Thesis for Habilitation

in

Epidemiology and Public Health

Dr. phil. Wilhelm Oberaigner

Institute of Public Health, Medical Decision Making and HTA Department of Public Health and Health Technology Assessment UMIT – University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria

November 2011, Hall i.T., Austria

TABLE OF CONTENTS

Summary (English)	5
Zusammenfassung (Deutsch)	9
Introduction	13
Results	15
Record Linkage	15
Data Quality Aspects at the Cancer Registry of Tyrol	17
Breast Cancer	18
Prostate Cancer	21
Department Volume and Cancer Patient Survival	23
Gender and Survival in Oncology	25
Short Summary of Second and Senior Authorship Papers	
Overall Summary and Conclusion	29
References	31
List of First Authorship Papers	35
List of Second and Senior Authorship Papers	
Acknowledgements	39
APPENDIX: Declaration	41
APPENDIX: Full Papers	
B1-1 Appendix Full Papers A	ppendix 1
B1-2 Appendix Full Papers A	ppendix 7
B1-3 Appendix Full Papers A	ppendix 17
B1-4 Appendix Full Papers A	ppendix 25
B1-5 Appendix Full Papers A	ppendix 33
B1-6 Appendix Full Papers A	ppendix 43
B1-7 Appendix Full Papers A	ppendix 49
B1-8 Appendix Full Papers A	ppendix 57
B1-9 Appendix Full Papers A	ppendix 63
B1-10 Appendix Full Papers A	ppendix 71
B2-1 Appendix Full Papers A	ppendix 77
B2-2 Appendix Full Papers A	ppendix 85
B2-3 Appendix Full Papers A	ppendix 91

SUMMARY (ENGLISH)

Background. This habilitation thesis covers the primary work of the habilitation, expressed by first authorship publications developed when setting up and operating the Cancer Registry of Tyrol. One important methodological aspect involved in operating a cancer registry in Austria is the issue of linking data from different sources, because to date, there is no unique person identifier in Austria. Another important issue is the quality of the cancer registry data. The incidence data collected at the Cancer Registry of Tyrol since 1988 made it possible to address important public health questions in oncology, for example whether offering PSA testing to men free of charge leads to a reduction in prostate cancer mortality, whether the spontaneous mammography screening program in Tyrol should be changed to an organised one, whether implementation of the newly established mammography screening program in Tyrol should be changed to an organise one, whether implementation of the newly established mammography screening program in Tyrol should be changed to an organised one, whether implementation of the newly established mammography screening program in Tyrol permits us to meet the goals defined in the EU guidelines, whether there are differences in survival between female and male cancer patients, and whether there is an association between department volume and survival for gynaecological cancer sites.

Material and Methods. Most investigations employed the incidence data collected by the Cancer Registry of Tyrol, the cancer mortality data collected by Statistics Austria and population data also collected by Statistics Austria. Finally, the mammography screening database served to evaluate the newly established organised mammography screening program in Tyrol. All studies were conducted in conformity with the Helsinki Declaration. Investigations based on register data needed no approval by the local Ethics Committee. Linkage between incidence data and the mammography database employed the pseudonym number, and linkage between incidence and mortality data deployed the record linkage method developed at the Cancer Registry of Tyrol. Methods of biostatistical analysis ranged from basic epidemiological methods like age-standardised incidence rates to more sophisticated methods such as age-period-cohort models for the analysis of time trend, multivariate Cox Proportional Hazard models and multivariate relative excess risk models for estimating survival. **Results.** For the performance of record linkage tasks in a cancer registry, we developed a new probabilistic record linkage system adapted to conditions in Austria. The system achieved adequate precision, however the results also depended on the precision of individual decisions made in unclear cases. In addition, we compared this new system with commonly used deterministic record linkage methods and showed that relative differences in five-year survival equalled 26% and 16% for females and males, respectively. Consequently, we recommended using a probabilistic record linkage system.

Data quality of incidence data collected by the Cancer Registry of Tyrol was shown to be good, with the exception of ovarian cancer which was partly miscoded in the early 1990s. The method of passive assessment of patient life status implemented by the Cancer Registry of Tyrol led to only small absolute errors of 0.5% and 1.0% in five- and ten-year survival rates, respectively.

Analysis of breast cancer incidence and mortality after fifteen years of opportunistic mammography screening in Tyrol showed that only part of the mortality reduction known for organised screening programs, which was derived from randomised controlled trials conducted in several countries, could be exploited. After introducing the organised screening program in Tyrol in a pilot phase in June 2007 and extending the program to all counties of Tyrol in June 2008, we analysed performance by comparing it with recommended levels given in the well-established EU guidelines. Most of the indicators were in or very close to the recommended range for age group 50-69, except the participation rate of 55% versus the recommended level of 70% and the proportion of stage II+ cases, namely 33% versus the recommended level of 25%.

In two studies performed in 2005 and 2010, we were able to demonstrate that after offering PSA testing free of charge to all men in Tyrol aged 45 to 74 since 1993, prostate cancer mortality decreased in Tyrol by 19% (95% CI 2%-32%) and 30% (95% CI 13%-43%) using mortality data up to 2003 and 2008, respectively, each compared to 1989-93. These studies lacked the strength of randomised controlled studies. However, bearing in mind that all randomised controlled studies of the effect of PSA screening were confronted with

methodological problems (e.g., contamination, generalizability), evidence from observational studies must be considered when determining whether PSA screening has the potential to bring about a significant reduction in prostate cancer mortality.

A study of survival differences between female and male cancer patients for all cancer sites combined showed a case mix-adjusted multivariate relative excess risk in survival of 0.95 (95% CI 0.91-0.99) for females as compared to males, reaching statistical significance. However, this benefit for women is restricted to patients aged <80. In addition, a statistically significant survival benefit for females as compared to males was shown for head and neck cancer, stomach cancer and lung cancer with multivariate relative excess risks of 0.72 (95% CI 0.56-0.93), 0.86 (95% CI 0.75- 0.95) and 0.82 (95% CI 0.75-0.90), respectively. All risks were not only adjusted for other factors but also for staging distribution.

A study of the association between department volume and survival for gynaecological cancer sites and breast cancer revealed a statistically significant multivariate hazard ratio of 1.39 (95% CI 1.22-1.58) and 1.27 (95% CI 1.05-1.54) for breast cancer and ovarian cancer, respectively, each comparing small departments defined as having fewer than 12 patients per year and large departments. For cervix cancer, we calculated a multivariate hazard ratio of 0.67 (95% CI 0.51-0.88).

An analysis of patients with relapsed or refractory Hodgkin's disease or non-Hodgkin's lymphoma showed that allogeneic is superior to autologous stem cell transplantation for patients with elevated serum LDH levels and bone marrow involvement only. A new regimen with CEOP/IMVP-Dexa for the treatment of patients with untreated aggressive lymphoma showed an eight-year cumulative survival of 70% for patients in age group < 60. Finally, we were able to show that incidence of end-stage renal disease in Tyrol as compared to Austria except Tyrol was lower for patients with diabetes type 2 only and the analysis of a number of surrogate databases for estimating the prevalence of diabetes mellitus 2 showed a consistently lower prevalence of diabetes type 2 in Tyrol as compared to Austria except Tyrol.

Conclusions. The probabilistic record linkage method is well established at high quality. This method is optimised for German typing names and must be adapted for migrant names,

7

reflecting the increasing cancer incidence in migrant groups. Quality of cancer registry data is high, in the future a linkage to residence registration data would help minimize the error in assessing patient life status. The quality of the newly established mammography system in Tyrol was shown to meet most of recommendations given in the EU guidelines except participation rate. The main challenge for the future is to integrate the mammography screening program in Tyrol into the nationwide mammography system without loss of performance. We were able to show that it is likely that PSA testing reduces prostate cancer mortality; our results are consistent with recently published studies. The main question now is to assess the harms of PSA screening and to weigh the balance between prostate cancer mortality reduction and harms of PSA screening, for example, due to overdiagnosis. The question of department volume and survival of cancer patients has high public health relevance; implementation of health system decisions taken in this respect is high in sensitivity and calls for special attention. We were able to show a small benefit in survival for female cancer patients as compared to male patients, but restricted to younger patients. An important question is why the gap in survival between younger and elderly patients increases, especially in women. This question is linked both to gender aspects and geriatric oncology.

ZUSAMMENFASSUNG (DEUTSCH)

Hintergrund. Diese Habilitationsschrift bildet den Hauptteil der Habilitation und umfasst hauptsächlich die Erstautorschaften, die aus dem Aufbau und dem Betrieb des Tumorregisters Tirol entstanden. Eine wichtige methodische Frage in einem Krebsregister ist die Verbindung von Daten aus unterschiedlichen Quellen, weil es bis jetzt in Österreich keine eindeutige Personennummer gibt. Eine andere wichtige Frage ist die Beurteilung der Qualität der Registerdaten. Aufbauend auf den Inzidenzdaten des Tumorregisters Tirol war es möglich, eine Reihe von wichtigen Public Health Fragen in der Onkologie zu untersuchen: Ob PSA-Untersuchungen in einer Bevölkerung zu einer Reduktion der Mortalität des Prostatakarzinoms führen, ob das spontane Mammographie-Screening Programm in Tirol in ein organisiertes Programm umgewandelt werden soll, ob die Qualität des neu eingeführten organisierten Mammographie-Screening Programms Tirol die Qualitätsvorgaben der EU Guidelines erfüllt, ob es Unterschiede im Überleben zwischen weiblichen und männlichen KrebspatientInnen gibt und ob ein Zusammenhang zwischen Abteilungsgröße und Überleben von KrebspatientInnen nachgewiesen werden kann.

Material und Methoden. Die meisten Untersuchungen basierten auf den Inzidenzdaten des Tumorregisters, auf den Mortalitätsdaten und Bevölkerungsdaten, die von der Statistik Austria registriert werden, und auf der Screening-Datenbank des Mammographie-Screening Programms Tirol. Alle Untersuchungen wurden unter Einhaltung der Deklaration von Helsinki durchgeführt. Studien, die auf Registerdaten basieren, benötigten kein Votum der Ethik-Kommission. Inzidenzdaten und Mammographie-Datenbank wurden auf Basis der Pseudonyme verbunden, für die Verbindung der Inzidenz- mit den Mortalitätsdaten wurde das neu entwickelte Record-Linkage Programm verwendet. Die eingesetzten biostatistischen Methoden reichten von Basismethoden der Epidemiologie wie altersstandardisierte Inzidenzraten zu sophistizierteren multivariaten Modellen, wie zum Beispiel Alter-Perioden-Kohorten-Modelle für die Analyse von Zeittrends und multivariate Cox-Proportional-Hazard Regressionsmodelle und relative Exzess-Risiko-Modelle für die Analyse von Überlebensraten. Ergebnisse. Um die Record-Linkage Aufgaben in einem Krebsregister durchzuführen, haben wir ein probabilistisches Record-Linkage System entwickelt. Das System erfüllte die gesetzten Qualitätsansprüche, allerdings hing die Qualität der Ergebnisse auch von individuellen Entscheidungen von unklaren Fällen ab. Außerdem haben wir das probabilistische Record-Linkage System verglichen mit deterministischen Record-Linkage Systemen, die bis 2005 immer noch häufig angewandt wurden. Die Anwendung der deterministischen Record-Linkage Systeme bei Überlebensanalysen führte zu einer Überschätzung von relativen Fünfjahresüberlebensraten von 26% (Frauen) und 16% (Männer), daher haben wir die Anwendung von probabilistischen Record-Linkage Systemen empfohlen.

Eine umfassende Analyse der Datenqualität im Tumorregister Tirol konnte die sehr gute Qualität belegen mit Ausnahme der Ovarialkarzinome, die in der ersten Hälfte der 1990er Jahre zum Teil falsch kodiert waren (Borderline-Karzinome). Die passive Methode der Erhebung des Überlebensstatus von PatientInnen führte nur zu kleinen absoluten Fehlern von 0.5% in der Fünfjahresüberlebensrate und 1.0% in der Zehnjahresüberlebensrate.

Die Analyse der Mammakarzinom-Mortalität nach fünfzehn Jahren spontanem Mammographie-Screening in Tirol hat nicht die Reduktion der Mammakarzinom-Mortalität gebracht, die man nach den Ergebnissen der randomisierten Screening-Studien von gut organisierten Screening-Programmen erwartet. Im Juni 2007 wurde das organisierte Mammographie-Screening Modell Tirol eingeführt, zuerst in einer Pilotphase in den zentralen Bezirken Innsbruck-Stadt und Innsbruck-Land, und anschließend im Juni 2008 ausgedehnt auf ganz Tirol. Wir haben die Qualitätsindikatoren mit den Vorgaben der EU-Guidelines verglichen. Die meisten Qualitätsindikatoren lagen im Bereich der EU-Empfehlungen für die Altersgruppe 50-69. Die zwei Ausnahmen waren die Teilnahmerate von 55% (Empfehlung von 70%) und der Anteil der entdeckten Karzinome mit Stadium II+ mit 33% (Empfehlung von 25%).

In zwei Studien konnten wir nachweisen, dass sich nach dem Angebot von kostenlosen PSA-Untersuchungen seit 1993 für alle Männer in Tirol im Alter 45 bis 74 die Prostatakarzinom-Mortalität reduziert hat, und zwar in den Jahren 1999-2003 um 19% (95% KI 2%-32%) und in den Jahren 2004-2008 um 30% (95% KI 13%-43%), jeweils im Vergleich zu 1989-93. Es handelte sich dabei um eine Beobachtungsstudie, die aber angesichts der methodischen Probleme der kontrollierten randomisierten Studien (wie Kontamination der Kontrollarme, Verallgemeinerbarkeit) einen wichtigen Beitrag zum Wissen über den Zusammenhang zwischen PSA-Screening und Prostatakarzinom-Mortalität liefern kann.

Eine Studie zu Überlebensunterschieden zwischen weiblichen und männlichen KrebspatientInnen basierend auf den Tiroler Inzidenzdaten von 1988 bis 2003 hat für die Zusammenfassung aller Krebslokalisationen für Frauen ein um 5% (95% KI 1%-9%) erniedrigtes Exzess-Risiko für Überleben gezeigt (statistisch signifikant), allerdings nur bis zum Alter 80. Außerdem konnte ein statistisch signifikant erniedrigtes multivariates Exzessrisiko aus Sicht der Frauen für HNO-Karzinome (28%, 95% KI 7%-44%), für Magenkarzinome (14%, 95% KI 5%-25%) und für Lungenkarzinome (18%, 95% KI 10%-25%) nachgewiesen werden.

Eine Studie über den Zusammenhang von Abteilungsgröße und Krebsüberleben für Mammakarzinom, Ovarialkarzinom, Zervixkarzinom und Endometriumkarzinom zeigte für kleine Abteilungen (definiert durch weniger als 12 Patientinnen pro Jahr) ein multivariates Hazard-Ratio von 1.39, 95% KI 1.22-1.58 (Mammakarzinom), 1.27, 95% KI 1.05-1.54 (Ovarialkarzinom) und 0.67, 95% KI 0.51-0.88 (Zervixkarzinom), jeweils statistisch signifikant.

Eine Analyse von PatientInnen mit rezidivierenden oder refraktären Hodgkin und Non-Hodgkin Lymphomen zeigte, dass nur bei PatientInnen mit einem erhöhten LDH-Wert und gleichzeitigem Knochenmarksbefall eine allogene Knochenmarkstransplantation besser war als eine autologe. Für ein neues Behandlungsschema mit CEOP/IMVP-Dexa bei der Behandlung von unbehandelten aggressiven Lymphomen konnte gezeigt werden, dass die Überlebensrate nach acht Jahren bei PatientInnen mit Alter bis 60 bei ausgezeichneten 70% lag. Schließlich wurde die Frage der niedrigen Nierenersatztherapie-Inzidenzrate in Tirol untersucht. Erstens konnte nachgewiesen werden, dass die einzige Gruppe mit einer erniedrigten Inzidenzrate PatientInnen mit Diabetes Mellitus Typ 2 als Grunderkrankung waren. Zweitens konnte durch eine Analyse von einer Reihe unterschiedlicher Datenquellen ein konsistentes Bild von erniedrigter Prävalenz von Diabetes Mellitus Typ 2 in Tirol im Vergleich zu den anderen österreichischen Bundesländern nachgewiesen werden. Schlussfolgerungen. Das System des probabilistischen Record-Linkage ist gut etabliert, die gute Qualität der Ergebnisse konnte belegt werden. Die Methode enthält wichtige Komponenten, die spezifisch für deutsche Namen sind und diese müssen an die Migrationsgruppen angepasst werden. Die Qualität der Inzidenzdaten des Tumorregisters ist hoch, es waren nur minimale Optimierungen der Abläufe notwendig. Für die Zukunft wäre ein Zugriff auf die Meldedatenbank äußerst wichtig, um die ohnehin kleinen Fehler in der Erhebung des Überlebensstatus zu minimieren. Die gute Qualität des neu eingeführten organisierten Mammographie-Screening Programms Tirol konnte nachgewiesen werden. Bis auf die Teilnahmerate und den Anteil der Karzinome mit Stadium II+ liegen alle Qualitätsindikatoren im Bereich der EU-Empfehlungen. Hauptfrage wird es sein, das Tiroler Modell ohne Qualitätsverlust in das österreichweite Screening-Programm zu integrieren. Wir konnten belegen, dass sich in Tirol nach dem Angebot von PSA-Untersuchungen für Männer im Alter 45-74 die Prostatakarzinom-Mortalität um 30% erniedrigt hat. Hauptfrage ist jetzt die Beurteilung des Schadens der PSA-Untersuchungen. Die Frage des Zusammenhangs zwischen Abteilungsgröße und Behandlungserfolg für KrebspatientInnen ist von großer Public Health Relevanz, die gesundheitspolitische Umsetzung solcher Ergebnisse ist allerdings äußerst komplex und sollte wissenschaftlich begleitet werden. Die Frage nach Überlebensunterschieden zwischen weiblichen und männlichen KrebspatientInnen führt zu wichtigen Fragen der geriatrischen Onkologie, weil sich der Überlebensvorteil nur für Frauen bis Alter 80 zeigt.

INTRODUCTION

This *habilitation* thesis summarises the primary work of the *habilitation* consisting of research projects led by Wilhelm Oberaigner when setting up and operating the Cancer Registry of Tyrol. The results of this work have been published as first authorship publications in international peer-reviewed journals. The work reports on applications of epidemiological methods to public health questions in oncology. In addition, work in which W. Oberaigner was involved as second or senior author of publications in peer-reviewed international journals is summarised in brief.

All first authorship papers address research questions relevant to public health and make use of data from the Cancer Registry of Tyrol (CRT). The research questions ranged from methodological aspects when running population-based cancer registries to relevant public health questions in oncology. We investigated both questions concerning specific cancer sites and questions for all cancer cases combined, for example the question of survival differences between female and male cancer patients. Mainly two cancer sites were analysed that were of special public health interest in Tyrol. The first was prostate cancer: the main question being whether offering prostate-specific antigen (PSA) testing free of charge to all men aged 45-75 leads to a reduction in prostate cancer mortality. The second was breast cancer, where two questions were investigated. Firstly, whether it is necessary to change the spontaneous mammography screening program in Tyrol to an organised program, and secondly, whether the newly established organised mammography screening program can reach the performance goals defined in the EU guidelines. This analysis was based mainly on the screening database established in Tyrol. Moreover, the screening database was linked to the breast cancer cases registered in the CRT.

All studies were conducted in conformity with the Helsinki Declaration.¹ Investigations based on registry data needed no approval by the local Ethics Committee.

In all first authorship papers, the author was either the principal investigator or substantially contributed to developing the research questions, selecting the study design, performing the epidemiological analysis and writing the manuscript.

The chapter Record Linkage deals with one methodological aspect that is very important in a cancer registry in Austria, namely how to conduct linkage between various data sources for patients lacking a unique person identifier. The chapter Data Quality Aspects at the Cancer Registry of Tyrol contains a comprehensive analysis of data quality aspects at the CRT. Next, the chapter Breast Cancer covers the situation of breast cancer incidence and mortality after fifteen years of a spontaneous mammography screening program in Tyrol as well as evaluation of the performance of the newly established organised mammography screening program in Tyrol. The chapter Prostate Cancer analyses the question of prostate cancer mortality reduction after offering PSA testing free of charge to all men in Tyrol aged 45-74. The chapter Gender and Survival in Oncology focuses on the question whether there is a difference in survival between female and male cancer patients. The chapter Department Volume and Survival addresses the question of an association between department volume and survival for gynaecological cancer sites. Finally, the chapter Short Summary of Second and Senior Authorship Papers gives a brief overview of the second and senior authorship publications.

RESULTS

Record Linkage B1-1, B1-10

The methodological aspect of record linkage is very important for cancer registries in countries without a unique person identifier, because in many registries patient life status is assessed by applying a passive method.² This means that patient life status results from linking incidence data to mortality data: if the incident patient is found in the mortality file, the patient is treated as dead. Otherwise, it is assumed that the patient is still alive at the close date of the mortality file. When lacking a unique patient identifier, record linkage must be based on various components characterising a person like surname, first name, date of birth, gender and residence. We assume that no single component identifies the person and we also assume it to be likely that some components in the data sources differ for the same person, for example with commonly used surnames with very similar spelling like the German names "Mair", "Maier", "Mayr", or typing errors in date of birth. For this reason, it was necessary to develop a system of record linkage that takes into account possible errors in components and minimises the number of false-positive and false-negative linkage results.

The theory of probabilistic record linkage describes how to calculate the probability that a pair of components identifies the same person.³ Starting from a historical database of record linkage results that contains the pairs of components as well as the correct result (which was identified in time-consuming one-on-one enquiries), we computed the probabilities defined by the theory of probabilistic record linkage for the components surname, first name, date of birth, gender and place of residence.⁴ In addition, to properly deal with typical errors in German language names, we applied a so-called phonetic transformation ("Kölner Transformation")⁵⁻⁶ and we also defined some rules for common typing errors made in dates of birth. The resulting method was validated and we were able to demonstrate adequate precision. However, when applying this probabilistic record linkage method, there are always a certain number of cases that the method classifies as "unclear," and these cases must be resolved one by one by a clerk who tediously collects additional information. Thus, the process is not completely computerized and the overall result also depends on the precision of the individual decisions. One factor influencing the correctness of the final decisions is the number of unclear cases.

As probabilistic record linkage methods were infrequently applied in Austria up to 2005, we analysed in a second paper the errors resulting from deterministic record linkage methods.⁷ A common rule of deterministic record linkage is that a pair of components describes the same person if and only if the components surname and first name and date of birth are identical. Consequently, the proportion of false-negative results is considerable. A false-negative result in the framework of passive assessment of patient life status means that the life status of a patient, who has in fact died, is assessed as alive. A comparison with probabilistic record linkage (as gold standard) resulted in error rates in relative five-year survival of 26% in female and 16% in male patients (relative differences).

Limitations

The main limitations of the record linkage method are the restriction of components to German names and the fact that the historical database used to compute the probabilities for the record linkage method was not developed in a systematic way.

Conclusions

The newly developed method of probabilistic record linkage is widely applied at the CRT and we showed adequate precision resulting from this method.

In the future, two aspects will gain importance. Firstly, the proportion of cancer patients with migration status will increase, bearing in mind that the proportion of immigrants in Austria is between 15% and 20% of the country's population. However, some part of the method is restricted to the German language, namely the phonetic transformation that is widely applied for routine data. Therefore, it will be necessary to upgrade the system of probabilistic record linkage by taking into account characteristics of mainly Turkish and Ex-Yugoslavian names, which reflect the main migrant groups. Secondly, in Austria a unique person identifier may be

introduced in the upcoming years to comply with strict data privacy laws. If such a unique person identifier is introduced and stored in the main data sources used by cancer registries, it will no longer be necessary to apply record linkage methods.

Data Quality Aspects at the Cancer Registry of Tyrol ^{B1-4}

Assessing the correctness of survival rates was only one of the aspects considered when analysing data quality at the CRT.⁸ The design of the comprehensive analysis was based on two overview papers by Max Parkin and Freddie Bray published in 2008.⁹⁻¹⁰ We compared incidence rates with those of neighbouring countries and analysed the time trends of incidence rates, because jumps in the time trend could be an indicator of changes in the completeness of the registration process.¹¹⁻¹² We also analysed quality indicators like proportion of histologically verified cases, mortality-to-incidence ratio, DCO proportion (DCO is an acronym for death certificate only cases, meaning incidence data registered on the basis of information from the mortality database only). Also analysed was the aspect of timeliness, meaning the delay between cancer diagnosis and registration of the incident cancer case. Finally, in order to evaluate the passive method for assessing patient life status,^{4, 7} patient life status was rechecked for every patient diagnosed in 1997 and alive when conducting the study. We rechecked patient life status by actively investigating life status for every patient in the particular municipal office. The year of diagnosis 1997 was chosen to be able to estimate the effect of errors on long-term survival, namely ten-year survival rates.

We re-checked a total of 1026 patients registered as alive when starting the analysis and identified 34 persons who had died. The majority of these cases were due to out-migration (therefore the person cannot be registered in the mortality file of Tyrol). Some very few cases were erroneously assessed as alive during the record linkage process, and a few persons who were still living in Tyrol died outside Austria and were not included in the official mortality file. The effect of these errors on survival rate was small: we observed an overestimation of 0.5% and 1.0% in relative five- and ten-year survival rates, respectively (differences in absolute percent).

