986–2005. Swiss Association of cancer registries 2007.

Qiu, Z, Hatcher, J, Wang, M, Cancer Projection Analytical Network Working Team. Review cancer projections methods for Canadian Partnership Against Cancer Analytical Network. Alberta Health Service.

Qiu, Z, Jiang, Z, Wang, M, Hatcher, J, 2010. Cancer Projections Network (C-Projections). Long-term projection methods: Comparison of age-period-cohort model-based approaches. Alberta Health Services.

Richardson, S, Abellan, JJ, Best, NG, 2006. Bayesian spatio-temporal analysis of joint patterns of male and female lung cancer risks in Yorkshire (UK). *Statistical Methods in Medical Research*, 15(4):385–407.

Samet, JM, Avila-Tang, E, Boffetta, P, Hannan, LM, Olivo-Marston, S, Thun, MJ, Rudin, CM, 2009. Lung cancer in never smokers: Clinical epidemiology and environmental risk factors. *Clinical Cancer Research*, 15(18):5626–5645.

Sánchez, MJ, Payer, T, De Angelis, R, Larrañaga, N, Capocaccia, R, Martinez, C, 2010. Cancer incidence and mortality in Spain: Estimates and projections for the period 1981–2012. *Annals of Oncology*, 21 (suppl 3):iii30–36.

Schüler, G, Bopp, M, 1997. *Atlas der Krebsmortalität in der Schweiz, 1970–1990.* Birkhäuser Verlag.

Schopper, D, Obrist, R, 2005. Nationales Krebsprogramm für die Schweiz 2005–2010. Oncosuisse, Bern.

Schrader, H, 1997. Entwicklungstendenzen ländlicher Räume in Deutschland und ihre Förderung durch die EU-Strukturpolitik. In: Die ländlichen Räume in Deutschland und deren Besonderheit in Mecklenburg Vorpommern. Universität Rostock.

Segi, M, 1960. Cancer mortality for selected sites in 24 countries (1950–57). Tohoku University of Medicine, Sendai.

Sitas, F, Douglas, AJ, Webster, EC, 1989. Respiratory disease mortality patterns among South African iron moulders. *British Journal of Industrial Medicine*, 46(5):310–315.

Sturtz, S, Ickstadt, K, 2007. A descriptive and model-based spatial comparison of the standardised mortality ratio and the age-standardized mortality rate. *Geospatial Health*, 1(2):255–266.

Subramanian, J, Govindan, R, 2007. Lung cancer in never smokers: A review. *Journal of Clinical Oncology*, 25(5):561–570.

Sue Bell, B, Broemeling, LD, 2000. A Bayesian analysis for spatial processes with application to disease mapping. *Statistics in Medicine*, 19(7):957–974.

Sun, DC, Tsutakawa, RK, Kim, H, He, ZQ, 2000. Spatio-temporal interaction with disease mapping. *Statistics in Medicine*, 19(15):2015–2035.

Taylor, R, Najafi, F, Dobson, A, 2007. Meta-analysis of studes of passive smoking and lung cancer: Effects of study type and continent. *International Journal of Epidemiology*, 36(5):1048–1059.

Thun, MJ, Day-Lally, C, Myers, DG, Calle, EE, Flanders, WD, Zhu, BP, Namboodiri, MM, Heath, CW, 1997. Trends in tobacco smoking and mortality from cigarette use in Cancer Prevention Studies I (1959 through 1965) and II (1982 through 1988).

Thun, MJ, Henley, SJ, Burns, D, Jemal, A, Shanks, TG, Calle, EE, 2006. Lung cancer death rates in lifelong nonsmokers. *Journal of the National Cancer Institute*, 98(10):691–699.

Thun, MJ, Hannan, LM, Adams-Campbell, LL, Boffetta, P, Buring, JE, Feskanich, D, Flanders, WD, Jee, SH, Katanoda, K, Kolonel, LN, Lee, IM, Marugame, T, Palmer, JR, Riboli, E, Sobue, T, Avila-Tang, E, Wilkens, LR, Samet, JM, 2008. Lung cancer occurrence in never-smokers: An analysis of 13 cohorts and 22 cancer registry studies. *PLoS Med*, 5(9):e185.

Torok, S, Hegedus, B, Laszlo, V, Hoda, MA, Ghanim, B, Berger, W, Klepetko, W, Dome, B, Ostoros, G, 2011. Lung cancer in never smokers. *Future Oncology*, 7(10):1195–1211.

Trichopoulos, D, Kalandidi, A, Sparros, L, MacMahon, B, 1981. Lung cancer and passive smoking. *International Journal of Cancer*, 27(1):1–4.

Tyczynski, JE, Bray, F, Parkin, DM, 2003. Lung cancer in Europe in 2000: Epidemiology, prevention and early detection. *Lancet Oncology*, 4(1):45–55.

United States Department of Health and Human Services, 2004. The health consequences of smoking: A report of the Surgeon General. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

United States Department of Health and Human Services, 2010. How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the Surgeon General. Centers for Disease Control and Prevention (US); National Center for Chronic

Disease Prevention and Health Promotion (US); Office on Smoking and Health (US). Atlanta (GA): Centers for Disease Control and Prevention (US).

United States Environmental Protection Agency, 1992. Respiratory Health Effects of Passive Smoking: Lung Cancer and Other Disorders. Office of Health and Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

United States Public Health Service, 1964. Smoking and health. United States Public Health Service. Office of the Surgeon General.

Ver Hoef, JM, Cressie, N, Fisher, RN, Case, TJ, 2001. Uncertainty and spatial linear models for ecological data. In *Spatial uncertainty in ecology: implications for remote sensing and GIS applications*, Chapter 10. Springer-Verlag, New York.

Waller, LA, Carlin, BP, Xia, H, Gelfand, AE, 1997. Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association*, 92(438):607–617.

Weinstein, ND, 2001. Public understanding of risk and reasons for smoking low-yield products, in risks associated with smoking cigarettes with low machine-measured yields of tar and nicotine. u.s. dept. of health and human services, public health service, national institutes of health, national cancer institute: Bethesda, md:193-197.

Whittle, P, 1954. On stationary processes in the plane. *Biometrika*, 41(3/4):434–449.

WHO, 2008. WHO Report on the global tobacco epidemic, 2008: The MPOWER Package. Geneva: World Health Organization.

WHO, 2009. *WHO handbook on indoor radon: A public health perspective. Geneva.* Geneva.

WHO, 2013a. WHO Framework Convention on Tobacco Control. `http://www.who.int/fctc/en/`.

WHO, 2013b. Management of substance abuse. Tobacco. `http://www.who.int/substance_abuse/facts/tobacco/en/`.

Willemsen, MC, Kiselinova, M, Nagelhout, GE, Joossens, L, Knibbe, RA, 2012. Concer about passive smoking and tobacco control policies in European countries: An ecological study. *BMC Public Health*, 12(1):876.

Wingo, PA, Landis, S, Parker, S, Bolden, S, Heath, CW, 1998. Using cancer registry and vital statistics data to estimate the number of new cancer cases and deaths in the United States for the upcoming year. *Journal of Registry Management*, 25(2):43–51.

Xia, H, Carlin, BP, Waller, LA, 1997. Hierarchical models for mapping Ohio lung cancer rates. *Environmetrics*, 8(2):107–120.