17

Limitations

The most severe limitation of this study is that we were not able to investigate all quality indicators proposed in ⁹⁻¹⁰ and we focused on those indicators related to possible bias in survival rates. Furthermore, we have limited data to assess the impact of emigration after cancer diagnosis and the impact of cancer patients in immigrants to Austria.

Conclusions

This comprehensive assessment of data quality aspects demonstrated good quality of incidence data at the CRT. This means that registry procedures needed only minor corrections. However, access to residence registration data to check out-migrant status of patients still alive would minimise errors in survival analysis because out-migration is the main reason for errors in assessing patient life status.

In recent years, more and more structured data on cancer patients has been collected in medical records, for example in clinical cancer registries or in breast cancer units. Therefore, one of the strategic questions for cancer registries will be whether to implement only a link to these data sources or whether to extend epidemiological cancer registries by directly integrating the information in the cancer registry database. The main difference is that in epidemiological cancer registries the cancer registry is fully responsible for data quality, while data used by clinical cancer registries are entered by hospital personnel. If epidemiological and clinical cancer registries are to be integrated, data quality in this new setting has to be re-evaluated.

Breast Cancer B1-5, B1-7, B1-9

In 2006, the Austrian health minister declared mammography to be one of the top health agendas, and in July 2006 the decision was made to implement an organised mammography screening program in a first step in pilot regions, of which Tyrol was the largest. Therefore, in Tyrol there was scientific and public health interest to investigate the time trend of breast cancer mortality. The main question was whether fifteen years of spontaneous mammography screening in Tyrol had already exploited the reduction in breast cancer mortality, as is known from organised programs.¹³ We conducted a historical time trend analysis by applying an ageperiod-cohort model that permits age and period and cohort effects to be analyzed in parallel.¹⁴⁻¹⁵ In the age group 40 to 79, namely the age group targeted by mammography screening in Tyrol, breast cancer mortality in 2006 was reduced by a relative difference of 26% (95% CI 13%-36%), by comparison to 1992, namely before spontaneous mammography screening was introduced. The main problem with this analysis is that without a detailed screening database the effect of screening cannot be assessed properly bearing in mind that mortality is influenced by different factors, especially improvements in therapy. We know from investigations in other countries that up to two-thirds of breast cancer mortality reduction during the 1990s was due to improved therapy.¹⁶ Therefore, we concluded that in Tyrol we could achieve only about half of the mortality reduction that is known to result from highquality screening programs, namely 20% to 25%, for an overview see ¹⁷. Therefore, we recommended that an organised mammography screening program be introduced, in particular a detailed screening database, in order to acquire detailed knowledge on the performance of the screening program and to compare this performance with well-accepted guidelines.

The organised mammography screening program was launched in Tyrol in 2007, starting with a pilot phase from June 2007 to May 2008 in the two central counties covering about 40% of the population. Most EU recommendations for organised mammography screening programs were followed, except double reading. Thus, the next important question was whether the new program reached the quality limits given in the well-accepted EU guidelines.¹⁸ Most quality parameters are analysed directly from the screening database, except for interval cancer cases which are assessed by record linkage between the screening database and the CRT data on breast cancer cases.¹⁹ In routine procedure, CRT publishes incidence data 18 months after the end of the respective year of diagnosis. In order to register breast cancer cases in time to assess interval cancer cases, it was necessary to change the registration process for breast cancer cases to now make them available five months after incidence date. We checked to ensure that the registration process for breast cancer achieves the same degree of

19

completeness as the routine data in CRT, which is essential for a complete assessment of interval cancer cases.

The cumulative participation rate was 35% and 57% in a one- and two-year observation period, respectively. Per 1000 mammograms, 18 were invited for further assessment and 16 for an intermediate screening test in six months. Per 1000 mammograms, nine biopsies were performed and four cancer cases were detected, 10% of which were in situ cases. Of all invasive cancer cases detected in screening, 35% were less than 10 mm in size and 76% were node-negative. In the first year of observation, we assessed six interval cancer cases, or 19% of the background incidence rate.

In a very recent paper, we reported the results of the first year of complete rollout of the mammography screening program to Tyrol.²⁰ From June 2008 to May 2009, 120,440 women were invited. Cumulative participation rate was 57% in two years of observation (with higher participation in younger women). Per 1000 mammograms, 14 were recalled for further assessment, nine biopsies were performed and four cancer cases were detected. Of invasive cancer cases 32% and 68% were smaller than 10 mm and 15 mm in size, respectively. Positive predictive value equalled 39% for core biopsy. Of all cases leading to an invasive cancer, 90% of assessments were performed within five working days after screening and 87% of surgeries were started within ten working days after assessment. Interval cancer rate in the first year of observation was 18% of background incidence rate. We also assessed interval cancer in the second year of observation for the pilot phase and established it to be 55% and 33% of background incidence rate in age groups 40-49 and 50-69, respectively. However, absolute numbers are small with a total of twelve interval cancer cases. Compared with recommended levels according to the EU guidelines (for age group 50-69 only), most indicators were in the recommended range with the exception of participation rate (55% versus the recommended level of 70%) and the proportion of stage II+ cases (33% versus the recommended level of 25%).

Limitations

The main limitation of the first analysis ¹³ was the lack of a screening database at time of publication. The main limitation of the mammography screening program in Tyrol is the lack of double reading. In addition, BI-RADS categories are used in assessing the screening outcome instead of a single yes/no rule and a number of screeners still use BI-RADS 0 (i.e., an unclear result) and BI-RADs 3 (invitation to an intermediate screen within six months). Furthermore, the fact that women aged 40 to 49 are invited is not in line with the EU guidelines.¹⁸

Conclusions

We were able to show that most EU guideline quality limits were met in the pilot phase and in the first year of complete rollout. The main exception is the participation rate at about 55%. Because the Austria-wide introduction of an organised mammography screening program has been announced for 2013, a political decision was made not to change the design of the Tyrol program. The main challenge will be to transform the Tyrol model to the nationwide screening program while maintaining the high quality of the Tyrol program. Analysing performance parameters and setting up a quality assurance group at the state level appears to be a good means of reacting to possible problems in a short time and interacting directly with the screeners. The knowledge acquired in setting up a screening database and evaluating a mammography screening program in Tyrol, which is reflected by high-quality publications, should be utilized in the nationwide program.

Prostate Cancer B1-2, B1-8

Tyrol was one of the first countries worldwide where PSA testing was offered free of charge to all men in the age group 45 to 74, namely since the beginning of the 1990s.²¹ Therefore, the overriding question was whether offering PSA testing brought about a reduction in prostate cancer mortality. It is worthwhile mentioning that PSA testing was not offered in the framework of a screening program, no invitation system was implemented and no countrywide screening database was set up. This is the reason why we use the wording "PSA testing" and not "PSA screening." We analysed the time trend for prostate cancer mortality in Tyrol and Austria except Tyrol based on the official prostate cancer mortality data, which are registered by Statistics Austria for all of Austria in a uniform way.²² An age-period-cohort model was applied that allows parallel modelling of age, time and cohort effects, see also the chapter Breast Cancer.¹⁴⁻¹⁵ Two comparisons were conducted, firstly a historical comparison within Tyrol, secondly a comparison between Tyrol and Austria except Tyrol. Prostate cancer is diagnosed and treated in a rather uniform way throughout Austria, the only main difference being that PSA testing started in Tyrol at the beginning of and in Austria except Tyrol in the late 1990s. In Tyrol, compared to 1989-93, we observed a 19% reduction (95% CI 2%-32%) in prostate cancer mortality in 1999-2003 (reaching statistical significance) and a constant time trend in Austria except Tyrol. An estimation of the cumulative attendance rate in Tyrol showed that about 75% of men aged 45 to 74 had at least one PSA test between 1993 and 2001. This estimation is based on data from PSA labs and called for some very strong assumptions, for example concerning identification of persons.

Five years later, in 2010, the relevant question was whether we could affirm the reduction in prostate cancer mortality.²³ In fact, by applying the same methodology, we were able to show a 30% (95% CI 13%-43%) reduction in prostate cancer mortality (reaching statistical significance) in Tyrol as compared to 1989-93 and an 8% reduction in Austria except Tyrol, both reductions in relative percents. The 30% reduction in Tyrol is consistent with other publications, see for example ²⁴⁻²⁶. After the publication of two large randomised trials in 2009,²⁷⁻²⁸ a number of detailed studies were published mainly in Europe dealing with attendance and non-compliance problems in the randomised trials and resulting in an average 30% reduction in prostate cancer mortality as a consequence of PSA screening.²⁴⁻²⁶ In 2010, the results of the Göteborg randomised trial showed an even larger reduction, namely 56%.²⁹

Limitations

The main limitation of this analysis is that nonrandomized studies are prone to several biases. Our study design is an observational one that does not allow for appropriate control of confounders. Therefore, any causal interpretations must be done with caution. In addition, our outcome is prostate cancer mortality and we have no validation of prostate cancer as cause of death. We also lack detailed knowledge of the volume of PSA testing, and finally, to date we have only very limited data on harms caused by PSA testing.

Our analysis could not overcome the problems of non-randomised trials, but it was able to provide further information on the mortality reduction resulting from PSA testing or screening.

Conclusions

A number of investigations showed a consistent approximately 30% reduction in prostate cancer mortality following PSA screening or testing. Now, the most relevant question is the balance between mortality reduction and harms of PSA screening. The harms are serious, ranging from a possible reduction in quality of life as a consequence of attending PSA screening to questions of over-diagnosis and over-treatment and serious side-effects of treatment like incontinence and impotence.

In addition to these public health questions, research groups worldwide are looking to improve PSA tests with the aim of differentiating between aggressive and slow-growing prostate cancer.

Department Volume and Cancer Patient Survival B1-3

The question as to a possible association between department volume and outcome, especially survival of cancer patients, has been studied for two decades.³⁰ Overall, it can be said that there is some agreement that large departments show better outcome if the therapy modality is complex. Therefore, we investigated whether this general association also holds in the framework of hospitals in Tyrol, namely for breast, ovarian, cervical and corpus cancer diagnosed between 1988 and 2003.³¹ Based on the observed number of patients treated per year, we defined small departments as treating fewer than 12 patients per year and large departments more than 36 patients per year and, due to small numbers, more than 24 patients for ovarian and corpus cancer. The rationale behind this definition was the rule of thumb "one patient per month." This definition was fixed a priori. The analysis was conducted with a multivariate Cox regression model; covariates were age at diagnosis, year of diagnosis, histological confirmation, cancer stage according to UICC and transfer to another hospital.³² After building a model for every cancer site following a backward strategy, we calculated multivariate hazard ratios of 1.39 (95% CI 1.22-1.58) and 1.27 (95% CI 1.05-1.54) for breast and ovarian cancer, respectively, each for small departments as compared to large departments. Both results reached statistical significance. For cervical cancer, we established a hazard ratio of 0.67 (95% CI 0.51-0.88), also reaching statistical significance.

When comparing the results with those published in the literature, see for example ³³⁻³⁷, we must remember that most publications set the definition for large departments much higher, namely at 100 cases per year or more.

Limitations

The main limitation of this analysis is the lack of precision in defining which department is responsible for initial treatment and we have only limited information on transfer of patients. In addition, we may be confronted with residual confounding because the information on inter-departmental differences in patient characteristics is limited.

Conclusions

Investigations into an association between department volume and cancer survival have been widely criticized, because there is great doubt as to the causality. In actual fact, residual confounding could play a role in the multivariate analysis. On the other hand, these associations correlate strongly with economic and health policy questions. Moreover, a look at the discussion surrounding, for example minimum caseload for breast cancer centers, gives an idea of the complexity of the decision process. It is our belief that the question of department or hospital volume is of great public health importance, but research is still needed on how to integrate scientific results into final health policy decisions.

Gender and Survival in Oncology ^{B1-6}

Since the 1990s research groups all over the world have investigated whether there is a difference in survival between female and male cancer patients, see for example ³⁸. For this reason, we explored this question using the CRT incidence database for 1988 to 2003 with patient life status up to end of 2006.³⁹ The analysis was based on a model for relative excess risk ⁴⁰ taking further life expectancy into account and adjusting for age, stage, period of diagnosis, histological confirmation and interaction between follow-up and stage. For all cancer sites combined, after adjusting for site mix, we calculated an excess risk of 0.95 (95% CI 0.91-0.99) for females as compared to males, reaching statistical significance. However, this excess risk is age-dependent and for patients age ≥80 changes to an excess risk of 1.10 (95% 0.98-1.22) at borderline statistical significance. A statistically significant reduced excess risk for females as compared to males was shown for head and neck cancers except larynx cancer (0.72, 95% CI 0.56-0.93), for stomach cancer (0.86, 95% CI 0.75-0.95) and for lung cancer (0.82, 95% CI 0.75-0.90). The results are consistent with the literature, see for example ³⁸, and to our knowledge this is the first analysis adjusting for staging distribution. The reasons for these differences are mainly unknown and many researchers are convinced that a better understanding of the underlying risks could help improve therapy for cancer patients.³⁸

Limitations

The main limitation of this analysis is the potential for residual confounding in this observational setting. In addition, cancer registry data include only limited data for controlling information bias. Whether gender has a direct effect on survival, whether the effect is mediated in a classical way by tumour stage or whether the effect is influenced by some unknown factor interacting with tumour stage and with survival needs to be discussed, see for example ⁴¹.

Conclusions

The fact that for patients up to the age of 80, female cancer patients have a reduced excess mortality compared to males is consistent with the literature. A very interesting aspect is the

age dependency of the excess risk. It was shown recently that the difference in survival between younger and elderly patients in Europe was on the increase during the 1990es. This imposes consequences for geriatric oncology bearing in mind that 40% and 17% of incident cancer patients in Tyrol are in age group \geq 70 and \geq 80 respectively. Due to the age shift in population towards the elderly, these proportions will increase in the future.

Short Summary of Second and Senior Authorship Papers B2-1, B2-2, B2-3

This chapter briefly summarises the work published in the three second and senior authorship papers: two of them are applications of biostatistical methods, one covers an epidemiologic registry topic.

The first paper entitled "Allogeneic or autologous stem cell transplantation (STC) for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma: a single-centre experience"^{B2-1, 42} investigated the question how patients with relapsed and/or refractory lymphoma might benefit most from stem cell transplantation. Patients who received a stem cell graft at the Division of Hematology and Oncology of Innsbruck University Hospital were analysed by means of a multivariate Cox regression model. We were able to show that allogeneic STC was superior to autologous SCT only in patients with elevated serum LDH levels and bone marrow involvement.

The second paper entitled "Long-term results of dose density therapy in patients with aggressive lymphoma"^{B2-3, 43} investigated the long-term outcome of patients with untreated aggressive lymphoma. Patients from two trials were analysed; this was a phase 2 trial assessing the feasibility, toxicity and efficacy of a new CEOP/IMVP-Dexa regimen and a randomised trial comparing CEOP/IMVP-Dexa versus CHOP. All patients had a previously untreated aggressive lymphoma. Observation time was eight years in the median to a maximum of thirteen years. Overall survival, time to treatment failure and time to relapse were estimated by applying the Kaplan-Meier method. Estimated cumulative overall survival at eight years after diagnosis was 0.71 and 0.30 for patients age < 60 and ≥ 60 years, respectively. In summary, we were able to

demonstrate that the new CEOP/IMVP-Dexa regimen affords excellent survival for patients in the age group up to 60.

The last paper entitled "Regional variability in the incidence of end-stage renal disease: an epidemiological approach"^{B2-2, 44} deals with the question of low incidence of end-stage renal disease in Tyrol as compared to other parts of Austria. Our first finding was that patients with type 2 diabetes mellitus (DM-2) are the only subgroup with lower incidence of end-stage renal disease in Tyrol as compared to Austria except Tyrol. The next question was whether in Tyrol, the prevalence of DM-2 is lower than in Austria. Because in 2003 there was no diabetes registry in Austria, the approach was to estimate DM-2 prevalence using several surrogate parameters derived from Health Interview Surveys conducted in 1991 and 1995, the National Hospital Discharge Registry in Austria, the National Mortality Registry and the National Drug Wholesale Registry. All available data sources supported low DM2 prevalence rates in Tyrol as compared to Austria except Tyrol. However, this investigation is subject to several limitations.

OVERALL SUMMARY AND CONCLUSION

In summary, we were able to address very relevant public health questions in Tyrol and to address the methodological aspects that are important in a cancer registry.

The newly developed method of probabilistic record linkage is widely applied, and we were able to show adequate precision as a result of applying this method. The comprehensive assessment of data quality aspects demonstrated good quality of incidence data in the CRT. We were able to show that the newly established organised mammography screening program in Tyrol, which was launched in 2007, meets most quality limits given in the well established EU guidelines. We also were able to show a 30% reduction in prostate cancer mortality following the introduction of prostate-specific antigen testing offered free of charge to all men aged 45 to 74. Finally, we showed an association between department volume and cancer survival in Tyrol for breast and ovarian cancer and we demonstrated that female cancer patients in Tyrol up to the age of 80 have reduced excess mortality as compared to males. Nevertheless, due to the observational setting and the limited data of registries, all such studies are limited by the potential of residual confounding and information bias.

Important research questions for the future are how to make use of cancer registry data for systematic quality assessment in oncology. This research question addresses both documentation aspects and methodological challenges bearing in mind the mean number of incident cancer cases treated in oncologic departments in Tyrol. Secondly, many questions regarding cancer patient survival in Tyrol are to be discussed. One key issue will be to refine methods for analysing survival for medium and small size cancer registries like the CRT. And thirdly, maybe the most important issue is the transition from results achieved by epidemiological research to decisions made in the public health area. A main goal should be a systematic approach to evaluate risks, benefits and costs in selected areas in oncology within the framework of comprehensive decision models. Cancer registry data are one important component in the development of decision models should be intensified. As a consequence, the network between medium and small size cancer registries should be intensified and the cooperation between cancer registries and interdisciplinary research teams working on decision models must be promoted.

REFERENCES

- World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. [http://www.wma.net/en/30publications/10policies/b3/index.html] (acessed 7 November 2009).
- Capocaccia R, Gatta G, Roazzi P, Carrani E, Santaquilani M, De Angelis R, et al. The EUROCARE-3 database: methodology of data collection, standardisation, quality control and statistical analysis. Ann Oncol. 2003;14 Suppl 5:v14-27.
- Jaro MA. Probabilistic linkage of large public health data files. Stat Med. 1995;14(5-7):491-8.
- 4. Oberaigner W, Stuhlinger W. Record linkage in the Cancer Registry of Tyrol, Austria. Methods Inf Med. 2005;44(5):626-30.
- 5. Schmidtmann I, Michaelis J. Untersuchungen zum Record Linkage für das Krebsregister Rheinland-Pfalz. Mainz: Techn. Bericht Tumorzentrum Rheinland-Pfalz; 1994.
- Stegmaier C, Ziegler H. Zusammenführung personenbezogener Informationen (Record-Linkage) - Probleme und Möglichkeiten am Beispiel des Krebsregisters Saarland. Krebsregister Saarland Vierteljahresheft. 1992;4:1-7.
- 7. Oberaigner W. Errors in survival rates caused by routinely used deterministic record linkage methods. Methods Inf Med. 2007;46(4):420-4.
- 8. Oberaigner W, Siebert U. Are survival rates for Tyrol published in the Eurocare studies biased? Acta Oncol. 2009;48(7):984-91.
- 9. Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009;45(5):747-55.
- 10. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. Eur J Cancer. 2009;45(5):756-64.
- 11. Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B. Cancer Incidence in Five Continents. Volume VIII. Lyon: IARC Sci Publ; 2002.
- 12. Curado MP, Edwards B, Shin H, Ferlay J, Heanue M, Boyle P, et al., editors. Cancer incidence in five continents. Volume IX. Lyon: IARC Sci Publ; 2008.
- Oberaigner W, Buchberger W, Frede T, Knapp R, Marth C, Siebert U. Breast cancer incidence and mortality in Tyrol/Austria after fifteen years of opportunistic mammography screening. BMC Public Health. 2010;10:86.

- 14. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: Age-period and age-cohort models. Stat Med. 1987;6(4):449-67.
- 15. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. Stat Med. 1987;6(4):469-81.
- 16. Peto R, Boreham J, Clarke M, Davies C, Beral V. UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years. Lancet. 2000;355(9217):1822.
- 17. Gotzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database Syst Rev. 2011(1):CD001877.
- 18. Perry N, Broeders M, de Wolf C (Eds.). European guidelines for quality assurance in breast cancer screening and diagnosis. Luxembourg: Office for Official Publications of the European Communities; 2006.
- 19. Oberaigner W, Buchberger W, Frede T, Daniaux M, Knapp R, Marth C, et al. Introduction of organised mammography screening in tyrol: results of a one-year pilot phase. BMC Public Health. 2011;11:91.
- Oberaigner W, Daniaux M, Geiger-Gritsch S, Knapp R, Siebert U, Buchberger W. Introduction of organised mammography screening in Tyrol: results following first year of complete rollout. BMC Public Health. 2011;11:673.
- 21. Bartsch G, Horninger W, Klocker H, Reissigl A, Oberaigner W, Schonitzer D, et al. Prostate cancer mortality after introduction of prostate-specific antigen mass screening in the Federal State of Tyrol, Austria. Urology. 2001;58(3):417-24.
- 22. Oberaigner W, Horninger W, Klocker H, Schonitzer D, Stuhlinger W, Bartsch G. Reduction of prostate cancer mortality in Tyrol, Austria, after introduction of prostate-specific antigen testing. Am J Epidemiol. 2006;164(4):376-84.
- Oberaigner W, Siebert U, Horninger W, Klocker H, Bektic J, Schafer G, et al. Prostatespecific antigen testing in Tyrol, Austria: prostate cancer mortality reduction was supported by an update with mortality data up to 2008. Int J Public Health. 2011 Jun 17. DOI 10.1007/s00038-011-0266-4.
- Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FH, et al. Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trial (ERSPC section Rotterdam). Int J Cancer. 2010;127(11):2639-44.
- 25. Roobol MJ, Kerkhof M, Schroder FH, Cuzick J, Sasieni P, Hakama M, et al. Prostate cancer mortality reduction by prostate-specific antigen-based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). Eur Urol. 2009;56(4):584-91.

- 26. van Leeuwen PJ, Connolly D, Gavin A, Roobol MJ, Black A, Bangma CH, et al. Prostate cancer mortality in screen and clinically detected prostate cancer: estimating the screening benefit. Eur J Cancer. 2010;46(2):377-83.
- Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009;360(13):1310-9.
- Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med. 2009;360(13):1320-8.
- 29. Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. Lancet Oncol. 2010;11(8):725-32.
- Hillner BE, Smith TJ, Desch CE. Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. J Clin Oncol. 2000;18(11):2327-40.
- Oberaigner W, Stuhlinger W. Influence of department volume on cancer survival for gynaecological cancers--a population-based study in Tyrol, Austria. Gynecol Oncol. 2006;103(2):527-34.
- 32. Cleves M, Gould W, Gutierrez R. An Introduction to Survival Analysis Using Stata. Revised Edition. Stata Press; 2004.
- 33. Skinner KA, Helsper JT, Deapen D, Ye W, Sposto R. Breast cancer: do specialists make a difference? Ann Surg Oncol. 2003;10(6):606-15.
- 34. Calman K, Hine D. Policy Framework for Commissioning Cancer Services: a Report by the Expert Advisory Group on Cancer to the Chief Medical Officers of England and Wales: Department of Health. London: HM Stationary Office; 1995.
- 35. Improving Outcomes in Breast Cancer: Guidance for Purchasers. London: Department of Health; 1996.
- 36. Elit L, Bondy SJ, Paszat L, Przybysz R, Levine M. Outcomes in surgery for ovarian cancer. Gynecol Oncol. 2002;87(3):260-7.
- 37. Kumpulainen S, Grenman S, Kyyronen P, Pukkala E, Sankila R. Evidence of benefit from centralised treatment of ovarian cancer: a nationwide population-based survival analysis in Finland. Int J Cancer. 2002;102(5):541-4.
- Micheli A, Ciampichini R, Oberaigner W, Ciccolallo L, de Vries E, Izarzugaza I, et al. The advantage of women in cancer survival: an analysis of EUROCARE-4 data. Eur J Cancer. 2009;45(6):1017-27.

- 39. Oberaigner W, Siebert U. Do women with cancer have better survival as compared to men after adjusting for staging distribution? Eur J Public Health. 2010;21(3):387-91.
- 40. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. Stat Med. 2004;23(1):51-64.
- 41. Cole SR, Hernan MA. Fallibility in estimating direct effects. Int J Epidemiol. 2002;31(1):163-5.
- Nachbaur D, Oberaigner W, Fritsch E, Nussbaumer W, Gastl G. Allogeneic or autologous stem cell transplantation (SCT) for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma: a single-centre experience. Eur J Haematol. 2001;66(1):43-9.
- 43. Fridrik MA, Hausmaninger H, Linkesch W, Greil R, Krieger O, Baldinger C, et al. Long-term results of dose density therapy in patients with aggressive lymphoma. Ann Hematol. 2005;84(4):217-22.
- 44. Wimmer F, Oberaigner W, Kramar R, Mayer G. Regional variability in the incidence of endstage renal disease: an epidemiological approach. Nephrol Dial Transplant. 2003;18(8):1562-7.

LIST OF FIRST AUTHORSHIP PAPERS

No.	Citation	Remarks	Contribution
B1-1	Oberaigner W , Stuhlinger W. Record linkage in the Cancer Registry of Tyrol, Austria. Methods Inf Med. 2005;44(5):626-30.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-2	Oberaigner W , Horninger W, Klocker H, Schonitzer D, Stuhlinger W, Bartsch G. Reduction of prostate cancer mortality in Tyrol, Austria, after introduction of prostate-specific antigen testing. Am J Epidemiol. 2006 Aug 15;164(4):376-84.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-3	Oberaigner W , Stuhlinger W. Influence of department volume on cancer survival for gynaecological cancersa population- based study in Tyrol, Austria. Gynecol Oncol. 2006 Nov;103(2):527-34.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-4	Oberaigner W , Siebert U. Are survival rates for Tyrol published in the Eurocare studies biased? Acta Oncol. 2009;48(7):984-91.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-5	Oberaigner W , Buchberger W, Frede T, Knapp R, Marth C, Siebert U. Breast cancer incidence and mortality in Tyrol/Austria after fifteen years of opportunistic mammography screening. BMC Public Health. 2010;10:86.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper

B1-6	Oberaigner W , Siebert U. Do women with cancer have better survival as compared to men after adjusting for staging distribution? Eur J Public Health. 2010;21(3):387-91.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-7	Oberaigner W , Buchberger W, Frede T, Daniaux M, Knapp R, Marth C, Siebert U. Introduction of organised mammography screening in Tyrol: results of a one-year pilot phase. BMC Public Health 2011, 11:91	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-8	Oberaigner W , Siebert U, Horninger W,Klocker H, Bektic J, Schäfer G, Frauscher F, Schennach H, Bartsch G. Prostate-specific antigen testing in Tyrol, Austria: prostate cancer mortality reduction was supported by an update with mortality data up to 2008. Int J Public Health 2011. DOI 10.1007/s00038-011-0266-4.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-9	Oberaigner W , Daniaux M, Geiger- Gritsch S, Knapp R, Siebert U, Buchberger W. Introduction of organised mammography screening in Tyrol: results following first year of complete rollout. BMC Public Health 2011, 11:673.	Full paper, peer-reviewed, Medline, English With UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
B1-10	Oberaigner W . Errors in survival rates caused by routinely used deterministic record linkage methods. Methods Inf Med. 2007;46(4):420-4.	Full paper, peer-reviewed, Medline, English Without UMIT affiliation	Framing the research question, development of study design, selection of biostatistical methods, performance of statistical analysis, writing of paper
LIST OF SECOND AND SENIOR AUTHORSHIP PAPERS

No.	Citation	Remarks	Contribution
B2-1	Nachbaur D, Oberaigner W , Fritsch E, Nussbaumer W, Gastl G. Allogeneic or autologous stem cell transplantation (SCT) for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma: a single-centre experience. Eur J Haematol. 2001 Jan;66(1):43-9.	Full paper, peer-reviewed, Medline, English Without UMIT affiliation	Performance of analyses, contributing to study design, critical review and comments on manuscript
B2-2	Wimmer F, Oberaigner W , Kramar R, Mayer G. Regional variability in the incidence of end-stage renal disease: an epidemiological approach. Nephrol Dial Transplant. 2003 Aug;18(8):1562- 7.	Full paper, peer-reviewed, Medline, English Without UMIT affiliation	Performance of analyses, contributing to study design, critical review and comments on manuscript
B2-3	Fridrik MA, Hausmaninger H, Linkesch W, Greil R, Krieger O, Baldinger C, Klocker J, Pont J, Oberaigner W . Long- term results of dose density therapy in patients with aggressive lymphoma. Ann Hematol. 2005 Apr;84(4):217-22.	Full paper, peer-reviewed, Medline, English Without UMIT affiliation	Performance of analyses, contributing to study design, critical review and comments on manuscript

ACKNOWLEDGEMENTS

This *habilitation* thesis is based on publications with many coauthors.

My very first thanks go to all coauthors for the contributions to these papers and for their stimulating discussion.

A special word of thanks goes to the many medical doctors and nurses in Tyrol who contribute so much to the incidence data collected by the Cancer Registry of Tyrol.

I am also indebted to all my IET colleagues, in particular

- Helmut Mühlböck and Christine Wartha, who bear the brunt of the hard work involved in the registration process and who deserve praise for the excellent quality of the incidence data over decades,
- Patricia Gscheidlinger, who assisted with typing and paper layout,
- Lois Harrasser for his excellent programming work and his mastery of STATA.

Thanks to Mary Margreiter for native-speaker editing of the English.

My very special thanks also go to important academic supporters,

- Prof. Wolfgang Gröbner (who died in 1980) for encouraging my early scientific work. Although this first work in algebraic geometry is not directly reflected in this *habilitation* thesis, many of his thoughts are still valid and stimulating,
- Prof. Albrecht Neiß, for offering me the opportunity to enter the field of epidemiology and set up a cancer registry from the outset,
- Prof. Wolfgang Buchberger for encouraging my scientific work on mammography screening,
- Prof. Uwe Siebert for promoting my scientific work in recent years.

Finally, a warm and heartfelt thank you to my family, who gave me the energy, time, calm, and everything else I needed to complete my *habilitation* process!

APPENDIX: DECLARATION

I herewith declare that I have never before submitted an application for *habilitation*, or that any other *habilitation* proceeding is pending.

I further declare that the publications listed in the cumulative *habilitation* thesis were written by me as first author or second author or last author and that I have used no other sources or resources than the ones cited.

The sources used are stated in accordance with the pertinent rules and regulations. I have indicated all quotes and citations that were literally taken from publications or that were in close accordance with the meaning of those publications. All sources and other resources used are stated in the bibliography.

November 2011, Hall i.T.

Dr. phil. Wilhelm Oberaigner

626

© 2005

Schattauer GmbH

Record Linkage in the Cancer Registry of Tyrol, Austria

W. Oberaigner^{1, 2}, W. Stühlinger²

¹Cancer Registry of Tyrol, Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Innsbruck, Austria

²Department of Quality Science, Medical Planning and Information Management, University for Health Informatics and Technology Tyrol, Innsbruck, Austria

Summary

Objective: Record linkage of patient data originating from various data sources and record linkage for checking uniqueness of patient registration are common tasks for every cancer registry. In Austria, there is no unique person identifier in use in the medical system. Hence, it was necessary and the goal of this work to develop an efficient means of record linkage for use in cancer registries in Austria.

Methods: We adapted the method of probabilistic record linkage to the situation of cancer registries in Austria. In addition to the customary components of this method, we also took into consideration typing errors commonly occurring in names and dates of birth. The method was implemented in a program written in DELPHITM with interfaces optimised for cancer registries.

Results: Applying our record linkage method to 130,509 linkages results in 105,272 (80.7%) identical pairs. For these identical pairs, 88.9% of decisions were performed automatically and 11.1% semi-automatically. For results decided automatically, 6.9% did not have simultaneous identity of last name, first name and date of birth. For results decided semi-automatically, 48.4% did not have an identical last name, 25.6% did not have an identical date of birth and 83.1% did not have simultaneous identity of last name and date of birth and first name. **Conclusions:** The method implemented in our cancer registry solves all record linkage problems in Austria with sufficient precision.

Keywords

Probabilistic record linkage, cancer registry, homonym rate, synonym rate

Methods Inf Med 2005; 44: 626-30

Introduction

The prime objective of population-based cancer registries is to document every incident of cancer cases diagnosed in the target population [1-5]. According to international guidelines, a cancer registry should take into account various data sources containing valid information on cancer cases. Consequently, in addition to data sent to the registry by treating physicians, data sources like pathology reports, department information systems (i.e. radiotherapy) and hospital information systems must be included in the registration process. Many cancer registries analyze survival rates as the most important outcome measure, and for this analysis patient life status has to be assessed. Most registries apply a passive method, meaning record linkage between incidence data and mortality data [6].

Summing up, record linkage is a central task to be solved by cancer registries. In Austria, there is no general use of unique person identifiers as, for example, in Scandinavian countries. There is a social insurance number that is known to not be unique in all cases and it is not widely used in medical information systems. Therefore, the decision on whether data describe the same person must be based on information like last name, first name, date of birth etc. and can be time-consuming when a high degree of precision is involved. All registries aim to obtain complete and reliable information needed for patient identification, but it must be remembered that in actual practice all the components mentioned above can be distorted by (registration as well as typing) errors.

Administrative workflow in cancer registries differs in some respect from that in administrative units in hospital departments. In contrast to hospital administration, in cancer registries there is no need to register patient data immediately. Since cancer registries collect data mostly on the basis of year of diagnosis, their data collection efforts are more thorough and generally ensure good quality of data needed for record linkage.

In order to develop an efficient, scientifically founded method for record linkage, we decided some years ago to implement a method based on the theory of probabilistic record linkage and taking into account common types of error sources in the German language.

Methods

Basics

This chapter presents the basics of the theory of probabilistic record linkage to the extent needed to understand the method developed for our cancer registry. Detailed descriptions of the theory can be found for example in [7, 8].

Data in a cancer registry consist of several components describing an individual person or cancer case. One part of these components, often called person data, identifies the person. We assume that no single component uniquely identifies a person.

If a person is described by n components k_1 to k_n , we assign standardized weights to each component, i.e. w_1 to w_n , where $w_1 + ... + w_n = 1$.

Downloaded from www.methods-online.com on 2011-11-30 | ID: 1000466760 | IP: 193.171.69.65 For personal or educational use only. No other uses without permission. All rights reserved ved: February 9, 2004; accepted: December 20, 2004 Appendix 1



Fig.1 Decision process

For linkage of two records with components k_i^{1} and k_i^{2} we define p_i for each component k_i as follows:

$$p_i = \begin{cases} 1 & \text{if } k_i^{\ l} = k_i^{\ 2} \\ 0 & \text{otherwise} \end{cases}$$
(1)

This gives a sum probability defined as

$$p = w_1 p_1 + ... + w_n p_n$$
 (2)

p (in the following often denoted by p probability) can be interpreted as a measure of whether two records describe the same person. Then, two cut points p^1 and p^2 are introduced with the following consequences:

- If p is smaller than p¹, it is assumed (without further checks) that the records describe different persons.
- If p is greater than p², it is assumed (again without further checks) that the records describe the same person.
- If p lies between p¹ and p², it must be decided on an individual basis whether the two records describe the same person or different persons. Usually, this means further information must be obtained.

The decision process is shown in Figure 1.

Choice of Weights

In order to choose weights according to the theory of probabilistic record linkage, two probabilities are computed, usually denoted as m and u probability.

For any component k_i , m_i is defined as the probability that k_i is equal for identical persons. u_i describes the probability that k_i is equal for non-identical persons. The weight w_i is then defined by the following formula:

$$v_i = \log_2(\frac{m_i}{u_i})$$

Table 1

Transformations according to the "Kölner Transformation"

Rule	Example
Eliminate diphthongs	Wimmer \rightarrow WIMER
Transform German "Umlaute"	Müller \rightarrow MUELER
Transform "c " in front of "e, i " to "z "	$Cicero \rightarrow ZIZERO$
Transform "c " in front of "a, o, u " to "k "	$Cugel \rightarrow KUGEL$
Otherwise transform "c " to "z "	Mucke \rightarrow MUZKE
Transform "v" to "f"	$Vogel \rightarrow FOGEL$
Transform "j" to "i"	$Deljc \to DELIZ$
Transform "ie" to "i"	Liederlich \rightarrow LIDERLIZH
Transform "ai" to "ei"	Aigner \rightarrow EIGNER
Transform "ae" to "e"	Jaeger \rightarrow IEGER
Transform "th" to "t"	Thaler \rightarrow TALER
Transform "tz" to "z"	Matzer \rightarrow MAZER
Transform "d" to "t"	Danner \rightarrow TANER
Delete silent "h"	Gehler \rightarrow GELER
Transform "qu" to "q"	Qualler \rightarrow QALER

From the experience in our cancer registry the components were chosen as follows [9]:

- Last name
- Phonetic transformation of last name (used only if last name is not identical for the two persons under investigation), see Table 1.
- Birth name
- Phonetic transformation of birth name (used only if birth name is not identical for the two persons under investigation), see Table 1.
- First name
- Date of birth
- Sex

(3)

• Zip code (or municipality code)

The German language contains typical transformations of names following certain rules. We thus introduced the concept of phonetic transformation defined by the rules given in Table 1 (derived from the so-called Kölner Transformation, see [10, 11]).

The probabilities m_i and u_i were calculated based on results obtained before introducing the method described here, when we performed record linkage by heuristic methods and individual checks. All results were stored in a meta-relation describing pairs of data to be linked as well as linkage results. Based on this relation, it is straight-

Table 2 Standardized weight for components

Component k _i	w _i (standardized)
Phonetic transformation last name	0.22
Phonetic transformation birth name	0.202
First name	0.139
Date of birth	0.289
Sex	0.075
Zip code (or municipality code)	0.075

forward to compute the probability m_i as follows:

$$m_i = \frac{\text{number of patients}}{\text{number of patients}} \tag{4}$$

In the same way, we can compute the probability u_i as follows (we assume that every patient in our database is unique, hence the Cartesian product Pat × Pat (denoting all possible combinations of patients) does not contain pairs of equal patients):

$$u_{i} = \frac{number of patients}{number of action different} (5)$$

$$u_{i} = \frac{with identical component k_{i}}{number of different}$$
pairs of patients

These computations gave the weights shown in Table 2.

Downloaded from www.methods-online.com on 2011-11-30 | ID: 1000466760 | IP: 193.171.69.65

For personal or educational use only. No other uses without permission. All rights reserved.

v

Table 3 Additional methods

Method	Example
Left part or right part of name identical	Müller and Müller-Westernhagen
1 character wrong	Maier and Mayer
1 character missing	Maier and Mair
2 neighboring characters exchanged	Maier and Miaer

Table 4 Correction factors for weights

Component or method for component	Weight
Last name: left or right part identical	w _{last name} *0.9
Last name: 1 character wrong	w _{last name} *0.8
Last name: 1 character missing	W _{last name} *0.8
Last name: 2 characters exchanged	w _{last name} *0.8
First three digits of last name identical	w _{last name} *0.4
First name: left or right part identical	w _{first name} *0.5
Last name and birth name exchanged	W _{last name}
Date of birth: 1 character wrong	w _{date of birth} *0.8
Date of birth: 2 characters exchanged	w _{date of birth} *0.8
Date of birth: day and month exchanged	W _{date of birth} *0.8
Date of birth: day identical	w _{date of birth} *0.3
Date of birth: month identical	W _{date of birth} *0.3
Date of birth: year identical	W _{date of birth} *0.3

After detailed analysis of our database and after investigating typical errors occurring in our registry, we found that our registry contains [9] common typing errors in last name and birth name and common typing errors in date of birth.

In order to properly deal with these errors, we added the methods described in Table 3 to the components described above and consequently extended the weights given in Table 2 by the weights defined in Table 4.

For every component k_i the maximum weight is limited by the weight for this component as defined in Table 2, even if all methods defined in Table 4 add up to a greater weight.

Table 5 Results of evaluation for years 1999-2003

Number of linkages	130,509	
Identical pairs	105,272 (80.7%)	
Decision automatic	93,627 (88.9%)	
semiautomatic	11,645 (11.1%)	
Applied rules	Decision automatic	Decision semiautomatic
Last name identical	91,835 (98.1%)	6,014 (51.6%)
Phonetic transformation of last name identical	1,692 (1.8%)	157 (1.3%
First name identical	88,949 (95%)	9,926 (85.2%)
Date of birth identical	93,627 (100%)	8,662 (74.4%)
Sex identical	91,750 (98%)	10,872 (93.3%)
Last name AND date of birth identical	91,835 (98.1%)	3,159 (27.1%)
Last name AND date of birth AND first name identical	87,160 (93.1%)	1,970 (16.9%)
One character rules (see Table 3) apply for last name	0	4,701 (40.4%)

Choice of Critical Bounds p¹ and p²

Our experience shows that $p^1 = 75$ and $p^2 = 95$ are good choices for cancer registries in Austria. This means that we inspect all cases with a p probability between 75 and 95 and assume without further inspection that pairs with $p \in (95, 100]$ describe the same person.

Inspection of all pairs with $p \in [75, 95]$ is a very time-consuming and tedious job. Scanning through the lists requires a great deal of concentration. However, there are usually some pairs describing the same person but with a smaller p probability (think, for example, of twins living in the same residence, perhaps with similar first names). Hence, in order to keep homonym and synonym rates low (see also the discussion on the consequences of wrong decisions) it is necessary to run through all parts of the resulting list with full concentration.

Implementation

The method described above was implemented as a program written in DELPHI. Interfaces for input are either plain text files with fields separated by "\", or Oracle tables (our cancer registry database is implemented in OracleTM). Results are written both in a plain text file and in an Oracle table. Output in either format can be imported for further analysis to any statistical package and contains original data as well as p probability (see equation (2)) and information on the rules applied. Pairs of data with p probability less than 70 are not included in the output. This information allows us to also do detailed analyses of the method.

The DELPHI program first transforms all names according to the Kölner Transformation and implements the methods defined in Tables 3 and 4. When comparing one person against 100,000 persons the program needs about two seconds on a common PC. The resulting computing times are acceptable for our typical projects. Therefore, we did not implement blocking techniques, which are known to reduce computing time by a quadratic factor [8].

The program runs well in practice and has proven advantages with regard to simplicity of interface and interpretation of results. From the point of view of our cancer registry its main advantage is that it takes into account typing errors that derive from the language used, thus here restricted to the German language.

Results

The program described above is applied in the Cancer Registry of Tyrol to join various

Downloaded from www.methods-online.com on 2011-11-30 | ID: 1000466760 | IP: 193.171.69.65 For personal or educational use only. No other uses without permission. All rights reserved.

629

data sources and check duplicates in the incidence database. Table 5 describes the main results for all linkages done in the years 1999 to 2003. A total of 130,509 linkages were conducted, of which 105,272 (80.7%) were identical pairs. Of these identical pairs, 88.9% of decisions were performed automatically and 11.1% semi-automatically (meaning they were made by the clerical staff).

For results decided automatically, 98.1% had identical last name and 1.8% had identical phonetic transformation of the last name; 95% of cases had identical first name and all cases had identical date of birth. Simultaneous identity of last name and date of birth and first name was observed for 93.1%.

For results decided semi-automatically, 51.6% had identical last name and 1.3% identical phonetic transformation of the last name. First name was identical in 85.2% of cases and date of birth was identical in 74.4%. Simultaneous identity of last name and date of birth and first name was observed for 16.9%. One-character rules (defined in Table 3) applied to last name for 40.4%.

Discussion

Choice of Critical Bounds

We use this program for two main purposes, namely for linking two different data sources and for identification of persons registered more than once in the database.

One of the key decisions during implementation was to choose specific values for the critical bounds p^1 and p^2 . In order to evaluate this decision, one must bear in mind the consequences of false-positive and false-negative decisions [12-17].

For medical applications, false-positive linkages cause wrong medical information to be assigned to a person. This must be avoided in all cases. The consequences of false-negative linkage (not assigning, for example, diagnoses or results to a patient) would mean that data available for a person are not recognized. Of course, this should also be avoided, but the consequences are not as dramatic as for falsepositive linkage.

In epidemiological studies, false-positive linkages generally result in underestimating true rates, whereas false-negative linkages result in overestimating rates. It is well known that small errors in record linkage (5%) can yield a substantial error in the estimated rates (see e.g. Pukkala, lecture at the IARC 1998 conference in Atlanta).

When applying our method, false-positive record linkage results (homonyms) can occur in the following situations based on p probability: For $p \in (p^2, 100]$ the decision is based only on the p probability. Based on our choice of $p^2 = 95$, a false-positive decision occurs only when there are minimal differences in a single component and all other components have identical values. For $p \in [p^1, p^2]$ all decisions are made by the user. The method can prompt false-positive decisions if the resulting list contains long parts with identical pairs interspersed by a few pairs describing different persons.

False-negative record linkage results (synonyms) can occur in the following situations based on the p probability: For $p \in (0, p^1)$, the pair is not included in the output file. For $p \in [p^1, p^2]$, all decisions are made by the user. The method can provoke false-negative decisions if the resulting list contains long parts with non-identical pairs interrupted by a few pairs describing the same person.

In order to reduce false-positive and false-negative results, the critical bounds p¹ and p^2 can be changed. It should be noted that every change in the critical bounds has consequences for the time needed to decide the unclear cases and in some respect also for the overall result, bearing in mind the potentially longer lists with unclear cases which can also provoke additional errors. Many decisions can be made just by taking a close look at the components. Other decisions require further information and in general a few minutes of time. Good decisions are based on proper knowledge of data origins, on knowledge of typical registration errors and on good knowledge of frequent last names and first names.

Validation of Method

The correctness of the method presented depends on three factors, namely the correct implementation of the probabilistic record linkage method, the proper choice of critical bounds and the thoroughness of the clerical staff working on the list of unclear cases.

Implementation of the method by writing a software program was checked and carefully tested by proper cross-reading of the code and by applying the program to suitable test data. The proper choice of critical bounds was discussed in the previous chapter.

By implicit assumption, the method also depends on the availability of the key information needed for the method. As described in the Introduction, the cancer registries usually collect these data accurately.

In order to check the overall result of the method, we reanalyzed two typical applications of the record linkage method. As mentioned, we use the program for two purposes, namely to detect persons registered multiple times and to combine two databases. Both functions were checked systematically.

Checking for persons registered multiple times was done for all incident cancer cases of the year of diagnosis 1996. Checking for errors when combining two databases was performed by linking the incidence data of the year 1996 and the mortality data for the years 1996 to 2001. We searched for falsepositive and false-negative pairs. This was done by means of a long list of heuristic checks, for example persons for whom the first three letters of their last name and their complete date of birth are identical, or persons for whom the first five letters of their last name and the month and year of birth are identical. In total, we could not find any false-positive or false-negative combination. Also, we could not find any person registered multiple times. It should be mentioned that one possible bias within this check is the fact that the re-evaluation was done by the same clerical staff, who therefore could make the same wrong decision a second time.

Practical Considerations

Our method needs additional time as compared to deterministic procedures. This is the case for every probabilistic record linkage procedure, because they result in cases that cannot be decided automatically per definition. Thus, when applying a probabilistic method, one has to decide how much time to spend on deciding the status of unclear cases. Both our main applications, namely detecting persons registered multiple times and assessing patient life status, have direct impact on main results and we therefore decided to invest the extra time in order to obtain reliable incidence and survival rates.

Table 5 shows that 11% of identical pairs were not decided automatically and that of those cases decided automatically 6.9% did not have simultaneous identity of last name and first name and date of birth. This means that around 15% of cases would not have been linked by the widely used rules of deterministic record linkage procedures.

One further aspect should be mentioned that is specific for our region: residential mobility is low. We know from studies that patients have on average only about three residences throughout their lifetime [20]. This means that change of patient address is rather unlikely to occur and so the component municipality code or zip code is very stable.

Commercial programs are available for record linkage, Automatch [10, 11, 18, 19] being one of the main programs used in this area. Automatch offers very good implementation of the methodology of probabilistic record linkage. The main difference between Automatch and our solution is the consideration of what we call additional methods defined in Table 3. In addition, our implementation is adapted to cancer registry data structure, and all decisions concerning choice of parameters are fix-coded so that all user interactions are minimized, resulting in a very time-efficient operation. A further reason was the rather high price of Automatch.

One of the problems encountered in practical record linkage is that more or less precise information is needed to identify a person while every registry must observe strict data privacy laws [21-23]. The legal basis for our cancer registry allows us to store all data on identification of patients, of course in compliance with strict guidelines to safeguard confidentiality. We hope that in future a unique person identifier will be introduced in our country, which would overcome record linkage problems and all data privacy concerns [24].

Conclusions

We have developed a record linkage method for cancer registries in Austria based on the theory of probabilistic record linkage adjusted for special conditions in the German language. The method serves two main purposes, namely record linkage of various data sources and identification of persons registered more than once in the database. Both goals were reached with adequate precision. The time needed to decide unclear cases is justifiable.

References

- Parkin DM, Whelan SL, Ferlay J, Raymond L, Yuen J. Cancer Incidence in Five Continents. Volume VII. IARC Scientific Publications No. 143. Lyon: IARC, 1997.
- Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B. Cancer Incidence in Five Continents. Volume VIII. IARC Scientific Publications No. 155. Lyon: 2002.
- Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG. Cancer Registration. Principles and Methods. Lyon: IARC, 1991.
- Schouten LJ, Hoppener P, van den Brandt P, Knottnerus JA, Jager JJ. Completeness of Cancer Registration in Limburg, The Netherlands. Int J Epidemiol 1993; 22 (3): 369-76.
- Ziegler H, Stegmaier C. Bevölkerungsbezogene Krebsregistrierung in Deutschland. Onkologie 1996; 19: 268-77.
- Calle EE, Terrell DD. Utility of the National Death Index for Ascertainment of Mortality among Cancer Prevention Study II Participants. Am J Epid 1993; 137 (2): 235-41.
- 7. Fellegi IP, Sunter AB. A Theory for Record Linkage. JASA 1969; 64: 1183-1210.
- Jaro MA. Probabilistic Linkage of Large Public Health Data Files. Stat Med 1995; 14: 491-8.
- Leitner H. Probleme und Lösungen beim Zusammenführen von Patientendaten unterschiedlicher Quellen. Innsbruck: 1995.
- 10. Stegmaier C, Ziegler H. Zusammenführung personenbezogener Informationen (Record-Link-

age) – Probleme und Möglichkeiten am Beispiel des Krebsregisters Saarland. Krebsregister Saarland, Vierteljahresheft 1992; 4:1-7.

- Schmidtmann I, Michaelis J. Untersuchungen zum Record Linkage f
 ür das Krebsregister Rheinland-Pfalz. 1994.
- Brenner H, Schmidtmann I. Determinants of Homonym and Synonym Rates of Record Linkage in Disease Registration. Methods Inf Med 1996; 35 (1): 19-24.
- Brenner H, Schmidtmann I, Stegmaier C. Effects of Record Linkage Errors on Registry-based Follow-up Studies. Stat Med 1997; 16 (23): 2633-43.
- 14. Newcombe HB, Smith ME, Howe GR, Mingay J, Strugnell A, Abbatt JD. Reliability of Computerized versus Manual Death Searches in a Study of Health of Eldorado Uranium Workers. Comput Biol Med 1983; 13 (3): 157-69.
- 15. van den Brandt P, Schouten LJ. Development of a Record Linkage Protocol for Use in the Dutch Cancer Registry for Epidemiological Research. Int J Epid 1990; 19 (3): 553-8.
- 16. MacLeod MC, Bray CA, Kendrick SW, Cobe S. Enhancing the Power of Record Linkage involving low quality Personal Identifiers: Use of the Best Link Principle and Cause of Death prior Likelihoods. Comput Biomed Res 1998; 31 (4): 257-70.
- Gomatam S, Carter R, Ariet M, Mitchell G. An empirical Comparison of Record Linkage Procedures. Stat Med 2002; 21 (10): 1485-96.
- Thomas B, Newman M. Use of Commercial Record Linkage Software and Vital Statistics to Identify Patient Deaths. J Am Med Inform Assoc 1997; 4 (3): 233-7.
- Kendrick SW, Douglas MM, Gardner D, Hucker D. Best-link Matching of Scottish Health Data Sets. Methods Inf Med 1998; 37 (1):64-8.
- 20. Oberaigner W, Kreienbrock L, Schaffrath-Rosario A, Kreuzer M, Wellmann J, Keller G, Gerken M, Langer B, Wichmann HE. Radon und Lungenkrebs im Bezirk Imst/Österreich. In: Wichmann HE, Schlipkoeter HW, Fuelgraff G (eds). Fortschritte in der Umweltmedizin. Landsberg/Lech: Ecomed Verlag; 2002.
- Pommering K, Miller M, Schmidtmann I, Michaelis J. Pseudonyms for Cancer Registries. Methods Inf Med 1996; 35 (2): 112-21.
- 22. Meux E. Encrypting Personal Identifiers. Health Serv Res 1994; 29(2):247–256.
- Rabinowitz J. A Method for Preserving Confidentiality when Linking Computerized Registries. Am J Public Health 1998; 88 (5): 836.
- 24. Kelman C, Smith L. It's time: Record Linkage The Vision and the Reality. Aust N Z J Public Health 2000; 24 (1):100-1.

Correspondence to: Dr. Willi Oberaigner Cancer Registry of Tyrol TILAK GmbH Anichstrasse 35 6020 Innsbruck, Austria E-Mail willi.oberaigner@iet.at

Downloaded from www.methods-online.com on 2011-11-30 | ID: 1000466760 | IP: 193.171.69.65

For personal or educational use only. No other uses without permission. All rights reserved.

Appendix 6

pendix 6

B1-2 Appendix Full Papers



American Journal of Epidemiology Copyright © 2006 by the Johns Hopkins Bloomberg School of Public Health All rights reserved; printed in U.S.A.

Original Contribution

Reduction of Prostate Cancer Mortality in Tyrol, Austria, after Introduction of Prostate-specific Antigen Testing

Willi Oberaigner^{1,2}, Wolfgang Horninger³, Helmut Klocker³, Dieter Schönitzer⁴, Wolf Stühlinger², and Georg Bartsch³

¹ Cancer Registry of Tyrol, Innsbruck, Austria.

² Institute of Quality Science, University of Health Informatics and Technology Tyrol, Innsbruck, Austria.

³ Department for Urology, Innsbruck Medical University, Innsbruck, Austria.

⁴ Institute of Immunology and Transfusion Medicine, Innsbruck Medical University, Innsbruck, Austria.

Received for publication August 5, 2005; accepted for publication February 21, 2006.

The objective of this study was to analyze in detail the time trend in prostate cancer mortality in the population of Tyrol, Austria. In Tyrol, prostate-specific antigen tests were introduced in 1988–1989 and, since 1993, have been offered to all men aged 45–74 years free of charge. More than three quarters of all men in this age group had at least one such test in the last decade. The authors applied the age-period-cohort model by Poisson regression to mortality data covering more than three decades, from 1970 to 2003. For Tyrol, the full model with age and period and cohort terms fit fairly well. Period terms showed a significant reduction in prostate cancer mortality in the last 5 years, with a risk ratio of 0.81 (95% confidence interval: 0.68, 0.98) for Tyrol; for Austria without Tyrol, no effect was seen, with a risk ratio of 1.00 (95% confidence interval: 0.95, 1.05). Each was compared with the mortality rate in the period 1989–1993. Although the results of randomized screening trials are not expected until 2008–2010, these findings support the evidence that prostate-specific antigen testing offered to a population free of charge can reduce prostate cancer mortality.

Austria; mortality; prostate-specific antigen; prostatic neoplasms

Abbreviations: APC, age-period-cohort; ASR age-standardized rate; PSA, prostate-specific antigen.

Prostate cancer is the second-leading cause of male cancer death in most industrialized countries. Thus, the discussion about whether prostate-specific antigen (PSA) testing should be offered in organized screening programs acquires great public health importance. Very large, randomized studies with more than 100,000 cases and controls per study are still ongoing in Europe and the United States; to our knowledge, only one smaller randomized study in Quebec has been concluded (1). These large studies reflect the exceptional interest in scientifically proven evidence on whether organized PSA screening reduces prostate cancer mortality. Until now, screening healthy men for prostate cancer has been shown to be feasible and acceptable in large studies (2). However, conclusive results are not anticipated until 2008–2010 (3), and one must bear in mind that randomized studies are expected to entail some problems with contamination of control groups (2).

PSA tests were introduced in Tyrol, Austria, in 1988– 1989 and, since 1993, have been offered to all men aged 45–74 years (4). In Tyrol, where PSA testing is free of charge and is widely accepted, more than three quarters of men in this age group had at least one PSA test in the period 1993–2003, and some of them have PSA tests regularly. In addition, free annual health checks, including a digital rectal examination, are offered not only in Tyrol but also in all of Austria. Roughly one fifth of men accept this offer of a general medical examination. However, in Austria without Tyrol, PSA tests are not included in the free annual checks

Reprint requests to Dr. Willi Oberaigner, Cancer Registry of Tyrol, Anichstr. 35, A-6020 Innsbruck, Austria (e-mail: willi.oberaigner@iet.at).

and must be paid for by the patient. Consequently, in Tyrol, the prostate has taken the lead among incident cancer sites for men, accounting for one third of all incident cancer cases, although, in terms of mortality, lung cancer is still much more frequent and accounts for one fourth of male cancer deaths. Prostate cancer is responsible for 12 percent of such deaths. The number of incident prostate cancer cases has more than doubled in the last decade, with up to 600 incident prostate cancer cases diagnosed annually in recent years (5-8).

These facts prompted us to conduct an in-depth analysis of time trends in cancer mortality. Our objective was to examine the time trend in prostate cancer mortality by using an age-period-cohort (APC) model to determine whether there was a significant change in the trend and to compare the results for Tyrol with those for Austria without Tyrol.

MATERIALS AND METHODS

Mortality data, which are collected by Statistics Austria (9), were analyzed for Tyrol and for Austria without Tyrol. In Austria, death certificates are issued by official, specially trained medical physicians, pathologists, and forensic medical experts. Specialists at Statistics Austria, the federal institution for statistics in Austria, follow international guide-lines and select one main diagnosis that led to death and assign it one *International Classification of Diseases* code (using the Ninth Revision until 2001, the Tenth Revision since 2002). All procedures concerning death certificates, data collection, and coding are applied uniformly throughout Austria and are not state specific. We analyzed all cases for whom prostate cancer was coded as the cause of death, as described above.

Population data are also collected by Statistics Austria. Census data are available for the years 1971, 1981, 1991, and 2001; for intercensus years, population figures are extrapolated based on births, deaths, and migration information. At the time of our analysis, we had no access to population data for 2003 and thus used the population data from 2002 for 2003 (the difference in population for 1 year is very small: about 0.6 percent for states in western Austria and even less in the eastern states, namely, about 0.2 percent). The male population of Tyrol in census year 2001 was 328,323. In Austria without Tyrol, it was 3,559,913.

The analysis of mortality time trends was based on APC modeling by fitting separate models for Tyrol and for Austria without Tyrol (10, 11). APC models allow separate effects to be estimated for age (A), period or year of death (P), and cohort (C) by means of Poisson regression. In a more formal sense, we fit a series of models as follows:

 $log(\rho_{APC}) = \alpha_A + \beta_P + \gamma_C, \text{ where } C = P - A,$ \rho denotes the mortality rate.

The model is often written in antilogs as follows:

$$\rho_{APC} = \alpha_A ' \beta_P ' \gamma_C ',$$

where $\alpha_A '$ denotes the antilog of α_A or
 $\alpha_A ' = \exp(\alpha_A)$, and so forth.

Am J Epidemiol 2006;164:376-384

As suggested by Clayton and Schifflers (10, 11), a series of models is fit until model fit is adequate. We start with A alone and proceed by including P and/or C in the model if the model fit is not sufficient without the extra term and inclusion of the term substantially improves goodness of fit. Goodness of fit is measured by deviance, which should be equal to or close to the degrees of freedom if the model fit is reasonably good.

For statistical analysis, the number of prostate cancer deaths was aggregated in 5-year age groups, 5-year period groups, and consequently 5-year cohort groups. In Tyrol, there are very few prostate cancer deaths in men less than age 60 years (3.3 percent of all prostate cancer deaths). We thus decided to build the model for age groups beginning with age 60–64 years and to continue by using 5-year age groups. We had access to mortality data beginning in 1970, so our first period group was 1970–1973. The others continued in 5-year period groups and ended with the period group 1999–2003. Our hypothesis was that the mortality rate decreases following PSA testing, so the reference category for period was 1989–1993. Consequently, because C = P - A, cohort groups began with 1882–1886 and continued in 5-year groups.

The analysis was performed with Stata version 8 software, using procedure poisson for Poisson regression (12).

RESULTS

We fitted separate models for prostate cancer mortality for Tyrol and for Austria without Tyrol according to the method suggested by Clayton and Schifflers (10, 11). If a model fits well, the deviance is chi-square distributed with degrees of freedom as given by the model. Therefore, if the deviance is equal to or near the degrees of freedom, the model fits rather well. For Tyrol, the AP model had 30 df and deviance 50.3; the AC model had 25 df and deviance 69.0. After adding period and cohort terms, the APC model reached 20 df with deviance 27.1, which seems reasonably good. For Austria without Tyrol, the AP model had 30 df and deviance 243.7, the AC model had 25 df and deviance 80.8, and the APC model had 20 df and deviance 61.4. We also applied the likelihood ratio test for parameters to test whether the effect of a new parameter was different from zero. For every step in model extension, the likelihood ratio test showed that the parameter effect was different from a zero effect. Thus, it was justified to add each parameter step by step.

One characteristic of the applied model is that period and cohort effects were divided into linear effect (called drift) and nonlinear effect (called nondrift). For Tyrol, we found that neither the AP model nor the AC model reached sufficient model fit, but the APC model fit fairly well. We could not distinguish between drift in period and drift in cohort because, when we modeled drift in cohort terms, we found a strong nondrift period effect, and when we modeled drift in period terms, we found a strong nondrift cohort effect. For Austria without Tyrol, none of the models reached sufficient model fit, so conclusions drawn from the model are to be interpreted with great caution. We modeled drift in period

		Tyrol	Austria	without Tyrol
	Estimator	95% CI*	Estimator	95% CI
Age (years)				
60–64	1	Reference	1	Reference
65–69	2.08	1.63, 2.65	2.29	2.15, 2.43
70–74	4.99	3.93, 6.34	4.62	4.35, 4.92
75–79	8.73	6.72, 11.33	8.96	8.38, 9.57
80–84	14.93	11.07, 20.12	15.44	14.31, 16.66
<u>≥</u> 85	23.83	17.06, 33.30	24.59	22.61, 26.74
Period				
1970–1973	0.71	0.53, 0.96	0.88	0.82, 0.95
1974–1978	0.66	0.51, 0.85	0.90	0.84, 0.96
1979–1983	0.81	0.66, 0.99	0.87	0.83, 0.92
1984–1988	0.83	0.71, 0.97	0.91	0.87, 0.95
1989–1993	1	Reference	1	Reference
1994–1998	0.94	0.81, 1.10	0.99	0.95, 1.03
1999–2004	0.81	0.68, 0.98	1.00	0.95, 1.05
Cohort				
1882–1886	1	Reference	1	Reference
1887–1891	1.50	0.94, 2.39	1.22	1.09, 1.38
1892–1896	1.62	1.07, 2.46	1.25	1.12, 1.39
1897–1901	1.80	1.22, 2.65	1.34	1.21, 1.48
1902-1906	1.65	1.13, 2.40	1.41	1.28, 1.55
1907–1911	1.52	1.04, 2.21	1.45	1.32, 1.59
1912–1916	1.57	1.07, 2.30	1.45	1.32, 1.60
1917–1921	1.31	0.87, 1.99	1.27	1.15, 1.41
1922-1926	1.13	0.72, 1.77	1.18	1.05, 1.32
1927–1931	1.02	0.62, 1.67	1.17	1.04, 1.33
1932–1936	1.06	0.60, 1.87	1.23	1.07, 1.41
1937–1941+				

TABLE 1. Model estimators for period and cohort given by the age-period-cohort model, drift in period, for Tyrol and for Austria without Tyrol, mortality data for Austria, 1970–2003

* CI, confidence interval.

+ Because there was drift in period, there is no estimator for this last cohort.

terms, so what we report here as estimators for period is the linear plus the nonlinear time trend.

Effects from the APC model are described in table 1. The reference category for age was 60–64 years; for period, the reference category was 1989–1993; and for cohort, the reference category was 1882–1886. Figures 1 and 2 show observed age-specific rates and predicted rates in an age-period graph and in an age-cohort graph, respectively.

Age effects were comparable for Tyrol and Austria without Tyrol. Compared with the age group 60–64 years, effects were about 2, 5, 9, 15, and 24 for the age groups 65–69, 70–74, 75–79, 80–84, and \geq 85 years, respectively.

Period effects, each compared with years of death 1989– 1993, were about 0.7–0.8 for the 1970s and 1980s in Tyrol and about 0.9 for both decades in Austria without Tyrol. Details are shown in table 1 and figure 3. For the years after 1993, which means after optional PSA testing was introduced for all men in Tyrol, Tyrol showed an effect of 0.94 (95 percent confidence interval: 0.81, 1.10) for 1994–1998 and a significantly reduced effect of 0.81 (95 percent confidence interval: 0.68, 0.98) for 1999–2003. For Austria without Tyrol, the effects were 0.99 for 1994–1998 and 1.00 for 1999–2003.

For Tyrol, cohort effects were about 1.5 until 1916, after which we found a decrease over the next decade, reaching 1.0 in 1927. For Austria without Tyrol, cohort effects were rather stable, with estimators of 1.20–1.40.

DISCUSSION

Our analysis was based on an observational study conducted among the population of Tyrol, where PSA testing has been offered to men free of charge since it was introduced in the early 1990s. Note that PSA testing is offered in



FIGURE 1. Predicted period effects, by age group (years), of prostate cancer mortality in Tyrol and in Austria without Tyrol, mortality data for Austria, 1970–2003.

an opportunistic way, not in the framework of an organized screening program. In addition, without a system of invitation and reinvitation, about three quarters of men aged 45–74 years underwent at least one PSA test for screening purposes in 1991–2003 (4). It seems justified to compare the time

trend in prostate cancer mortality in Tyrol with that in the other Austrian states (detailed figures are given in table 2) because time trends in prostate cancer mortality were quite comparable until 1990, and health services in general, as well as diagnosis and therapy for cancer patients, are



FIGURE 2. Predicted cohort effects, by age group (years), of prostate cancer mortality in Tyrol and in Austria without Tyrol, mortality data for Austria, 1970–2003.



FIGURE 3. Estimated period effects of prostate cancer mortality in Tyrol and in Austria without Tyrol, mortality data for Austria, 1970–2003.

uniform throughout Austria. PSA tests are also conducted in Austria without Tyrol but not on the same scale as in Tyrol.

The results of our model showed a statistically significant reduction in prostate cancer mortality during the last period (1999–2003) in Tyrol, but no reduction in Austria without Tyrol. Our final model for Tyrol fit well and also included cohort as an independent factor, so period effects were adjusted for cohort effects. In contrast, for Austria without Tyrol, model fit was not good.

Figure 4 and table 2 show an increase in the prostate cancer mortality rate in both geographic areas and higher rates for Tyrol compared with Austria without Tyrol between 1980 and 1990. We observed an increase of about 15 percent in prostate cancer mortality for most central European countries between 1980 and 1990 (13). In Tyrol, the age-standardized rate (ASR) was 11-17 in 1970-1975 and reached a peak between 1987 and 1995, with a mean ASR of 19; in the rest of Austria, the ASR was 13-15 in 1970-1975 and peaked at 18 in 1991. We found no clear reasons for this different increase in Tyrol and in Austria without Tyrol. In the model, we defined the reference category for time as 1989–1993. As a consequence, for both geographic areas, the estimator for this reference time period was 1 and the period estimators for Tyrol were smaller in the 1970s and 1980s than for Austria without Tyrol (figure 3).

A recent publication by Vutuc et al. (14) analyzed prostate cancer mortality data in Austria from 1970 to 2002. That study used a different method, namely, a joined-point regression model, which assumes linear segments and identifies points where the slope changes. Age groups were also defined in a slightly different way. Possibly its greatest difference from our method is that the joined-point regression model did not take cohort effects into account. Finally, Vutuc et al. analyzed mortality data up to 2002, whereas we considered mortality data up to 2003. For Austria without Tyrol, Vutuc et al. found a significant annual decrease of -2.36 for the age group 70–79 years beginning in the year 1989 and a significant annual increase of 1.64 for the age group 80–89 years (we report significant results only). For Tyrol, the authors reported a nonsignificant annual increase of 1.15 for the age group 50–59 years, a nonsignificant annual decrease of -0.60 for the age group 60–69 years, a significant annual decrease of -6.42 for the age group 70–79 years beginning in 1991 (after a nonsignificant annual increase of 1.16 for the age group 80–89 years.

When we looked at the age groups up to 80 years, about two thirds of prostate cancer deaths were found in the age group 70–79 years. However, the Vutuc et al. (14) results also showed a significant decrease.

One might argue that differences in age structure could be responsible for some of the differences in prostate cancer mortality; however, our model considered age groups. In addition, there were only slight differences in age structure between Tyrol and Austria without Tyrol. Whereas in Tyrol the percentages of men aged 65, 75, and 85 years or older were 12.3, 4.7, and 0.9, in Austria without Tyrol, the respective percentages were 10.9, 4.0, and 0.9.

Because we analyzed mortality data, the quality of death certificates was very important to the conclusions we drew. In general, the quality of mortality statistics in Austria has been high for decades (15). Nevertheless, we cannot rule out the possibility that PSA testing has had an influence on death certificates. As mentioned above, coding is performed

Year of		- yrol	Austria wit	hout Tyrol
death	No.	ASR*	No.	ASR
1970	56	17.1	677	13.5
1971	41	12.6	695	13.5
1972	39	11.0	700	13.8
1973	56	15.9	779	14.9
1974	51	13.8	777	15.1
1975	47	13.3	771	14.7
1976	52	14.7	809	15.4
1977	50	13.7	779	14.2
1978	57	15.2	865	16.3
1979	68	18.0	783	14.4
1980	90	22.9	845	15.3
1981	52	13.4	849	15.3
1982	69	17.1	874	16.0
1983	61	15.2	837	15.1
1984	68	17.1	835	15.0
1985	76	17.7	905	15.5
1986	70	16.7	925	15.8
1987	87	20.9	984	16.9
1988	71	15.7	941	16.2
1989	72	15.3	986	17.0
1990	96	20.3	1,014	16.8
1991	96	21.0	1,110	18.3
1992	91	18.3	1,048	17.1
1993	96	20.4	1,081	17.7
1994	95	19.5	993	16.1
1995	93	19.2	1,109	17.4
1996	91	17.8	1,079	16.9
1997	88	15.9	1,096	16.9
1998	60	11.3	1,079	16.1
1999	79	14.2	1,143	16.9
2000	79	13.7	1,150	16.5
2001	85	14.8	1,099	16.1
2002	79	14.0	1,059	15.2
2003	68	11.6	1,092	15.5

TABLE 2. Prostate cancer mortality in Tyrol and in Austria without Tyrol, mortality data for Austria, 1970–2003

* ASR, age-standardized rate per 100,000 using Segi weights.

by one central institution for the whole of Austria and thus is not state specific. Therefore, the only difference could be in how the death certificates are written. We can imagine a bias in each direction: a tendency to code either more prostate cancer deaths because of great public awareness of prostate cancer or fewer prostate cancer deaths because of more caution in denoting prostate cancer as the cause of death. In summary, although we cannot rule out a bias regarding death certificates attributable to the different time trends, our assessment is that if a bias exists, it is probably small

Am J Epidemiol 2006;164:376-384

and cannot explain the 19 percent reduction in prostate cancer mortality we observed in our model.

There are no approximate figures on the volume of PSA testing conducted in Austria without Tyrol. We tried to use sales figures collected by test kit companies, but all information was too imprecise to realistically estimate the PSA testing rate in Austria without Tyrol.

For Tyrol, we collected data from all PSA laboratories and estimated the PSA testing rate based on two assumptions. First, it was for only the Urology Department of Innsbruck Medical University that we knew whether a PSA test was for screening purposes; that is, 85 percent were screening tests, and we assumed the same percentage for all other laboratories. Second, there was no personal identifier for about 500,000 of the PSA tests, and we assumed that the first four digits of the surname and date of birth uniquely identified the person. Details are shown in table 3. After 9 years of intensive PSA testing, we estimated that 75.1 percent of all men aged 45–74 years in Tyrol had had at least one screening PSA test.

Because we had no valid information on the volume of PSA testing conducted in Austria without Tyrol, looking at the time trend in cancer incidence can provide some insight into the amount of such PSA testing. When we compared incidence time trends between Tyrol and Austria without Tyrol, we found an ASR of 40–53 for 1988–1991. Afterward, the incidence rate in Tyrol already had doubled by 1993 (ASR = 87), and we observed an ASR of 100–130 since 1997. In Austria, however, from 1988 to 1991, the ASR was identical to the rates in Tyrol; we observed an increase beginning in 1993 and an ASR of 79–90 since 1998. Details are shown in table 3. Thus, for Austria, we expect a smaller decrease in mortality, and we expect the decrease to begin some 5 years later.

Our estimation of the PSA testing rate shows that, in 1995 and 1997, more than one third and one half, respectively, of all men in the age group 45–74 years in Tyrol had at least one PSA screening test. However, our estimation did not consider PSA tests before 1993. Thus, we tended to underestimate the true PSA screening rate. In other words, the period when half of the men had at least one PSA screening test is likely to be 1 or 2 years earlier. The model shows a decrease in prostate cancer mortality in Tyrol by one third around 2000. These data would fit a screening latency period of 5–7 years, which has been shown for mammography screening programs.

The effect of screening programs depends on the sensitivity and specificity of the detection method but also on the efficacy of the therapy applied for the cases detected in the screening program. This second component should not be underestimated. In fact, in Tyrol, a large proportion of such patients are treated by high-quality radical prostatectomy. This high quality of outcome is also shown by the excellent survival figures, for example, in the EUROCARE study (16).

The reduction in prostate cancer mortality in Tyrol could be due to 1) prevention of the disease, 2) detection of the disease at a stage when it is more likely to be curable, or 3) improved outcome of therapy for metastatic disease (4). We discuss these possibilities in order to explain the



FIGURE 4. Age-standardized rate of prostate cancer mortality in Tyrol and in Austria without Tyrol, mortality data for Austria, 1970–2003. Segi weights, world population according to Segi, modified by Doll.

		Prostate cancer incidence		PSA screen	PSA screening test in Tyrol in the age group 45-74 years			
Year	Т	yrol	Aust	ria†	No. of men	Male	Testing rate‡ (%)	Cumulative
	No.	ASR*	No.	ASR	screening test	population (no.)		testing rate§ (%)
1988	203	49.2	2,125	44.1				
1989	216	51.9	2,322	48.5				
1990	179	40.6	2,309	47.6				
1991	202	47.2	2,285	46.5				
1992	291	68.4	2,422	48.9				
1993	365	87.3	2,709	54.5	9,474	86,067	11.0	11.0
1994	490	117.3	3,156	63.0	14,147	88,342	16.0	23.3
1995	403	94.2	3,487	68.4	20,309	90,153	22.5	34.6
1996	408	92.4	3,699	71.5	23,839	91,497	26.1	44.1
1997	480	109.1	4,001	76.2	26,796	92,607	28.9	51.0
1998	506	114.9	4,218	79.2	30,228	93,719	32.3	56.8
1999	471	103.5	4,593	84.6	36,366	95,000	38.3	63.3
2000	596	130.1	4,925	89.5	41,860	96,692	43.3	70.1
2001	600	130.9	5,131	91.5	44,400	98,638	45.0	75.1

TABLE 3. Prostate cancer incidence rates in Tyrol and in Austria without Tyrol, and the PSA* screening rate in Tyrol, mortality data for Austria, 1970–2003

* PSA, prostate-specific antigen; ASR, age-standardized rate per 100,000 using Segi weights.

† We had data for Austria only as a whole, not for Austria without Tyrol (the male population of Tyrol constitutes 8 percent of the Austrian male population).

‡ We had no unique personal identifier; therefore, we estimated men to be uniquely identified by the first four characters of their surname and date of birth. We had no information on screening intention in the PSA database, so we estimated the screener percentage as 85 on the basis of detailed data in the database of the Urology Department of Innsbruck Medical University; each man is counted only once per year.

§ For the cumulative testing rate, each man was counted only once from 1993 to the end of the respective period.

different results for Tyrol and for Austria without Tyrol. Obviously, as shown in table 3, the disease has not been prevented. PSA testing is known to result in a shift toward earlier stages of the disease (4, 17, 18). Incidence data from Tyrol show that the ASR for metastatic cancer decreased from 5.2 to 2.1 and for advanced cancer (stage IV according to Union International Contre Cancer) from 7.9 to 3.7, whereby each decrease was calculated from the period 1988–1992 to the period 1998–2002. We had only limited data for Austria without Tyrol showing a reduction in disseminated prostate cancers of 20 percent in the last decade (19). Labrie et al. (1) reported that, in the Quebec study, only one of 159 cancers (0.6 percent) was metastatic, and Hugosson et al. (18) found that 97 percent of the cancers detected by screening were clinically localized. Jani et al. (20) also reported stage shifts. Thus, the mortality reduction in Tyrol and the stage shift are in line with observations made in other studies, and the different sizes of stage shift are in line with the different results for Tyrol and for Austria without Tyrol.

With regard to improved outcome of therapy for metastatic disease, all patients in Austria have equal access to therapeutic resources; radiotherapy and hormonal therapy are offered in a similar way throughout Austria. In addition, aside from a small amount of money to be paid by hospital patients beginning recently, diagnosis and therapy are free of charge for everyone. Therefore, it is very unlikely that differences in therapy or differences in improvements in therapy caused the differences in mortality reduction between Tyrol and Austria without Tyrol. In conclusion, the main difference between Tyrol and Austria without Tyrol seems to be the high percentage of men in Tyrol who underwent a PSA test.

Other studies also show benefits of PSA screening. In a very detailed analysis, the Surveillance, Epidemiology, and End Results Program group showed possible benefits of PSA screening, although this study was also population based with known possible biases. The authors concluded that part of the decline in prostate cancer mortality in the United States could be due to PSA screening, although they did not rule out other interpretations (17, 21, 22). An analysis of data for England and Wales also showed a reduction in mortality, but there was little evidence that PSA screening was the main reason for that reduction; figures show that a change in therapy probably influenced mortality there (23).

The main problem with our analysis is that nonrandomized studies are prone to several biases. It is hoped that this problem will be solved by the large, randomized screening studies under way in both Europe (European Randomized Study of Screening for Prostate Cancer (2)) and the United States (24). Up to 2002, the European study—in part population based, in part volunteer based—had enrolled 220,000 men. Neither large study will perform its final analysis before 2008–2010 (25), and there is some concern about contamination of control groups (2). A small study with 46,486 participants was conducted in Quebec, Canada (31,133 men in the intervention arm and 15,353 in the control arm), and its last update showed a relative risk of 0.38 ($p \leq 0.0002$), in other words, a 62 percent reduction in

Am J Epidemiol 2006;164:376-384

prostate cancer deaths in the screened group. The 33 percent mortality reduction seen in our study 10 years after PSA testing was offered to all men in the age group 45–74 years is in line with findings from the Quebec study if we bear in mind that our result was derived from a population-based analysis.

While we wait for the conclusive results of the large randomized studies, there is great public health eagerness to know more details of the potential benefit of PSA screening. Our study concerned a well-defined population in Tyrol, where we had detailed knowledge of PSA testing rates and information on therapy offered to the population. The APC model fit well for Tyrol, and, in comparison to Austria without Tyrol, the PSA testing rate seemed to be the main factor explaining the difference in time trends between Tyrol and Austria without Tyrol. Of course, our analysis could not overcome the problems of nonrandomized studies, but it can provide further information on the potential benefits of PSA testing or screening.

ACKNOWLEDGMENTS

Conflict of interest: none declared.

REFERENCES

- Labrie F, Candas B, Cusan L, et al. Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. Prostate 2004;59: 311–18.
- de Koning HJ, Auvinen A, Bereguer Sanchez A, et al. Largescale randomized prostate cancer screening trials: program performances in the European Randomized Screening for Prostate Cancer trial and the Prostate, Lung, Colorectal and Ovary cancer trial. Int J Cancer 2002;97:237–44.
- Makinen T, Tammela TL, Stenman UH, et al. Second round results of the Finnish population-based prostate cancer screening trial. Clin Cancer Res 2004;10:2231–6.
- 4. Bartsch G, Horninger W, Klocker H, et al. Prostate cancer mortality after introduction of prostate-specific antigen mass screening in the Federal State of Tyrol, Austria. Urology 2001;58:417–24.
- Oberaigner W, Mühlböck H, Harrasser L. Biannual report on cancer incidence and mortality in Tyrol/Austria, 1997/1998. (In German). Innsbruck, Austria: IET-Bericht, 2003.
- Oberaigner W, Mühlböck H, Harrasser L. Biannual report on cancer incidence and mortality in Tyrol/Austria, 1999/2000. (In German). Innsbruck, Austria: IET-Bericht, 2004.
- Parkin DM, Whelan SL, Ferlay J, et al, eds. Cancer incidence in five continents. Vol VII. Lyon, France: International Agency for Research on Cancer, 1997. (IARC scientific publication no. 143).
- Parkin DM, Whelan SL, Ferlay J, et al, eds. Cancer incidence in five continents. Vol VIII. Lyon, France: International Agency for Research on Cancer, 2002. (IARC scientific publication no. 155).
- 9. Hansluwka H. Cancer mortality statistics: availability and quality aspects of mortality data worldwide. Recent Results Cancer Res 1989;114:163–75.

- Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: age-period and age-cohort models. Stat Med 1987;6:449–67.
- Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: age-period-cohort models. Stat Med 1987;6: 469–81.
- 12. Stata Corporation. Stata statistical software, release 8. College Station, TX: Stata Corporation, 2003.
- WHO Mortality Database (WHO Statistical Information System (WHOSIS)). Table 1: Numbers and rates of registered deaths, Austria—2002. World Health Organization, 2005. (http://www3.who.int/whosis/mort/table1_process.cfm).
- 14. Vutuc C, Schernhammer ES, Haidinger G, et al. Prostate cancer and prostate-specific antigen (PSA) screening in Austria. Wien Klin Wochenschr 2005;117:457–61.
- Holzner JH. The role of autopsy in the control of mortality in Austria. In: Riboli E, Delendi M, eds. Autopsy in epidemiology and medical research. Lyon, France: International Agency for Research on Cancer, 1991:25–35. (IARC scientific publication no. 112).
- 16. Coleman MP, Gatta G, Verdecchia A, et al. EUROCARE-3 summary: cancer survival in Europe at the end of the 20th century. Ann Oncol 2003;(14 suppl 5):v128–49.
- Hankey BF, Feuer EJ, Clegg LX, et al. Cancer surveillance series: interpreting trends in prostate cancer—part I: evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. J Natl Cancer Inst 1999;91: 1017–24.

- Hugosson J, Aus G, Bergdahl S, et al. Population-based screening for prostate cancer by measuring free and total serum prostate-specific antigen in Sweden. BJU Int 2003; 92(suppl 2):39–43.
- 19. Statistics Austria (ed). Annual health statistics for Austria. (In German). Vienna, Austria: Statistics Austria, 2004.
- 20. Jani AB, Vaida F, Hanks G, et al. Changing face and different countenances of prostate cancer: racial and geographic differences in prostate-specific antigen (PSA), stage, and grade trends in the PSA era. Int J Cancer 2001;96:363–71.
- Etzioni R, Legler JM, Feuer EJ, et al. Cancer surveillance series: interpreting trends in prostate cancer—part III: quantifying the link between population prostate-specific antigen testing and recent declines in prostate cancer mortality. J Natl Cancer Inst 1999;91:1033–9.
- Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer—part II: cause of death misclassification and the recent rise and fall in prostate cancer mortality. J Natl Cancer Inst 1999;91:1025–32.
- Oliver SE, Gunnell D, Donovan JL. Comparison of trends in prostate-cancer mortality in England and Wales and the USA. Lancet 2000;355:1788–9.
- 24. Andriole GL, Reding D, Hayes RB, et al. The prostate, lung, colon, and ovarian (PLCO) cancer screening trial: status and promise. Urol Oncol 2004;22:358–61.
- Auvinen A, Hugosson J. The rationale for the ERSPC trial: will it improve the knowledge base on prostate cancer screening? BJU Int 2003;92(suppl 2):14–16.



Gynecologic Oncology 103 (2006) 527-534

Gynecologic Oncology

www.elsevier.com/locate/ygyno

Influence of department volume on cancer survival for gynaecological cancers—A population-based study in Tyrol, Austria

Willi Oberaigner^{a,b,*}, Wolf Stühlinger^b

^a Cancer Registry of Tyrol, Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd. Innsbruck, Austria ^b Department of Quality Science, Medical Planning and Information Management, University for Health Informatics and Technology Tyrol, Innsbruck, Austria

> Received 17 January 2006 Available online 26 May 2006

Abstract

Objective. The objective of this study was to assess the effect of department volume on survival of patients with gynaecological cancer. *Methods.* We conducted an observational population-based study in Tyrol, Austria. The analysis includes all patient data on incident gynaecological cancer collected by the Cancer Registry of Tyrol. Data were collected since 1988 on a population-based perspective; publication of incidence data since 1988 in Cancer Incidence in Five Continents gives evidence for good completeness and validity of the database. Patient survival status is assessed in a passive way by probabilistic record linkage between incidence data and official mortality data. We applied a multivariate Cox regression with variables age, sex, stage, year of diagnosis, histological verification of diagnosis, transfer to other hospital and department volume. Department volume was categorised in ≤11/12–23/24–35/≥36 patients per year reflecting one/two/three/more than three patients per month; categories were computed separately for every site we analysed. Departments with up to 11 patients per year were called small departments.

Results. For 4191 breast cancer patients, we found a negative effect for small departments; hazard ratio (HR) 1.39, 95% confidence interval (CI) 1.22, 1.58. For ovarian cancer patients, we also found a negative effect for small departments (HR 1.27, 95% CI 1.05, 1.54). For cervical cancer patients, we found a positive effect for small departments (HR 0.67, 95% CI 0.51, 0.88). No effect was shown for corpus cancer (HR 0.80, 95% CI 0.63, 1.01).

Conclusion. The results indicate that, in our country, rules on minimum department case-load can further improve survival for breast and ovarian cancer patients.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Gynaecological cancer; Survival rate; Department volume; Minimum caseload; Cancer epidemiology

Introduction

The question whether for cancer patients department volume has an influence on overall survival and other outcome parameters has been investigated for more than a decade. Answers are of great relevance for health planning and policy in the respective countries. An overview published in 2000 [1] concludes that there is an association between centre size and survival for all solid cancer sites for which therapy is complex. One group of publications concentrates on specific cancer sites and/or specific therapy modalities, while another group analyses this question primarily on a population basis. In addition, some authors discuss interesting methodological questions like publication bias or self-interest bias.

In our country, about 25% of patients with gynaecological cancer are treated in small departments (with less than 11 patients per year), ranging from 16% to 52% of patients depending on specific cancer site. Hence, studying the association between department volume and survival was of special public health interest. We consequently analysed the question on a population basis taking into the study all cancer patients diagnosed in the population of Tyrol, not only patients qualifying, for example, for clinical trials. Also, we analysed all major gynaecological cancer sites. In this way, we tried to avoid both biases mentioned above

Abbreviations: CI, confidence interval; HR, hazard ratio; pat/year, patients per year; DCO, death certificate only.

^{*} Corresponding author. Cancer Registry of Tyrol, Austria, TILAK GmbH, Anichstrasse 35, Innsbruck, Austria. Fax: +43 512 504 22315.

E-mail address: willi.oberaigner@iet.at (W. Oberaigner).

by reporting results for all investigated cancer sites, irrespective of the kind of results.

Material and methods

The Cancer Registry of Tyrol was established in 1986. Cancer data for the population of Tyrol have been registered on a population basis since 1988. Also, since 1988, data have been published in Cancer Incidence in Five Continents [2], thus giving evidence of good completeness for the incidence data.

Registration is performed from a standardised questionnaire including sex, age, cancer site and histology, date of diagnosis, stage and basic information on primary treatment. Information on co-morbidity is not collected routinely. There are strict rules for collecting these variables in accordance with international guidelines (see for example [3]). The questionnaire is either completed by a physician, or a Cancer Registry clerk collects data directly from clinical records in the treating hospital. Two independent data bases are built up, one incidence database and one we call search database including all information on possible cancer diagnoses (mainly pathology reports, but also information from radiotherapy units and various other data sources) allowing the registry to check completeness. Cancer cases are attributed to treating departments according to place of initial treatment.

Patient life status is assessed in a passive way. We do a probabilistic record linkage between incidence data and the official mortality data set for Tyrol collected by Statistics Austria [4]. In Austria, there is no general use of unique person identifiers as, for example, in Scandinavian countries. Therefore, the Cancer Registry of Tyrol developed a method for probabilistic record linkage based on probabilistic record linkage theory. Using the components last name, birth surname, first name, date of birth, sex and municipality code or zip code, a probability of identity is computed for every pair of persons (denoted p-val), also taking into account phonetic translations and documentation and typing errors. If p-val is greater than 0.95, we assume without further checks that the components describe the same person; for a p-val smaller than 0.75, we assume, again without further checks, that the components describe different persons. A p-val between 0.75 and 0.95 calls for a decision on a case-by-case basis. In general, this means that further information is needed to describe the persons more precisely.

Closure of this study was end of 2003. For a few cases, we received information on out-migration, but only by chance. We cannot systematically check for out-migrant status due to data privacy constraints. However, aggregated data on out-migrants in the population of Tyrol show that, in the age classes above 50, which are the relevant age classes for cancer survival, the out-migrant rate is less than one percent of the population.

We analysed the main gynaecological cancer sites: breast, ovary, cervix and corpus. From 1988 to 2000, 4366 breast cancer cases, 976 ovarian cancer cases, 819 cervical cancer cases and 923 corpus cancer cases were registered in the Cancer Registry. Of these, 169 breast cancer cases, 64 ovarian cancer cases, 15 cervical cancer cases and 16 corpus cancer cases were excluded from analysis because of death certificate only (DCO) status and six breast cancer cases and one ovarian cancer case because of other reasons, mainly due to loss of follow-up. Thus, the final study included 4191 breast cancer cases, 911 ovarian cancer cases, 804 cervical cancer cases and 907 corpus cancer cases.

Care is provided by gynaecologists, medical oncologists and radiation oncologists for ovarian, cervix and corpus cancer and, in addition, by general surgeons for breast cancer. There is no training available in gynaecologic oncology in Austria. Radiotherapy is offered by one Department of Radiotherapy of Innsbruck Medical University and by a radiotherapy unit within the Department of Gynaecology of Innsbruck Medical University. Transfer to another hospital was defined as transfer during primary treatment.

A multivariate Cox model was applied using the variables age at diagnosis, year of diagnosis, histological confirmation, stage according to UICC, transfer to another hospital and residence. Age was categorised in groups $0 \le 54/55-64/65-74/\ge 75$ and year of diagnosis in groups 1988-1992/1993-1996/1997-2000. Follow-up time is shorter for more recent periods. From a theoretic point of view, this should not bias the results under the assumption that events are evenly distributed over time for all three period groups. The study area is

Appendix 18

served by one university hospital treating about half of the patients and nine regional hospitals. Department size was defined as average number of incident patients per year (pat/year) and categorised in groups $\leq 11/12-23/24-35/\geq 36$ pat/year; department size was computed for every site separately. We defined categories a priori according to the rationale one, two, three or more than three patients per month. Departments with ≥ 36 pat/year are called large departments and departments with 1–11 pat/year are called small departments. In Cox analysis, reference group is defined by large departments except for ovarian cancer and corpus cancer, for which the largest departments had no more than 24–35 pat/year.

Residence was grouped in the capital city Innsbruck and surroundings (Ibk), the western part of Tyrol (OL), the eastern past of Tyrol (UL) and East Tyrol (LZ), which is a county geographically separate from the main part of the state.

Statistical analysis was done with Stata Version 8.0 [5]. After univariate analysis, we fitted a multivariate Cox model separately for every cancer site by initially entering all variables into the model and then removing variables without significant influence (backward elimination). To check the influence of variables, the likelihood ratio test was applied. After the model was set up, we checked proportional hazard ratio assumption first graphically and then by procedure stphtest of Stata.

Significance was tested at the alpha level of 5%. We present hazard ratios (HR) together with 95% confidence intervals (95% CI).

The population of Tyrol was 612,309 in the year 1988, of which 316,057 were females (51.6%). The female population increased to 342,728 in the year 2000.

Results

Fig. 1 and Table 1 show an overview of all cancer sites investigated. For following cancer sites, we found a significant negative effect for small departments as compared to large departments: breast cancer with HR 1.39 (95% CI 1.22, 1.58) and ovarian cancer with HR 1.27 (95% CI 1.05, 1.54). For cervical cancer, we found a positive effect with HR 0.67 (95% CI 0.51, 0.88). A nonsignificant effect was found for corpus cancer at HR 0.80 (95% CI 0.63, 1.01), although the effect was near significance.

The following section describes results for individual cancer sites in more detail.

Of 4191 breast cancer patients, 1/3 were age 54 or younger and 22% were age 75 or older; see Table 2. Multivariate analysis was adjusted for age, histological confirmation, stage, year of diagnosis and department volume; see Table 1.

Of all cases, 3% had no histological verification (HR 2.68, 95% CI 2.17, 3.30), while 33% were stage I (reference category),



Fig. 1. Hazard ratio for small departments, by cancer site (HR: Hazard ratio adjusted for age, stage, histological confirmation and year of diagnosis. Reference category is large departments \geq 36 pat/year for breast cancer and cervical cancer and 24–35 pat/year for ovarian and corpus cancer).

Table 1 Multivariate hazard ratios with 95% confidence interval^a

	Breast	Ovarian	Cervical	Corpus
	cancer	cancer	cancer	cancer
Age group	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
≤54	1.00	1.00	1.00	1.00
55-64	1.19 (1.02, 1.39)	2.07 (1.55, 2.78)	1.35 (0.98, 1.86)	2.64 (1.35, 5.14)
65-74	1.37 (1.19, 1.58)	2.93 (2.23, 3.85)	1.96 (1.44, 2.66)	5.97 (3.18, 11.21)
≥75	2.51 (2.20, 2.87)	4.34 (3.29, 5.72)	3.08 (2.28, 4.16)	16.34 (8.75, 30.51)
No histological verification	2.68 (2.17, 3.30)	2.77 (2.01, 3.80)	10.14 (6.38, 16.10)	4.08 (2.27, 7.34)
UICC Stage				
Ι	1.00	1.00	1.00	1.00
II	2.11 (1.83, 2.44)	2.61 (1.68, 4.04)	1.98 (1.34, 2.94)	1.91 (1.30, 2.81)
III	4.16 (3.51, 4.93)	4.02 (2.99, 5.42)	4.19 (3.02, 5.82)	3.16 (2.10, 4.76)
IV	9.89 (8.26, 11.83)	7.64 (5.60, 10.42)	9.45 (6.04, 14.76)	6.62 (4.40, 9.97)
Х	4.51 (3.69, 5.53)	3.27 (2.32, 4.63)	3.77 (2.66, 5.34)	2.62 (1.95, 3.51)
Year of diagnosis			b	e
1988–1992	1.00	1.00		1.00
1993-1996	0.85 (0.76, 0.95)	0.89 (0.73, 1.09)		0.83 (0.63, 1.09)
1997-2000	0.78 (0.68, 0.90)	0.76 (0.61, 0.95)		0.66 (0.48, 0.91)
Department volume				
\geq 36 pat/year	1.00	c	1.00	d
24-35 pat/year	1.07 (0.93, 1.23)	1.00	d	1.00
12–23 pat/year	1.10 (0.96, 1.27)	с	d	d
\leq 11 pat/year	1.39 (1.22, 1.58)	1.27 (1.05, 1.54)	0.67 (0.51, 0.88)	0.80 (0.63, 1.01)

^a Transfer to other department and Study Region make no significant contribution to multivariate model for any cancer site.

^b Year of diagnosis makes no significant contribution to multivariate model for cervical cancer.

^c There are no departments with \geq 36 or 12–23 pat/year for ovarian cancer or corpus cancer.

^d There are no departments with 12-35 pat/year for cervical cancer.

^e The global effect of year of diagnosis is significant and so remains in the multivariate model, P = 0.0492 (Likelihood Ratio Test) and P = 0.0498 (Wald Test).

43% were stage II (HR 2.11, 95% CI 1.83, 2.44), 11% were stage III (HR 4.16, 95% CI 3.51, 4.93), 7% were stage IV (HR 9.89, 95% CI 8.26, 11.83) and 7% were stage X (HR 4.51, 95% CI 3.69, 5.53). Of all breast cancer patients, 34% were diagnosed in the years 1988–1992 (reference category), 32% in 1993–1996 (HR 0.85, 95% CI 0.76, 0.95) and 34% in 1997–2000 (HR 0.78, 95% CI 0.68, 0.90). Of these patients, 51% were treated in large departments, 17% in departments with 24–35 pat/year (HR 1.07, 95% CI 0.93, 1.23), 16% in departments with 12–23 pat/year (HR 1.10, 95% CI 0.96, 1.27) and 16% in small departments (HR 1.39, 95% CI 1.22, 1.58).

We analysed a total of 911 ovarian cancer patients, of whom 29% were age 54 or younger and 24% were age 75 or older, see Table 3. Multivariate analysis was adjusted for age, histological confirmation, stage, year of diagnosis and department volume; see Table 1.

Of all ovarian cancer patients, 7% had no histological verification (HR 2.77, 95% CI 2.01, 3.80), while 26% were stage I (reference category), 6% were stage II (HR 2.61, 95% CI 1.68, 4.04), 34% were stage III (HR 4.02, 95% CI 2.99, 5.42), 20% were stage IV (HR 7.64, 95% CI 5.60, 10.42) and 14% were stage X (HR 3.27, 95% CI 2.32, 4.63). Of these patients, 36% were diagnosed in the years 1988–1992 (reference category), 31% in 1993–1996 (HR 0.89, 95% CI 0.73, 1.09) and 32% in 1997–2000 (HR 0.76, 95% CI 0.61, 0.95); 50% of patients were treated in departments with 24–35 pat/year (reference category), 50% in small departments (HR 1.27, 95% CI 1.05, 1.54). We observed no patients in other size categories.

Of 804 cervical cancer patients, 58% were age 54 or younger and 13% were age 75 or older; see Table 4. Multivariate analysis was adjusted for age, histological confirmation, stage and department volume; see Table 1.

Of all cervical cancer patients, 4% had no histological verification (HR 10.14, 95% CI 6.38, 16.10), while 47% were stage I (reference category), 13% were stage II (HR 1.98, 95% CI 1.34, 2.94), 18% were stage III (HR 4.19, 95% CI 3.02, 5.82), 4% were stage IV (HR 9.45, 95% CI 6.04, 14.76) and 17% were stage X (HR 3.77, 95% CI 2.66, 5.34). Of these patients, 44% were diagnosed in the years 1988–1992, 29% in 1993–1996 and 27% in 1997–2000 (year of diagnosis had no significant influence in the multivariate model); 64% of patients were treated in large departments and 36% in small departments (HR 0.67, 95% CI 0.51, 0.88).

We analysed 907 corpus cancer patients, of whom 15% were age 54 or younger and 27% were age 75 or older; see Table 5. Multivariate analysis was adjusted for age, histological confirmation, stage, year of diagnosis and department volume; see Table 1.

Of all corpus cancer patients, 2% had no histological verification (HR 4.08, 95% CI 2.27, 7.34), while 64% were stage I (reference category), 7% were stage II (HR 1.91, 95% CI 1.30, 2.81), 7% were stage III (HR 3.16, 95% CI 2.10, 4.76), 4% were stage IV (HR 6.62, 95% CI 4.40, 9.97) and 18% were stage X (HR 2.62, 95% CI 1.95, 3.51). Of these patients, 34% were diagnosed in the years 1988–1992 (reference category), 32% in 1993–1996 (HR 0.83, 95% CI 0.63, 1.09) and 35% in 1997–2000 (HR 0.66, 95% CI 0.48,

Table 2			
Patient characteristics and univariat	e HR for breast cancer	by department size (N=4191)

	Department size			Totals	Univariate HR	
	≥36 pat/year	24-35 pat/year	12-23 pat/year	≤11 pat/year		(95% CI)
Age group						
≤54	791 (37.0%)	224 (31.5%)	202 (29.6%)	177 (26.9%)	1394 (33.3%)	1.00
55-64	460 (21.5%)	142 (20.0%)	157 (23.0%)	132 (20.0%)	891 (21.3%)	1.18 (1.01, 1.38)
65-74	474 (22.2%)	164 (23.1%)	187 (27.4%)	145 (22.0%)	970 (23.1%)	1.52 (1.32, 1.75)
≥75	413 (19.3%)	137 (20.1%)	137 (20.1%)	205 (31.1%)	936 (22.3%)	3.35 (2.94, 3.81)
No histological verification	20 (0.9%)	5 (0.7%)	17 (2.5%)	101 (15.3%)	143 (3.4%)	8.13 (6.78, 9.74)
UICC stage						
Ι	752 (35.2%)	240 (33.8%)	222 (32.5%)	180 (27.3%)	1394 (33.3%)	1.00
П	968 (45.3%)	317 (44.6%)	311 (45.5%)	200 (30.3%)	1796 (42.9%)	2.14 (1.86, 2.48)
III	208 (9.7%)	93 (13.1%)	82 (12.0%)	55 (8.3%)	438 (10.5%)	4.77 (4.03, 5.65)
IV	122 (5.7%)	40 (5.6%)	31 (4.5%)	92 (14.0%)	285 (6.8%)	12.47 (10.47, 14.86)
Х	88 (4.1%)	21 (3.0%)	37 (5.4%)	132 (20.0%)	278 (6.6%)	6.83 (5.66 8.24)
Year of diagnosis						
1988–1992	706 (33.0%)	251 (35.3%)	211 (30.9%)	269 (40.8%)	1437 (34.3%)	1.00
1993–1996	670 (31.3%)	251 (35.3%)	227 (33.2%)	192 (29.1%)	1340 (32.0%)	0.74 (0.66, 0.83)
1997-2000	762 (35.6%)	209 (29.4%)	245 (35.9%)	198 (30.0%)	1414 (33.7%)	0.62 (0.54, 0.71)
Transfer to other department	28 (1.3%)	11 (1.6%)	8 (1.2%)	21 (3.2%)	68 (1.6%)	1.78 (1.28, 2.49)
Study region						
Ibk	1653 (77.3%)	3 (0.4%)	10 (1.5%)	318 (48.3%)	1984 (47.3%)	1.00
UL	257 (12.0%)	353 (49.6%)	438 (64.1%)	154 (23.4%)	1202 (28.7%	1.00 (0.95, 1.10)
OL	203 (9.5%)	355 (49.9%)		145 (22.0%)	703 (16.8%)	1.02 (0.89, 1.16)
Lz	25 (1.2%)		235 (34.4%)	42 (6.4%)	302 (7.2%)	1.10 (0.90, 1.33)
Department volume						
≥36					2138 (51.0%)	1.00
24–35					711 (17.0%)	1.14 (0.99, 1.31)
12–23					683 (16.3%)	1.13 (0.98, 1.31)
≤11					659 (15.7%)	1.94 (1.71, 2.20)

0.91); 48% of patients were treated in departments with 24–35 pat/year (reference category) and 52% in small departments (HR 0.80, 95% CI 0.63, 1.01). We observed no patients with other department size categories.

Discussion

Departments

The main conclusion of the analysis is highly influenced by how we define which department is responsible for initial treatment. For cancer patients treated by only one department, this definition was clear. But for some of the patients, more than one department was involved in initial treatment. Our country has no strict rules governing prime responsibility for cancer treatment. So we assigned the chronologically first treating department, and this rule seems to be rather straightforward and make sense. The percentage of patients treated by more than one department is rather small. In addition, when subsetting the analysis of patients treated by only one department, the effects were of similar size. We thus conclude that errors made in defining who holds prime responsibility for cancer treatment did not disturb our results.

The variable for transfer is defined as transfer during primary treatment. Our cancer register is an incidence register that does not collect information on the whole period from diagnosis to death. Therefore, we are not able to present more in-depth information on transfer and we especially do not have data on co-morbidities.

We attempted to estimate the percentage of patients treated by more than one department. Since full data are lacking, we can present only a rough estimate for primary treatment, namely 6% of breast cancer patients, 17% of cervical cancer, 18% of corpus cancer and 9% of ovarian cancer patients.

Staging and other confounders

Every survival analysis depends on how well we can adjust for inter-departmental differences in patient characteristics. In Cox regression, we can adjust for patient characteristics if the information is available. Our Cancer Registry contains information on sex, age at diagnosis, staging, year of diagnosis, histological verification of cancer, transfer to other departments, and residence. We set up a model specific for every cancer site by starting with all parameters in the model and then eliminating parameters with no significant influence on the effects (backward elimination). This is a standard procedure described in many textbooks; see for example [6].

Staging is collected as either TNM stage or FIGO for ovarian cancer sites. Because there were too many combinations of TNM values, we transformed TNM stage to stages I to IV according to UICC rules [7]. For example, the Finnish

Table 3 Patient characteristics and univariate HR for ovarian cancer by department size (N = 911)

	Department size		Totals	Univariate
	24-35	≤11		HR (95% CI)
	pat/year	pat/year		
Age group				
≤54	155 (34.2%)	108 (23.6%)	263 (28.9%)	1.00
55-64	100 (22.1%)	96 (21.0%)	196 (21.5%)	2.27 (1.70, 3.03)
65-74	123 (27.2%)	112 (24.5%)	235 (25.8%)	3.46 (2.64, 4.54)
≥75	75 (16.6%)	142 (31.0%)	217 (23.8%)	5.69 (4.36, 7.43)
No histological	3 (0.7%)	62 (13.5%)	65 (7.1%)	4.43 (3.36, 5.85)
verification				
UICC stage				
Ι	128 (28.3%)	108 (23.6%)	236 (25.9%)	1.00
II	26 (5.7%)	30 (6.6%)	56 (6.2%)	2.73 (1.76, 4.22)
III	215 (47.5%)	93 (20.3%)	308 (33.8%)	3.72 (2.78, 4.99)
IV	43 (9.5%)	141 (30.8%)	184 (20.2%)	9.24 (6.81, 12.52)
Х	41 (9.1%)	86 (18.8%)	127 (13.9%)	5.04 (3.63, 7.00)
Year of diagnosis				
1988-1992	163 (36.0%)	168 (36.7%)	331 (36.3%)	1.00
1993-1996	137 (30.2%)	149 (32.5%)	286 (31.4%)	0.87 (0.71, 1.06)
1997-2000	153 (33.8%)	141 (30.8%)	294 (32.3%)	0.75 (0.60, 0.94)
Transfer to other	6 (1.3%)	27 (5.9%)	33 (3.6%)	1.61 (1.07, 2.42)
department				
Study region				
Ibk	225 (49.7%)	197 (43.0%)	422 (46.3%)	1.00
UL	154 (34.0%)	117 (25.5%)	271 (29.8%)	0.88 (0.72, 1.08)
OL	70 (15.5%)	81 (17.7%)	151 (16.6%)	1.11 (0.88, 1.41)
Lz	4 (0.9%)	63 (13.8%)	67 (7.4%)	1.16 (0.84, 1.60)
Department volur	ne			
24-35			453 (49.7%)	1.00
≤11			458 (50.3%)	1.78 (1.50, 2.12)

nationwide cancer registry categorises staging information as localised/nonlocalised/unknown, and the European Network of Cancer Registries recommends collecting not detailed TNM stage but only what they call condensed TNM. If one of the TNM components is missing, this transformation results in stage X, thus counting unknown as well as imprecise stages. Percentage of stage X depends heavily on cancer site but also on department, because we have indications that some departments have a higher percentage of imprecise staging information (at least imprecise staging information documented in the Cancer Registry). Thus, some problems are encountered when comparing stage X between departments. The percentage of stage X is in line with other publications when we consider that our data set contains all cancer patients of a population, and not only histologically verified cases or cases treated in the framework of clinical trials [8,9].

What remains is the question whether our adjustment for staging effects was precise enough. Following international studies and well-established registries, adjusting for UICC stage seems to be precise enough. For certain cancer sites and for clinical aspects, our analysis may be too imprecise, but on a population basis, this was the best we could achieve. We also tried to find a surrogate measure for terminal cases (meaning cases with very poor prognosis), which were also part of our population-based analysis. We believe that the combination of histological verification, age and stage IV should allow adjustment for terminal patients.

Age at diagnosis was modelled in categories also allowing adjustment for nonlinear effects in age. Age categories were defined a priori. For all cancer sites, the reference category "≤54" was large enough to provide stable estimates. We observed poorer survival in older patients, also in multivariate analysis. In general, our cancer register contains only limited information that can help to explain this fact, especially since we do not collect data on co-morbidities. Information about primary treatment for patients aged 75 and older as compared to patients up to age 74, reveals less radiotherapy and chemotherapy for breast cancer, less surgery and chemotherapy for cervical cancer, less radiotherapy for corpus cancer and less chemotherapy for ovarian cancer. Our results are in line with, for example, those of the EUROCARE working group, who found survival rates to decrease with increasing age for almost all cancer sites [10].

Period effects were modelled to adjust for time effects, for example departments changing their treatment guidelines over the years. Reference category was defined as years of diagnosis 1988–1992. Thus, HR can be interpreted as change in treatment as compared to years 1988–1992. Multivariate analysis shows improved survival over time for breast cancer and ovarian

Table 4

Patient characteristics and univariate HR for cervical cancer by department size (N = 804)

	Department size		Totals	Univariate HR
	≥36 pat/year	≤11 pat/year		(95% CI)
Age group				
≤54	291 (56.7%)	175 (60.1%)	466 (58.0%)	1.00
55-64	89 (17.3%)	41 (14.1%)	130 (16.2%)	1.93 (1.41, 2.64)
65-74	67 (13.1%)	39 (13.4%)	106 (13.2%)	3.41 (2.54, 4.58)
≥75	66 (12.9%)	36 (12.4%)	102 (12.7%)	6.02 (4.53, 7.99)
No histological		31 (10.7%)	31 (3.9%)	17.62 (11.73, 26.47)
verification				
UICC Stage				
Ι	233 (45.4%)	148 (50.9%)	381 (47.4%)	1.00
II	85 (16.6%)	22 (7.6%)	107 (13.3%)	2.70 (1.85, 3.95)
III	124 (24.2%)	18 (6.2%)	142 (17.7%)	5.83 (4.28, 7.94)
IV	17 (3.3%)	18 (6.2%)	35 (4.4%)	14.12 (9.22, 21.64)
Х	54 (10.5%)	85 (29.2%)	139 (17.3%)	5.55 (4.05, 7.61)
Year of diagnosis				
1988-1992	233 (45.4%)	123 (42.3%)	356 (44.3%)	1.00
1993-1996	144 (28.1%)	90 (30.9%)	234 (29.1%)	0.88 (0.68, 1.13)
1997-2000	136 (26.5%)	78 (26.8%)	214 (26.6%)	0.68 (0.50, 0.92)
Transfer to other	10 (1.9%)	41 (14.1%)	51 (6.3%)	1.15 (0.75, 1.77)
department				
Study region				
Ibk	274 (53.4%)	80 (27.5%)	354 (44.0%)	1.00
UL	158 (30.8%)	112 (38.5%)	270 (33.6%)	0.82 (0.64, 1.05)
OL	75 (14.6%)	57 (19.6%)	132 (16.4%)	1.03 (0.77, 1.39)
Lz	6 (1.2%)	42 (14.4%)	48 (6.0%)	0.49 (0.27, 0.88)
Department volur	ne			
≥36			513 (63.8%)	1.00
≤11			291 (36.2%)	0.87 (0.69, 1.09)

Table 5 Patient characteristics and univariate HR for corpus cancer by department size (N = 907)

	Department size		Totals	Univariate HR
	24-35	≤11		(95% CI)
	pat/year	pat/year		
Age group				
≤54	70 (15.9%)	62 (13.2%)	132 (14.6%)	1.00
55-64	116 (26.4%)	126 (26.9%)	242 (26.7%)	2.35 (1.21, 4.55)
65-74	140 (31.9%)	153 (32.7%)	293 (32.3%)	4.80 (2.57, 8.95)
≥75	113 (25.7%)	127 (27.1%)	240 (26.5%)	014.18 (7.66, 26.24)
No histological	1 (0.2%)	17 (3.6%)	18 (2.0%)	3.78 (2.16, 6.59)
verification				
UICC Stage				
Ι	297 (67.7%)	281 (60.0%)	578 (63.7%)	1.00
II	27 (6.2%)	38 (8.1%)	65 (7.2%)	2.23 (1.52, 3.27)
III	31 (7.1%)	31 (6.6%)	62 (6.8%)	2.66 (1.77, 3.99)
IV	21 (4.8%)	18 (3.8%)	39 (4.3%)	7.51 (5.04, 11.20)
Х	63 (14.4%)	100 (21.4%)	163 (18.0%)	2.40 (1.81, 3.19)
Year of diagnosis				
1988-1992	159 (36.2%)	145 (31.0%)	304 (33.5%)	1.00
1993-1996	132 (30.1%)	156 (33.3%)	288 (31.8%)	1.08 (0.83, 1.41)
1997-2000	148 (33.7%)	167 (35.7%)	315 (34.7%)	0.86 (0.63, 1.18)
Transfer to other	17 (3.9%)	71 (15.2%)	88 (9.7%)	1.06 (0.71, 1.59)
department				
Study region				
Ibk	261 (59.5%)	153 (32.7%)	414 (45.6%)	1.00
UL	104 (23.7%)	146 (31.2%)	250 (27.6%)	1.00 (0.77, 1.31)
OL	69 (15.7%)	104 (22.2%)	173 (19.1%)	0.77 (0.55, 1.07)
Lz	5 (1.1%)	65 (13.9%)	70 (7.7%)	1.32 (0.89, 1.98)
Department volur	ne			
24-35			439 (48.4%)	1.00
≤11			468 (51.6%)	0.98 (0.78, 1.23)

cancer. For cervical cancer, an improvement was seen in univariate analysis, but not in multivariate analysis. Finally, corpus cancer does not show an improvement in univariate analysis, but only in multivariate analysis. As already mentioned in the discussion of age effects, our cancer register has limited data and therefore is not able to fully explain the observed time trends. Information on primary treatment shows an increase in chemotherapy for breast cancer, but we do not have detailed information on chemotherapy regimen. For cervical cancer, our data show an increased surgical volume in early stages. In general, we see a clear shift towards early stages for breast cancer and cervical cancer which of course improves outcome.

Our general judgement is that the limited data available in our cancer register restrict our ability to analyse in depth some of the trends observed. This is beyond the scope of cancer registers and should be dealt with in specially designed studies.

We did not adjust for treatment. The reason was that adjustment should compensate for factors influencing survival which cannot be influenced by departments, but not for factors chosen by departments. If we consider the case of a department that offers poor treatment, adjusting for treatment could eliminate differences in outcome, which we feel would not be justified. It can be argued that therapy could heavily confound our analysis. Thus, a subanalysis examined basic variables for primary treatment in the model. Compared to our main analysis, the effects for breast cancer (HR for small departments 1.29, 95% CI 1.13, 1.47) and ovarian cancer (HR for small departments 1.26, 95% CI 1.04, 1.53) are a little smaller but still statistically significant. The effect for corpus cancer (HR 0.87, 95% CI 0.67, 1.12) is smaller and not statistically significant, as already shown in the analysis without therapy, and the effect for cervical cancer (HR 0.81, 95% CI 0.61, 1.09) loses statistical significance. In summary, our main results, namely negative effects for small departments for breast and ovarian cancer, change only slightly and remain statistically significant.

External comparison

The question studied here has been of public interest for more than a decade. Some articles also discuss methodological problems like self-interest bias and publication bias. Our general goal was to analyse all gynaecological cancer sites and report all results. Hence, neither biases mentioned above was relevant for our study. Another bias able to distort results is selection bias in the departments, meaning not all cancer patients are included in the analysis, for example only patients qualifying for certain trials. It is well known that patients treated in clinical trials differ in their survival from other patient groups. Again, this did not play a role in our analysis because we analysed a population-based cancer registry data set covering all cancer patients in our population.

Some articles deal with different outcome measures, for example hospital mortality or 30-day mortality and complications after surgery. This was not possible in our analysis, because we included all cancer patients, namely also patients who did not undergo surgery or even curative therapy. Moreover, we had no information on surgeon; this was never part of the Cancer Registry data set.

When comparing our results with published results, one must consider whether a specific study region employs guidelines which contribute to standardised diagnostics and therapy, or whether a country uses the best treatment principle, as in our country. Such guidelines would tend to minimise outcome differences, because small departments usually should not treat patients with advanced cancer.

This analysis obviously cannot answer the question whether department size per se influences patient outcome or whether department size is merely a surrogate measure counting for various factors influencing results.

When comparing our results with published results, we use the term effect as shorthand for negative effect for small departments.

Our results for breast cancer are consistent with published results. Roohan [11] reported an analysis from New York State with a total of 47890 patients hospitalised between 1984 and 1989. In addition to hospital volume, the investigators had information on patient age, surgery type, stage, co-morbidity, race, socioeconomic status and distance to the hospital. For fiveyear survival, they reported a risk ratio of 1.6 for very low hospital volume (10 or fewer patients) as compared to high hospital volume (151 or more per year). In addition, the investigators discuss a "dose–response" relationship between volume and survival. However, the study period dates back quite far, which means the results might not be applicable for the most recent decade. Skinner [12] reports a negative effect for small departments too. Two main factors are discussed as being responsible for the effect, namely hospital caseload [11] and surgical specialisation [13]. The UK has guidelines for minimum case-load [14,15]. Our database lacks detailed information on surgical specialisation, so we cannot distinguish between these factors.

Our results for ovarian cancer are also in line with most relevant publications. An analysis from Austria [16] found an effect of similar size. This study also includes information on residual cancer after surgery and covers more than half of the gynaecological units in Austria. Elit [17] for Canada investigated academic status and surgical speciality for cases diagnosed in Ontario from 1992 to 1998 in a total of 3355 patients. Analysis was adjusted for age, co-morbidity and metastatic status. The authors reported an HR of 0.7 for gyn-oncologist and of 0.65 for gynaecologist, each compared to general surgeon. Woodman [18] also investigated effects for surgeons as compared to gynaecologists and transfer to oncologists and found no effect for surgeon volume. For these two studies, the focus is not directly comparable to our study. Ioka [19] investigated 3523 patients newly diagnosed in 1975-1995 in Osaka, Japan. By adjusting for age, histological type and cancer stage, the authors report an HR of 1.6 for very low volume (less than one operation per year) as compared to high volume (average of 9 operations per year). Kumpulainen [20] did a population-based study in Finland with 3851 ovarian cancer patients diagnosed from 1983 to 1994. Hospitals were categorised as university, central or other, and by volume quartile. After adjusting for age and stage, the authors reported a relative risk of 1.06 for other hospitals as compared to university hospitals (nonsignificant) and a relative risk of 1.13 for smallest as compared to largest hospitals when categorised by quartile (significant). Du Bois [21] reported results from a German study group and found an 82% elevated risk for nonstudy hospitals versus study hospitals, but no effect for hospital volume. The discussion mentioned that about 15% of German hospitals participated, and a bias towards participation by centres more interested in quality assurance cannot be ruled out. The main reasons discussed for benefits in large centres are that teaching hospitals are reported to do more accurate staging [22] and, in general, cancer management should be done by a multidisciplinary team [18]. Recommendations for centralisation have been given in England [23,24], Scotland [25] and the United States [26].

For cervical cancer, we found a significant positive effect with an HR of 0.67 for small departments. At first view, this result was unexpected. When breaking down the analysis by stage, we found a nonsignificant positive effect for all stages but stage II (data not shown). Departments in Tyrol have agreed that stages II and III are not expected to be treated in small departments. If we repeat our analysis excluding stages II and III, the result remains unchanged (HR 0.63, 95% CI 0.45, 0.89). Patients are not younger in small departments. All nonhistologically verified cases were observed in small departments and slightly more stage I cases were also observed in small departments. We see many more cases with unknown stage in small departments (29.2% versus 10.5% in large departments). Consequently, the adjustment for staging might not be able to fully compensate differences in stage distribution for cervical cancer. We found no recent publications dealing with centre volume and survival for cervical cancer. This might be attributed to publication bias.

For corpus cancer, we found an HR of 0.82 (95% CI 0.65, 1.01), the effect being near significance. Again, this result was unexpected. When we split the analysis by stage, we saw nonsignificant positive effects for stages III, IV and X, no effect for stage I and a negative effect for stage II, each for small departments as compared to larger departments. Excluding stages II and III (which are unlikely to be treated in small hospitals, as for cervix cancer), the resulting HR is nearly unchanged (0.86, 95% CI 0.66, 1.12). There are no differences in age structure between larger and small departments; all but one nonhistologically verified case were diagnosed in small departments (21.4% versus 14.4%). Again, adjustment for staging might not be able to fully correct differences in stage distribution due to misclassification of stages.

For both cervix cancer and corpus cancer, additional information is needed in order to shed more light on the unexpected results. This means that we would need more precise information on therapy and multidisciplinary treatment. In our interpretation, we have doubts whether these results are chance findings and would need detailed information on therapy (not only information on whether surgery or chemotherapy was applied but also more details on treatment regimens) as well as on the degree of coordination by various departments, which seems to occur for these cancer sites.

Conclusion

Our analysis demonstrated for small departments significant negative effects for breast cancer and ovarian cancer and significant positive effects for cervical cancer. The analysis is based on Cancer Registry data sets and hence information on confounders is limited. As in every epidemiological analysis, possible confounders are subject to some limitation. However, most of our results are in line with published results. Consequently, it is necessary to carefully discuss results with clinicians and set up guidelines for minimum department case-load, at least for breast cancer and ovarian cancer.

References

- Hillner BE, Smith TJ, Desch CE. Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. J Clin Oncol 2000;18(11):2327–40.
- [2] Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B. Cancer incidence in five continents. Volume VIII. IARC Scientific Publications No. 155 Lyon: IARC, 2002.
- [3] Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG. Cancer registration. Principles and methods. Lyon: IARC; 1991.
- [4] Oberaigner W, Stühlinger W. Record linkage in the cancer registry of Tyrol. Methods Inf Med 2005;44(5):626–30.

- [5] StataCorp. Stata Statistical Software: Release 8. College Station, TX: StataCorp LP; 2003.
- [6] Cleves M, Gould W, Gutierrez R. An introduction to survival analysis using Stata, Revised edition. Stata Press; 2004.
- [7] Wittekind Ch, Wagner G. TNM Klassifikation maligner Tumoren. 5. Auflage. Berlin: Springer-Verlag; 1997.
- [8] Katalinic A, Holzmann M, Uhlenkamp T, Bartel C, Raspe H. Krebs in Schleswig-Holstein 1998. Lübeck: Institut für Krebsepidemiologie i.V.— Registerstelle des Krebsregisters Schleswig-Holstein; 2000.
- [9] Schmidtmann I, Husmann G, Krtschil G, Seebauer G. Krebs in Rheinland-Pfalz: Inzidenz und Mortalität im Jahr 2000. Mainz: Krebsregister Rheinland-Pfalz; 2002.
- [10] Vercelli M, Capocaccia R, Quaglia A, Casella C, Puppo A, Coeberg JW. Relative survival in elderly European cancer patient: evidence for health care inequalities. The EUROCARE working group. Crit Rev Oncol Hematol 2000;35(3):161–79.
- [11] Roohan PJ, Bickell NA, Baptiste MS, Therriault GD, Ferrara EP, Siu AL. Hospital volume differences and five-year survival from breast cancer. Am J Public Health 1998;88(3):454–7.
- [12] Skinner KA, Helsper JT, Deapen D, Ye W, Sposto R. Breast cancer: do specialists make a difference? Ann Surg Oncol 2003;10(6):606–15.
- [13] Gillis CR, Hole DJ. Survival outcome of care by specialist surgeons in breast cancer: a study of 3786 patients in the west of Scotland. BMJ 1996;312(7024):145–8.
- [14] Calman K, Hine D. Policy framework for commissioning cancer services: a report by the Expert Advisory Group on Cancer to the Chief Medical Officers of England and Wales. Department of Health. London HM Stationary Office; 1995.
- [15] Improving outcomes in breast cancer: guidance for purchasers. London: Department of Health; 1996.

- [16] Sevelda P, Obermair A, Oberaigner W, Marth C. Qualitätserfassung bei Ovarialkarzinom. Krebs! Hilfe! 2004;4:48–9.
- [17] Elit L, Bondy SJ, Paszat L, Przybysz R, Levine M. Outcomes in surgery for ovarian cancer. Gynecol Oncol 2002;87(3):260–7.
- [18] Woodman C, Baghdady A, Collins S, Clyma JA. What changes in the organisation of cancer services will improve the outcome for women with ovarian cancer? Br J Obstet Gynaecol 1997;104(2):135–9.
- [19] Ioka A, Tsukuma H, Ajiki W, Oshima A. Influence of hospital procedure volume on ovarian cancer survival in Japan, a country with low incidence of ovarian cancer. Cancer Sci 2004;95(3):233–7.
- [20] Kumpulainen S, Grenman S, Kyyronen P, Pukkala E, Sankila R. Evidence of benefit from centralised treatment of ovarian cancer: a nationwide population-based survival analysis in Finland. Int J Cancer 2002;102 (5):541–4.
- [21] Du Bois BA, Rochon J, Lamparter C, Pfisterer J. Pattern of care and impact of participation in clinical studies on the outcome in ovarian cancer. Int J Gynecol Cancer 2005;15(2):183–91.
- [22] Wolfe CD, Tilling K, Raju KS. Management and survival of ovarian cancer patients in south east England. Eur J Cancer 1997;33(11): 1835–40.
- [23] Improving outcomes in gynecological cancers: the research evidence. London: NHS Executive, Department of Health; 1999.
- [24] National Cancer Guidance Steering Group: improving outcomes in gynecological cancers: the manual. London: NHS Executive, Department of Health; 1999.
- [25] Fighting the silent killer: optimizing ovarian cancer management in Scotland. Accounts Commission for Scotland: The Commission; 1998.
- [26] Guidelines for referral to a gynecologic oncologist: rationale and benefits. The Society of Gynecologic Oncologists. Gynecol Oncol 2000;78(3 Pt 2): S1–S13.

ORIGINAL ARTICLE

Are survival rates for Tyrol published in the Eurocare studies biased?

WILLI OBERAIGNER^{1,2,3} & UWE SIEBERT ^{2,3,4,5}

¹Cancer Registry of Tyrol, Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Innsbruck, Austria, ²Department of Public Health, Information Systems and Health Technology Assessment, University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria, ³Oncotyrol – Center for Personalized Cancer Medicine, Innsbruck Medical University, Innsbruck, Austria, ⁴Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA and ⁵Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Objective. To investigate whether survival rates published in the EUROCARE studies for Tyrol are distorted, we evaluated data quality in the Cancer Registry of Tyrol. *Material and methods.* Potential errors in completeness of Tyrolean incidence data were assessed by applying semi-quantitative and quantitative methods, in part by comparing indices for Tyrol with those of neighboring countries published in Cancer Incidence in Five Continents. Validity of patient survival status was checked for all cancer patients diagnosed in 1997 (n = 2556). For all 1 026 of these patients still alive at end of 2007, we reassessed survival status. Finally, we re-abstracted date of diagnosis for a subset of 295 patients. *Results.* Quality indices on completeness showed no greater bias with the exception of borderline ovarian cancer, which was in part miscoded in the early nineties. Some differences for bladder cancer and prostate cancer between Tyrol and neighboring countries are due to PSA testing and pathology diagnosis. Concerning patient survival status, four cases were erroneously assessed as alive, five cases died outside Austria, three cases were proven not to belong to the population of Tyrol at time of diagnosis and 21 cases emigrated. Absolute errors in survival rates were less than 0.5 for up to five-year survival rates and less than 1.0 for ten year survival rates. *Conclusions.* Evaluation of data quality in the Cancer Registry of Tyrol demonstrated that the survival rates published for Tyrol are only minimally biased by registration or analysis procedures. However, access to data on emigration, which until now is not possible because of data protection restrictions, would reduce the bias in patient survival status, bearing in mind that the extent of emigration of cancer patients is expected to increase in Austria over the coming years.

Key Words: Cancer registry, record linkage, survival, survival status

The EUROCARE studies published survival rates for many European countries including Austria [1–3]. There has been a broad discussion of the advantages and problems involved in this group of studies. In EUROCARE-3, Tyrol was the only Austrian state to contribute data, while EUROCARE-4 included incidence data from all of Austria. For most cancer sites, Tyrol in EUROCARE-3 and Austria in EURO-CARE-4 were among the countries showing the best survival rates in Europe. For example, cohort survival analysis for years of diagnosis 1995–1999 showed for Austria relative five-year survival rates of 13.9% for lung cancer, 84.9% for prostate cancer and 40.0% for ovarian cancer. For some of the authors/ editors and international experts, these survival results were unexpectedly good and raised scepticism about methodology and possible bias in incidence data and in assessing patient survival status.

A recently published review [4,5] grouped data quality for cancer registries into comparability, validity, timeliness and completeness aspects. We will focus on selected aspects that are directly associated with possible bias in survival rates. First, completeness of incidence data is a selection bias for survival rates [6,7]. This bias can influence survival rates in both directions, towards better survival rates if cases with poor prognosis are not included in the incidence dataset, or towards poorer survival rates if cases with good prognosis are not registered. In total, the impact of problems in under-ascertaining cases is somewhat

(Received 8 August 2008; accepted 12 July 2009)

ISSN 0284-186X print/ISSN 1651-226X online © 2009 Informa UK Ltd. (Informa Healthcare, Taylor & Francis AS) DOI: 10.1080/02841860903188635

Correspondence: Willi Oberaigner, Cancer Registry of Tyrol, TILAK GmbH Anichstrasse 35, Innsbruck, Austria. Tel: +43 512 504 22310. Fax: +43 512 504 22315. E-mail: willi.oberaigner@iet.at

complex. Cases with poor prognosis are more likely to lack histological confirmation and also cause fewer hospital admissions. Hence, it is more difficult to trace these cases. Problems involving completeness can be caused by registration processes but also by errors in diagnosis, due to both pathology diagnosis and coding errors. For example, there are well-recognized differences in the classification and registration of bladder tumors [8]. Secondly, bias regarding the date of diagnosis will clearly influence survival rates. If the registered date of diagnosis is later than the true date of diagnosis, survival is shortened, and vice versa. The definition of the correct date of diagnosis is non-trivial: IARC and ENCR guidelines are followed by many cancer registries [9]. The third bias we will investigate concerns misclassification of patient survival status or by an error in determining the correct date of death. If a patient who has in fact died is registered as alive, this clearly biases towards better survival.

Our objective was to investigate data quality in the Cancer Registry of Tyrol and its impact on the survival rates published for Tyrol.

Materials and Methods

The Cancer Registry of Tyrol

The Cancer Registry of Tyrol was established in 1986. Cancer data for the population of Tyrol have been registered on a population basis since 1988. Also since 1988, data have been published in Cancer Incidence in Five Continents (CI5C) [10–12]. The population of Tyrol in the year 1988, the first year for which incidence data are available, was 612 309, of which 316 057 were females (51.6%), and increased to 674 080 in the year 2001 with a female proportion of 51.3%.

Registration is performed from a standardized questionnaire including sex, age at diagnosis, cancer site and histology, date of diagnosis, stage and basic information on primary treatment. Information on co-morbidity is not collected routinely. There are strict rules for collecting these variables in accordance with international guidelines, see for example [13]. The questionnaire is either completed by a physician, or a Cancer Registry clerk collects data directly from clinical records in the treating hospital. In addition to the incidence database, we also generate a so-called search database, which includes all information on possible cancer diagnoses (mainly pathology reports, but also information from radiotherapy units and various other data sources). Then, all entries in the search database are traced, which results either in an entry in the incidence database or in rejection of the potential cancer diagnosis.

The Cancer Registry of Tyrol routinely assesses patient survival status in a passive way. We employ a probabilistic record linkage method to combine incidence data and the official mortality dataset for Tyrol collected by Statistics Austria [14]. In Austria, there is no general use of unique person identifiers as, for example, in Scandinavian countries. Therefore, the Cancer Registry of Tyrol developed a method for probabilistic record linkage based on probabilistic record linkage theory using the components last name, birth surname, first name, date of birth, sex and municipality code or zip code [15]. Pairs of person identifiers that cannot be automatically identified as identical or different persons must be individually checked by registry personnel.

Evaluation of bias

As we argued in the Introduction, the first bias selected by us for analysis is under-ascertainment or in other words completeness of the incidence dataset. There is no gold standard for assessment of completeness in a cancer registry [5]. We followed the suggestions in [5] and selected both semiquantitative and quantitative methods for estimating possible bias in completeness.

Concerning semi-quantitative methods, we included a) the historic data method (figure of time trend for the four most frequent cancer sites per sex plus all cancer sites combined except non-melanoma skin cancer (NMSC)), b) methods based on mortality:incidence (M:I) ratio (by comparing M:I ratio with that of neighboring countries whose data were published in CI5C and plotting the M:I ratio (2002-2006) versus one minus relative five-year survival (1999-2003), and c) a method based on the microscopically verified (MV) cases method (figure of MV proportion in Tyrol as compared to that of neighboring countries whose data are published in CI5C). For comparison with neighboring countries, we selected the registries for Vorarlberg in Austria, Saarland in Germany, St. Gallen and Graubünden in Switzerland, and Northern Italy. The comparisons were based on the latest edition of CI5C, namely Volume IX covering years of diagnosis 1998 to 2002.

Concerning quantitative methods, we estimated completeness of incidence data by applying the flow method proposed by Bullard et al. [16]. The flow method estimates completeness of incidence data by taking into consideration the logical flow of the registration process and requires information on data from first registration of a cancer case, a copy of all death certificates with cancer as cause of death ("mentioning cancer") and the knowledge whether or not a cancer case was death certificate-initiated (DCI). This method estimates the probability of a patient diagnosed with cancer still being alive at time t after diagnosis, the probability of the death certificate of a patient including a mention of cancer, and the probability of a patient surviving until time t after diagnosis still being unregistered. Using these three probabilities the completeness at time t after diagnosis is estimated; details can be found in [16,17]. Our analysis was performed for year of diagnosis 1999. We used a statistical procedure programmed in STATA [18] that applies Bullard's method and was provided by the Thames Registry.

In a second step, we investigated validity of patient survival status and date of diagnosis, both of which have direct impact on survival rates. Our general goal was to study the validity of patient survival status for a complete year of diagnosis, namely 1997. This year was chosen so that we were able to estimate the impact of possible errors on five- and ten-year survival rates. A total of 2674 cancer cases were registered for year of diagnosis 1997; NMSC cases were excluded. Thirty-four patients had multiple tumors and 81 were DCO cases, thus leaving a total of 2559 cancer patients. Of these, 1026 were alive at end of 2007 according to the registry database. For all of these 1 026 cases we contacted the respective municipal office to obtain up-to-date information on life and migrant status. For some cases, we had to contact other municipal offices if the case had emigrated from the municipality. Impact on survival was investigated by comparing uncorrected and corrected observed and relative survival rates. Survival rates were computed using the STATA procedure strs provided by Paul Dickman [18]. Finally, for a subset of 295 cases drawn for other purposes we also checked the date of diagnosis (which is registered in strict compliance with IARC and ENCR guidelines [9]) by inspecting the pathology reports and/or the hospital records and deriving a re-abstracted date of diagnosis.

Statistical analysis was performed using STATA, Version 9.0 [19].

Results

The results of our completeness analysis applying the flow method are displayed in Figure 1. For year of diagnosis 1999, completeness after three years reached 95%, and after four years, when we finished the registration process, it was 96%.

Results obtained when applying some semi-quantitative methods are shown in Table I and Figures 2–5. To analyze age-specific rates for childhood cancer, we aggregated data for ten years in order to have more stable numbers. For age group 10–14, rates in Tyrol are outside the upper decile of the reference interval [5,10], also for boys aged 5–9 years. The deciles were



Figure 1. Estimated completeness of incidence dataset for Tyrol, year of diagnosis 1999.

derived from data published in Cancer Incidence in Five Continents [10]. Overall, there seems to be a tendency towards higher rates in Tyrol as compared to the reference.

Application of the historical data method by inspecting the incidence time trend is shown in Figure 2. There are of course gradients that vary with cancer, especially for prostate cancer. PSA screening was introduced in the 1990s and caused prostate cancer rates to double. The time trends do not seem to fluctuate in a systematic way.

Plotting M:I ratios against 1-survival shows very good correlation. Some deviations exist for ovarian cancer, kidney cancer and prostate cancer. Next, we compared the M:I ratio with that of neighboring countries published in Cancer Incidence in Five Continents, Vol. IX. Females showed some greater differences for ovarian cancer, while differences in males are greater for bladder cancer, prostate cancer and all sites combined. Finally, we compared the proportion of microscopically verified cases with that reported for the same neighboring countries as above. In total, we observed small differences between Tyrol and its neighboring registries, but some larger MV proportions in Tyrol for lung cancer and pancreas cancer. A statistical test by applying the test-statistics described in [5,11] for Tyrol and the neighboring countries did not flag any of the sites investigated, neither for M:I ratio nor for MV proportion, as statistically significant.

To check patient survival status, we traced back all cancer patients diagnosed in 1997 (NMSC cases were excluded) and still alive at end of 2007, namely 1026. Of these 1026 cases, 992 (96.7%) were proven to have been alive at end of 2007, four cases (0.4%) died in Austria before end of 2007, five cases (0.5%) died before end of 2007 on holiday outside Austria and such cases are not registered in the

	Girls		Boys	
Age group	Age-specific rate	Reference deciles	Age-specific rate	Reference deciles
0-4	16.3	< 9.7 > 21.4	14.5	<12.3>24.7
5–9	9.4	<6.9>12.0	16.0	< 8.5 > 15.6
10–14	17.2	< 6.8 > 13.6	16.9	< 8.5 > 15.0

Table I. Childhood cancer in Tyrol: Age-specific rates (years of diagnosis 1997-2006) and reference deciles.

mortality files. Seven cases (0.8%) emigrated to other Austrian states, 11 cases (1.1%) emigrated to foreign countries and three cases emigrated with unknown destination. Finally, four cases were proven not to have belonged to the population of Tyrol at time of diagnosis. Details are shown in Table II.

The uncorrected observed one, three, five and ten year survival rate was 73.3%, 60.7%, 53.9% and 41.1% and the corrected observed survival rate for one, three, five and ten years was 73.1%, 59.9%, 53.0% and 39.9%, respectively. The differences in relative survival rates are of similar magnitude, details are shown in Table III.



Figure 2. Time trend of age-standardized incidence rate for all sites combined except NMSC and for the most frequent sites (SEGI weights).

For a subset of 295 cases chosen for other purposes we also re-abstracted the date of diagnosis by inspecting pathology reports and hospital records. For 168 cases (56.9%) the re-abstracted date of diagnosis was completely identical to the registered



Figure 3. M:I ratio (2002–2006) versus 1- relative 5-year survival rate (1998–2002).



Figure 4. M:I ratio Tyrol versus that of neighboring countries* by cancer site; data from Cancer Incidence in Five Continents, Vol. IX (1998–2002).

*Northern Italy, Germany (Saarland), Austria (Vorarlberg), Switzerland (St. Gallen), Switzerland (Graubünden/Glarus)

date of diagnosis, and for a total of 286 (96.9%) cases the re-abstracted date of diagnosis was within one month of the registered date of diagnosis. For four cases the re-abstracted date of diagnosis was one to four months too late (thus underestimating published survival time), and for five cases it was one to two months too early (thus overestimating published survival time). Details are shown in Table IV.

Discussion

We investigated data quality in the Cancer Registry of Tyrol. Completeness was studied by applying selected quantitative and semi-quantitative methods for assessing the completeness of incidence data. Furthermore, we studied patient survival status and the impact on survival rate for all cancer patients diagnosed in 1997 and the validity of date of diagnosis for a subset of 297 patients.

First, we will discuss completeness of the incidence data. There is no gold standard or any one simple indicator for assessing the completeness of a

Microscopically verified (%) Females



Microscopically verified (%) Males



Figure 5. Proportion of microscopically verified cases in Tyrol versus in neighboring countries^{*} by cancer site; data from Cancer Incidence in Five Continents, Vol. IX (1998–2002).

cancer registry [5]. Hence, it is necessary to apply various methods and discuss completeness by forming an opinion on the basis of all information. Application of the flow method gave an estimation of completeness of 97% after four years. The flow method relies, among other things, on the fact that

Table II. Corrected patient survival status for all malignant cancer cases diagnosed in 1997 and still alive at 31.12.2007 (N = 1026).

Corrected Survival Status	N (%)
Alive at 31-12-2007	992 (96.7%)
Deceased before 31-12-2007 in Tyrol	4 (0.4%)
Deceased before 31-12-2007 outside Austria ^a	5 (0.5%)
Emigration ^b	21 (2.0%)
No regular residence in Austria at diagnosis ^c	4 (0.4%)

^aPersons living in Tyrol, who die outside Austria, are not registered in the mortality file.

^bOf 21 emigrants, seven moved to other Austrian states, four to Germany, two to former Yugoslavia, one to Italy, one to Great Britain, one to Spain, two to South America and for three cases we were not able to determine the destination.

^cIn Austria, we have two types of residence, a primary and a secondary residence. All four cases had only a secondary residence in Austria, and mortality information is provided only for persons holding a primary residence.

	Observed Survival Rate ^a		Relative Survival Rate ^a	
	Uncorrected	Corrected	Uncorrected	Corrected
One-year survival	73.4	73.2	76.0	75.9
Three-year survival	60.6	60.3	66.8	66.4
Five-year survival	53.8	53.4	63.4	62.9
Ten-year survival	41.1	40.4	58.1	57.1

Table III. Uncorrected and corrected observed and relative survival rates for patients diagnosed in year 1997 (n = 2559).

^aBoth observed and relative survival rates were calculated with the STATA procedure strs written by Paul Dickman, which is used by our registry to compute survival rates.

the exact date of first registration and the DCI status for a cancer case were recorded. We register this information carefully, because it is directly linked to the search database and thus essential for the registration procedures.

In addition to the flow method as a quantitative method, we also applied four semi-quantitative methods, namely we looked at age-specific rates of childhood cancer and compared these to reference values, looked at time trends for frequent cancer sites, compared the M:I ratio to survival estimates in the registry and compared the M:I ratio to that of neighboring countries and finally compared MV proportions to those of neighboring countries.

Childhood cancer age-specific rates are at the upper limit of reference deciles and in part exceed the upper decile. Underestimation is thus unlikely. One possible reason for high rates could be duplicates. This was carefully checked and we found no errors. Pediatricians told us that all cases are treated within clinical studies and diagnoses are crosschecked with a reference institution in Austria.

When looking at time trends, the most striking effect is seen for prostate cancer, where the agestandardized rate doubled in the early 1990s as a

Table IV. Analysis of a subset of 295 patients for re-abstracting date of diagnosis, difference between documented and corrected date of diagnosis.

Difference in months ^a	N (%)		
-2	1 (0.3%)		
-1	4 (1.4%)		
0	286 (96.9%)		
1	2 (0.7%)		
3	1 (0.3%)		
4	1 (0.3%)		

^aThe difference was defined by subtracting the corrected date of diagnosis from the uncorrected date of diagnosis. For a positive difference, this means that the corrected date of diagnosis was before the uncorrected date of diagnosis (thus underestimating the published survival rate), and for a negative difference the corrected data of diagnosis was after the uncorrected date of diagnosis (thus overestimating the published survival time).

result of intensive PSA testing; this phenomenon has been described elsewhere [21–23]. We observed that prostate cancer accounts for about one-third of all male cancer sites. Thus, prostate cancer also has a great impact on survival rates for all cancer sites taken together.

Investigation of the M:I ratio shows some differences for ovarian cancer, bladder cancer and prostate cancer by comparison with neighboring countries. Prostate cancer was described above. We observed that for females, the rates for ovarian cancer and for bladder cancer are higher than those in neighboring countries. For ovarian cancer, the proportion of borderline tumors might explain this phenomenon. It is known that borderline ovarian cancer accounts for up to one-quarter of all ovarian cancer cases [24], and indeed we noticed that in the nineties, our registry erroneously coded some ovarian cancer cases (it should be mentioned that version 1 of ICD-O involved great problems in correctly coding borderline ovarian cancer). In the meantime, this error has been corrected.

Bladder cancer is known to be strongly influenced by pathology definitions and coding errors, see for example [8]. We checked our bladder cancer cases and came to the conclusion that a coding error is unlikely. However, it is known that the one main pathology institute that performs diagnostic tests for most of our bladder cancer cases follows a rather strict rule for diagnosing this cancer (personal communication).

It is worth noting that the registry area is quite small and its population is served by not more than ten hospitals in the state. Few patients are treated in neighboring regions of Austria; these cases are traced back, because we know the most likely treating hospitals. Furthermore, Innsbruck Medical University attracts cases because of its academic status and, consequently, immigration of patients is far stronger than is emigration. Also, for certain diagnoses like head and neck, neoplasms in the hematopoietic and the lymphatic system the predominant majority of patients is treated at Innsbruck Medical University plus one or two additional hospitals. This fact
facilitates registration procedures as compared to registries covering larger regions.

In summary, some coding and/or other diagnostic problems for ovarian cancer could have produced a small bias in survival rates. With regard to bladder cancer, there seem to be some differences in pathology diagnosing procedures that are beyond the influence of the Cancer Registry. Finally, the high prostate cancer incidence rate clearly influences the survival rate for men in total, bearing in mind that about one-third of all male cancer cases are prostate cancer and a large part of these are at a very early stage with favorable prognosis.

Date of diagnosis shows minimal errors. The small effects of over- and underestimating survival caused by errors in the date of diagnosis mostly cancel out each other. Therefore, a relevant bias of survival rates caused by date of diagnosis is unlikely.

To check validity of patient survival status we investigated all patients diagnosed in 1997: of 2559 patients diagnosed in 1997, all 1026 patients alive at end of 2007 were actively followed up. These cases were traced by phoning the respective municipal office for every case. Only few cases (two) were missed by the record linkage program, and two cases that we proved to have died before end of 2007 could not be identified in the mortality files. It was interesting to learn that five cases died on holiday outside Austria, whereby there is no formal procedure for registering such cases in Austrian mortality files. About 1% of cases emigrated; this fits to data provided by Statistics Austria showing that emigration in the year 1999 was below 3% for age up to 50 and below 1% for age 50 and above [25]. The Austrian Ministry of the Interior keeps a migration database, but because of its very strict data protection rules provides no access for cancer registries. As a consequence of our study, we will enforce our efforts to obtain access to information on emigrants and to secure registration of persons who die outside Austria.

The impact of these errors on survival rates was relatively small in the subset we investigated: the absolute error was less than 0.5 for up to five-year survival rates and less than one for ten-year survival rates. This fits well with results from, for example, the Ontario Cancer Registry [26].

While this analysis involves some strengths, it also presents limitations. Possibly the most severe limitation is that we were not able to investigate all quality indicators proposed in [4,5]. Because of our limited resources, we concentrated on those indicators related to possible bias in survival rates.

Although the impact of errors resulting from emigration after cancer diagnosis was shown to be small in the analysis subset, we expect this effect to increase over the next years. About 10% of the population of Tyrol is composed of immigrants, mainly from Turkey and former Yugoslavia, and our personnel and record linkage procedures have problems with non-German-language names. In addition, part of the Tyrolean population consists of seasonal workers. We must make it a point to correctly count persons with cancer diagnosis as long-stay residents. However, most of these persons are younger and thus far not so relevant for cancer diagnosis.

There are several factors that we expect to contribute to improving data validity in the future. In recent years, Austria introduced an electronic health card system. This system is already used by medical practitioners and hospital outpatient departments and will be introduced to the inpatient departments. We expect that in a few years this system will be employed by all partners in the healthcare system and should thus provide an electronic identifier. As a consequence, errors from record linkage procedures will be avoided in future.

Conclusion

The potential for selection and information biases in survival rates in the Tyrolean cancer registry was carefully investigated. Only minor problems were identified. In total, the rate of error in the registration procedures influencing survival rates is very low and is unlikely to have caused a relevant bias in published survival rates. However, access to data on emigration, which is by now impossible because of data protection restrictions, would reduce the bias in patient survival status if we remember that the extent of cancer patient emigration in Austria is expected to increase over the next years.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- Sant M, Aareleid T, Berrino F, Bielska Lasota M, Carli PM, Faivre J, et al. EUROCARE-3: Survival of cancer patients diagnosed 1990–94–results and commentary. Ann Oncol 2003;14(Suppl 5):61–118.
- [2] Berrino F, Angelis R, Sant M, Rosso S, Bielska Lasota M, Coebergh JW, et al. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995– 99: Results of the EUROCARE-4 study. Lancet Oncol 2007; 8:773–83.
- [3] Verdecchia V, Francisci S, Brenner H, Gatta G, Micheli A, Mangone L, et al. Recent cancer survival in Europe: A 2000– 02 period analysis of EUROCARE-4 data. Lancet Oncol 2007;8:784–96.

RIGHTSLINKA)

Appendix 31

- [4] Bray F, Parkin DM. Evaluation of data quality in the cancer registry: Principles and methods. Part I: Comparability, validity and timeliness. Eur J Cancer 2008, doi:10.1016/ j.ejca.2008.11.032
- [5] Parkin DM, Bray F. Evaluation of data quality in the cancer registry: Principles and methods. Part II: Completeness. Eur J Cancer 2009, doi:10.1016/j.ejca.2008.11.0334.
- [6] Brenner H, Hakulinen T. Reduction in selective underascertainment bias in population-based estimates of cancer patient survival by age adjustment. Eur J Cancer 2005;41: 1788–93.
- [7] Brenner H, Hakulinen T. Population-based monitoring of cancer patient survival in situations with imperfect completeness of cancer registration. Br J Cancer 2005;92:576–9.
- [8] Quinn MJ, Wood H, Cooper N. Cancer atlas of the United Kingdom and Ireland 1991–2000. London: Office for National Statistics (ONS); 2005.
- [9] ENCR[Internet]. Empfehlungen zur Verschlüsselung des Diagnosezeitunktes. [cited 2. Feb 2009]. Available from: URL: http://www.encr.com.fr/incidger.pdf
- [10] Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J, editors. Cancer incidence in five continents, Volume VII. IARC Scientific Publications No. 143. Lyon: IARC; 1997.
- [11] Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B, editors. Cancer incidence in five continents, Volume VIII. IARC Scientific Publications No. 155. Lyon: IARC; 2002.
- [12] Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al., editors. Cancer incidence in five continents, Vol. IX. IARC Scientific Publications No. 160. Lyon: IARC; 2007.
- [13] Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, editors. Cancer registration principles and methods. Lyon: IARC-Press; 1991.
- [14] Leitner B. Qualitätsaspekte der Todesursachenstatistik. Statist Nachrichten 2005;9:790–7.
- [15] Oberaigner W, Stühlinger W. Record linkage in the cancer registry of Tyrol. Methods Inf Med 2005;44:626–30.

- [16] Bullard J, Coleman MP, Lutz JM, Bell J, Peto J. Completeness of cancer registration: new method for routine use. Br J Cancer 2000;82:111–6.
- [17] Silcocks PBS, Robinson D. Completeness of ascertainment by cancer registries: Putting bounds on the number of missing cases. J Public Health 2004;26:161–7.
- [18] Dickman PW[Internet]. strs Version 1.2.8. [cited 21. June 2007]. Available from: http://www.pauldickman.com/rsmo del/
- [19] Stata Statistical Software: Release 9.0. College Station, Tx: StataCorp LP. 2005.
- [20] Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. Eur J Cancer 2008, doi:10. 1016/j.ejca.2008.10.037.
- [21] Bartsch G, Horninger W, Klocker H, Reissigl A, Oberaigner W, Schonitzer D, et al. Prostate cancer mortality after introduction of prostate-specific antigen mass screening in the federal state of Tyrol, Austria. Urology 2001;58:417–24.
- [22] Oberaigner W, Horninger W, Klocker H, Schonitzer D, Stuhlinger W, Bartsch G. Reduction of prostate cancer mortality in Tyrol, Austria, after introduction of prostatespecific antigen testing. Am J Epidemiol 2006;164:376–84.
- [23] Bartsch G, Horninger W, Klocker H, Pelzer A, Bektic J, Oberaigner W, et al. Tyrol Prostate Cancer Demonstration Project: Early detection, treatment, outcome, incidence and mortality. BJU Int 2008;101:809–16.
- [24] Du Bois A, Rochon J, Lamparter C, Pfisterer J. Pattern of care and impact of participation in clinical studies on the outcome in ovarian cancer. Int J Gynecol Cancer 2005;15: 183–91.
- [25] Oberaigner W. Errors in survival rates caused by routinely used deterministic record linkage methods. Methods Inf Med 2007;46:420–4.
- [26] Hall S, Schulze K, Groome P, Mackillop W, Holowaty E. Using cancer registry data for survival studies: The example of the Ontario Cancer Registry. J Clin Epid 2006;59:67–76.

B1-5 Appendix Full Papers

Oberaigner *et al. BMC Public Health* 2010, **10**:86 http://www.biomedcentral.com/1471-2458/10/86

RESEARCH ARTICLE



Open Access

Breast cancer incidence and mortality in Tyrol/ Austria after fifteen years of opportunistic mammography screening

Willi Oberaigner^{1,2,3*}, Wolfgang Buchberger^{2,4†}, Thomas Frede^{5†}, Rudolf Knapp^{6†}, Christian Marth^{7†}, Uwe Siebert^{2,3,8,9†}

Abstract

Background: The aim of this study was to analyse breast cancer incidence and mortality in Tyrol from 1970 to 2006, namely after performing more than a decade of opportunistic mammography screening and just before piloting an organised screening programme. Our investigation was conducted on a population level.

Methods: To study time trends in breast cancer incidence and mortality, we applied the age-period-cohort model by Poisson regression to the official mortality data covering more than three decades from 1970 to 2006 and to the incidence data ranging from 1988 to 2006. In addition, for incidence data we analysed data on breast cancer staging and compared these with EU guidelines.

Results: For the analysis of time trend in breast cancer mortality in age groups 40-79, an age-period-cohort model fits well and shows for years 2002-2006 a statistically significant reduction of 26% (95% Cl 13%-36%) in breast cancer mortality as compared to 1992-1996.

We see only slight non-significant increases in breast cancer incidence. For the past five years, incidence data show a 10% proportion of in situ cases, and of 50% for cases in stages II+.

Conclusions: The opportunistic breast cancer screening programme in Tyrol has only in part exploited the mortality reduction known for organised screening programmes. There seems to be potential for further improvement, and we recommend that an organised screening programme and a detailed screening database be introduced to collect all information needed to analyse the quality indicators suggested by the EU guidelines.

Background

Breast cancer (BC) is the leading cause of female cancer death in all industrialised countries (and also worldwide) and the breast is also the leading incident cancer site for females [1]. Therefore, screening methods for BC are of greatest public health importance. Efficiency and efficacy of organised mammography screening programmes have been proven in large randomised trials conducted in Europe and North America. For several years already, organised mammography screening programmes have been recommended in the EU[2]. Austria is one of the European countries where up to 2006 no organised

+ Contributed equally

programmes were implemented, but where coverage in spontaneous mammography screening could have been rather high. In a micro-census conducted in Austria in 2006-2007, more than 80% of women aged 40-59 answered that they had had at least one mammography (ever) and more than 40% had had one in the past year [3]. However, it is known that self-reporting of screening usage overestimates true coverage [4], and first preliminary data from the organised mammography screening programme in Tyrol strongly confirm this interpretation. In 2006, the Austrian health minister declared mammography to be one of the top health agendas, and in July 2006 a decision was made to implement organised mammography screening programmes, in a first step in pilot regions, of which Tyrol is the largest.



© 2010 Oberaigner et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

^{*} Correspondence: willi.oberaigner@iet.at

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd, Cancer Registry of Tyrol, Innsbruck, Austria

In Tyrol, spontaneous mammography screening was set up around 1993. In July 1998 the "Working Group for Early Breast Cancer Detection for Tyrol" was established. Since that year recommendations including monthly breast self-examination, annual examination by a physician and annual mammography, if necessary with adjunctive breast sonogram, beginning at age 35-40, were formulated. Assessment is offered centrally by eight hospitals. Results of this strategy have been published [5]. Generally, in Austria spontaneous mammography screening is offered in the framework of general health exams done by general practitioners and of gynaecologic exams performed by gynaecologists in private practice. Women are referred for screening mammography mostly to radiologists in private practice. Both services are free of charge for women as of age 40. In 2007 and 2008, five pilot projects were launched in several of Austria's federal states in order to evaluate how to implement organised mammography screening. In Tyrol, a state screening programme was started in June 2007. In a one-year pilot phase, the methods were tested in two counties of Tyrol and in June 2008 the programme was extended to the whole state of Tyrol. The basic goal was to smoothly change the existing opportunistic screening system. Main characteristics are personal invitation, screening offered by radiologists and hospitals (out-patient departments), assessment at two hospitals, training of all partners and careful quality control by collecting all data in a central screening database and periodic inspection of data by a medical quality assurance group. No double reading was implemented.

We feel there is a need to publish the baseline characteristics of incidence and mortality in order to give a transparent public view of the situation in Tyrol before changing the mammography system. Although programme characteristics have not been collected to date, we can roughly judge the outcome achieved with the former spontaneous system by analysing time trends in incidence and mortality and by looking at stage shifts in BC cases. To our knowledge, it is not only in our country that spontaneous mammography screening is offered to women broadly, and there is ongoing discussion about whether the mammography system should be changed [6]. Therefore, it is of general interest to analyse the effects of spontaneous mammography screening offered free of charge to all women in a population. The analysis was only to be published now, because mortality and incidence data for female BC in Tyrol were published just a few months ago for the period to 2006 [7].

It was our aim to analyse BC incidence and mortality before changing the mammography system in Tyrol and to estimate the effects of the spontaneous programme offered free of charge to women for about fifteen years in order to have a public discussion of results before making the decision on whether and how to change the mammography system in Tyrol.

Methods

Mortality Data

Mortality data are collected by Statistics Austria for the whole of Austria [8]. In Austria, death certificates are issued by official, specially trained medical doctors, pathologists and forensic medical experts. Specialists at Statistics Austria, the federal institution for statistics in Austria, follow international guidelines and select one main diagnosis that led to death and assign it one ICD code (ICD9 up to 2001, ICD10 since 2002). All procedures concerning death certificates, data collection and coding are applied in a uniform way throughout Austria and are not state-specific. We analysed all female cases coded for cause of death BC as described above.

Incidence Data

Incidence data have been collected by the Cancer Registry of Tyrol since year of diagnosis 1988 on a population-based perspective. Publication of incidence data in Cancer Incidence in Five Continents gives some hints for good completeness and validity of the database [9,10]. Registration is performed from a standardised questionnaire including sex, age, cancer site and histology, date of diagnosis, stage and basic information on primary treatment. Information on co-morbidity is not collected routinely. There are strict rules for collecting these variables in accordance with international guidelines. Either the questionnaire is completed by a physician or a Cancer Registry clerk collects data directly from clinical records in the treating hospital.

Modelling of time trends

Time trends were analysed by fitting age-period-cohort (APC) models [11,12]. APC models allow separate effects to be estimated for age (A), period or year of death (P) and cohort (C) by means of Poisson regression. In a more formal sense we fit a series of models

$$\log(\rho_{\rm APC}) = \alpha_{\rm A} + \beta_{\rm P} + \gamma_{\rm C}$$

where C=P-A, and ρ , denotes the mortality rate The model is often written in antilogs as follows:

$$\rho_{\rm APC} = \alpha'_{\rm A} \beta'_{\rm P} \gamma'_{\rm C}$$

where α_A ' denotes the antilog of α_A or α_A ' = exp(α_A) etc.

As suggested by Clayton and Schifflers, a series of models is fit until adequate model fit is attained. We start with A alone and proceed by including P and/or C in the model if model fit is not sufficient without the extra term and inclusion of the term substantially improves goodness of fit. Goodness of fit is measured by deviance, which should be equal or close to the degrees of freedom (DF) if model fit is reasonably good.

For statistical analysis the number of BC deaths was aggregated in ten-year age groups and five-year period groups. We started with age group 40-49 and continued in ten-year age groups ending with age group 70-79, because we expected mammography screening to also affect this age group, bearing in mind that women aged 60-69 at the beginning of the screening programme around 1993 are now in age group 70-79. We have access to mortality data beginning in 1970. Our first period group was 1972-1976 in order to finish with the five year group 2002-2006. Our hypothesis was that the mortality rate decreased following the introduction of mammography screening around 1993. Thus, the reference category was the period 1992-1996. Consequently, because C = P-A, cohort groups begin with a cohort group centred at 1899.

To analyse the incidence time trend, we fitted an APC model for age groups 40 to 69, namely the age groups aimed at by the screening programme. The incidence data set begins with 1988. Therefore, we defined period groups 1988-1991 and then five-year period groups ending with 2002-2006.

For incidence data, we also analysed the proportion of in situ cases and the proportion of stages according to UICC and compared these with accepted levels given by EU guidelines [2].

Age-specific rates were calculated using official population numbers as denominators. Population data are also collected by Statistics Austria. Census data are available for the years 1971, 1981, 1991 and 2001; for intercensus years population figures are extrapolated based on birth, death and migration information. The female population of Tyrol in the census year 2001 was 345,757. The analysis was performed with Stata, Version 9; the APC model was set up using the procedure poisson for Poisson regression [13].

This study was conducted in conformity to the Helsinki Declaration [14].

Results

For an impression of overall BC mortality and incidence, Figure 1 shows the time trend in age-standardised mortality and incidence rates (for all age groups); age standardisation was based on world population proposed by SEGI and modified by Doll et al. [9]. The line of moving averages suggests a decline in mortality since 1998 and an increase in incidence until 2003, however on a purely descriptive level.

For a formal analysis of time trend, we fitted an APC model separately for mortality and incidence. For

mortality, the final model includes terms for period and cohort because there are statistically significant cohort effects, model fit is very good (8 degrees of freedom, deviance 3.5). The resulting estimates for the APC model are described in Table 1 and Figure 2. Age effects, each compared to age group 40-49, are 2.15 for age group 50-59, 3.67 for age group 60-69 and 5.75 for age group 70-79. Period effects, each compared to 1992-1996, are about 1.05 before 1992, but 0.83 (95% CI 0.57, 1.21) for 1997-2001 and 0.74 (95% CI 0.64, 0.87) for 2002-2006. In general, the effects we report can be interpreted a change in mortality compared to the reference period, for example the effect of 0.83 for year 1997-2001 means a mortality reduction of 17% in 1997-2001 as compared to 1992-96. We also observe a strong cohort effect for cohorts born around 1920 and between 1930 and 1950 with relative risks at 1.4-1.8, each compared to the cohort centred at 1899.

For incidence, the time period from 1988 to 2006 is much shorter. We modelled the time trend for age groups for which mammography screening was recommended, namely 40 to 69. The AP model reaches sufficient model fit with 8 degrees of freedom and a deviance of 8.3. Since adding a cohort parameter does not cause a significant improvement, we accepted the AP model. Period effects, each compared to 1992-1996, show a non-significant increase in BC incidence up to 1992, a slight but non-significant increase of 1.05 after 1996 and a steady situation during the last five years, see Table 2.

In addition, we also analysed some of the quality indicators proposed by EU guidelines [2]. The proportion of in situ cancers out of the total in situ and invasive cancers shows a steady increase from 5% around 1990 to 13% around 2000 and a slight decrease to 10% in recent years (see Figure 3). This time trend is consistent for all three age decades investigated (data not shown) and meets the 10% acceptable level given by EU guidelines. Figure 3 shows staging groups according to UICC. We see a clear stage shift towards early stages I and II up to around 2000 and a slight decrease afterwards. The EU acceptable proportion of stages II+ (30%) is clearly missed; in the last years the proportion of stages II+ in Tyrol was about 50%.

A more detailed analysis for the last five years shows a proportion of very small cancer with a size of less than 1 cm (TNM staging T1a,b) of 24% (age group 40-49), 22% (age group 50-59) and 19% (age group 60-69), the acceptable level according to EU guidelines being 25%. However, this information is not available for the 1990s.

The proportion of node-negative cases also increased to 58% (age group 40-49) and 53% (age group 50-69) at about 2000, the acceptable proportion in EU guidelines being 70%.



Applying a COX model to analyse the effect of adjuvant hormonal therapy resulted in a hazard ratio of 0.87 (95% CI 0.77, 0.99) adjusted for age and stage, this means that patients receiving adjuvant hormonal therapy have a 13% lower risk of death than patients without this therapy.

Discussion

Our results indicate a significant decrease in BC mortality over the past five years, a nonsignificant slight increase in incidence and a stage shift towards early stages over the past fifteen years with some proportions in the range of the accepted levels given by EU guidelines.

Strengths and Limitations

This observational study was conducted in the population of Tyrol. Mortality and incidence data were collected on a population level. Mortality data were provided by Statistics Austria. The quality of death certificates was very important for the conclusions drawn. In general, mortality statistics in Austria have been of high quality for decades [8]. Coding of cause of death is done according to international guidelines by specialists who attend international benchmarking exercises. As already stated above, death certificates are written by specially trained doctors. Data on BC incidence on a population level are provided by the Cancer Registry of Tyrol, which is a member of IACR and whose data are published in Cancer Incidence in Five Continents, thus giving some evidence for good quality of incidence data [9,10,15]. Figures on completeness of incidence data show that for BC, in the past decade the proportion of death certificate-notified cases was 3.2% and the proportion of death certificate only cases 1.4% [7]. Both proportions allow the conclusion that completeness is good as compared to international data.

In addition, the proportion of cases with unknown or unspecified stage is less than 5% in age groups 40 to 69. When we analyse incidence data on a population level, we always encounter some cases that lack detailed information for various reasons. Since year of diagnosis 2004, the cancer registry includes a variable for mode of detection. However, the information is very incomplete, because in many cases we cannot obtain sufficient information from the hospital discharge records.

The model we fitted for analysis of the time trend shows very good model fit. This means we can trust the time trend parameters and can therefore draw reliable conclusions from the model. Moreover, the staging information used to describe stage shift should be reliable.

Table 1 Model estimators for age, period and cohort	
given by the APC model, drift in period, for breast	
cancer mortality in Tyrol 1972-2006	

	Estimator	95% CI	
Age			
40-49	1	Refe	erence
50-59	2.15	1.85	2.50
60-69	3.67	3.05	4.42
70-79	5.75	4.58	7.22
Period			
1972-1976	1.12	0.92	1.37
1977-1981	1.07	0.79	1.44
1982-1986	1.07	0.91	1.26
1987-1991	1.05	0.74	1.48
1992-1996	1	Refe	erence
1997-2001	0.83	0.57	1.21
2002-2006	0.74	0.64	0.87
Cohort (centred at)			
1899	1	refe	erence
1904	Collinearity		
1909	1.06	0.84	1.35
1914	1.20	0.92	1.56
1919	1.41	1.11	1.77
1924	1.32	0.98	1.77
1929	1.27	0.99	1.62
1934	1.60	1.13	2.28
1939	1.61	1.21	2.15
1944	1.80	1.18	2.74
1949	1.48	1.04	2.09
1954	1.55	0.90	2.66
1959	Drift		

* Because there is drift in period, there is no estimator for the last cohort centered at 1959

The main limitation is the lack of a screening database. Consequently, we do not have detailed information on screening performance parameters. Another weakness is that we have only some limited information on coverage from a micro-census performed in 2006/2007 and from a publication by Frede [5]. Both sources indicate a coverage of 70%. In Catalonia, Spain, Baré et al [16] reported also a coverage of 70% prior to introducing a screening programme and the authors investigated reasons for non-participation which can be very helpful in improving coverage. On the other side, it is known that self-reporting overestimates true coverage, [4] and a more realistic estimation could be a coverage of about 50%. This would fit to first preliminary data from the organised programme in Tyrol (data not shown). The lack of information on coverage and, of course, also on many other screening details was one of the reasons for changing the screening system, because we are convinced that a detailed knowledge of screening parameters is essential to draw valid conclusions in future. For staging distribution, the only source of information is the Cancer Registry dataset, whose focus was not to obtain information on screening indices but to concentrate on cancer cases.

Time trend for mortality and incidence data, model fit

We applied an APC model that takes age, period and cohort effects into account and models time trends that differ from a linear trend. Such models are widely used in epidemiology, see for example [17,18]. Each of the models we applied for both mortality and incidence fits well on its own, and all parameters allowing judgment of model fit are reasonably good. Also, the graphs showing observed and predicted rates give additional evidence that the model describes the data very well and hence that we can rely on estimated parameters (graphs not shown). In summary, the time trends given by the models should adequately describe the situation we observe.

Concerning the decrease in BC incidence in recent years, Ravdin et al. [19] hypothesized for the USA that the reduction in hormone replacement therapy (HRT) is the main cause of the rapid decrease in BC incidence seen in the USA from 2003 to 2004. Also in Tyrol, we observe a decrease in BC incidence only in age groups 50+, although the decrease is not as sharp as in the USA. According to local experts, it is likely that also in Tyrol, part of the decrease in breast cancer incidence in the age group 50+ between 2004 and 2006 is due to a reduction in HRT.

The main question remains whether the significant 26% reduction in BC mortality over the past five years as compared to 1992-1996 is associated with opportunistic mammography screening. Both randomised trials and data from population-based organised mammography screening programmes provide clear evidence that organised mammography screening can reduce BC mortality. This was also communicated at an IARC international expert conference [20]. The extent of mortality reduction differs in detail, but in general is estimated to be between 20% and 25% [21-32]. However, for population-based organised programmes it is necessary to distinguish between various factors influencing BC mortality [33]. Some authors[33] estimate that a great part of BC mortality reduction (approximately 2/3 of reduction in England and Wales) is related to improvements in therapy, mainly the introduction of tamoxifen. For the USA, Berry et al. [34] found a range of 28% to 65% (median 46%) for the proportion of BC mortality reduction attributed to screening by modelling this proportion by seven independent investigators. For Tyrol, this would imply a mortality reduction of 7% to 17% attributable to screening. In addition, when comparing BC mortality trends between countries, stage



Table 2 Model estimators for age and period given bythe AP model, for breast cancer incidence in Tyrol 1988-2006

	Estimator	95% Cl	
Age			
40-49	1	Refe	rence
50-59	1.65	1.53	1.79
60-69	2.08	1.93	2.25
Period			
1988-1991	0.92	0.83	1.01
1992-1996	1	Refe	rence
1997-2001	1.05	0.96	1.14
2002-2006	1.05	0.96	1.14

Remark: There is no significant cohort effect. Therefore, the model was set up without cohort terms.

distribution and differences in therapy also have to be discussed as factors influencing BC mortality at the population level.

Adjuvant therapy with tamoxifen was routinely introduced in Tyrol around 1985. We do not collect detailed information on BC therapy in the Cancer Registry, but we have an overall variable for adjuvant hormonal therapy. When we analysed the effect of adjuvant hormonal therapy in a COX model adjusted for age and stage, an overall effect of 13% was seen. With regard to time trend in survival rates, over the past fifteen years we observed an increase in relative five-year survival rates split by staging groups according to UICC (5% increase in stage I, 13% in stage II, and 5% in stage IV). Both observations are consistent with an estimated therapy effect on survival of between 10% and 15%, which is in line with the UK estimate [21]. Furthermore, as compared with EU guidelines, we miss some of the accepted levels (coverage, proportion of small cancers, proportion of II+ cancers, proportion of node-negative cancers). In conclusion, we estimate that less than half of the mortality reduction should be due to screening. This would mean that the screening effect is less than 13% and that, consequently, the opportunistic screening programme does not realise the potential of organised programmes, namely a mortality reduction of 20% - 25%.

However, when we compare BC data for Tyrol with quality indicators for mammography screening, we must remember that the BC data we analysed included all BC cases diagnosed in the population of Tyrol, not only those detected by opportunistic mammography screening. For example, Paci et al. [35] show a proportion of 53% for II+ breast cancer in the total population as opposed to 29% in the screen-detected subgroup.

Vutuc et al. recently analysed BC mortality in Austria [6]. The authors argue that BC screening is a plausible explanation for BC mortality reduction and doubt that a change in screening policy (meaning changing from opportunistic screening to an organised programme) would significantly improve the situation in Austria. We agree that BC screening is indeed one possible explanation for BC mortality reduction. However, if we take into consideration the fact that we have no detailed



information on diagnostic performance or coverage for opportunistic BC screening in Austria, we feel it is absolutely necessary that detailed information on mammography screening be collected, at least for several years. We need to know all the well-established quality indices for BC screening[2] before we can draw a final conclusion on how to proceed with mammography screening in Austria.

Interestingly, the greatest reduction in BC mortality was observed in the age group 40 to 49. This differs somewhat from international data, where doubts still prevail on the efficacy of mammography screening in the age groups below 50, see for example [36,37]. Surrogate performance indicators like stage shift, cancer size less than 1 cm and proportion of node-negative cancer also showed a clear tendency towards better performance in the age group 40 to 49 as compared to the age groups 50 to 59 and 60 to 69. In addition, during the past decade, these indicators improved more quickly in the age group 40 to 49 (details not shown). One possible explanation is the wide-spread use of sonography as an adjunct to mammography in Tyrol. It has been shown by various authors that the additional use of sonography can improve cancer detection rates, especially in younger women and women with dense breasts. The relative percentage of carcinomas found in supplemental breast ultrasound examinations as a fraction of the total number of detected cancers was reported by four studies, with a mean percentage of 22.5% (15%-34%) [38].

In opportunistic screening in Tyrol, sonography was offered to women with dense breasts (ACR density grades 3 and 4) and with inconclusive findings on mammography [39]. In addition, women in the younger age groups are likely to go more frequently to their general practitioner or gynaecologist, which results in higher coverage by opportunistic screening [3].

The discussion in the USA after publishing the revised recommendation by U.S. Preventive Services Task Force [40,41] shows that it is very challenging and hard to understand by women to remove a service that was recommended for several years. Without well founded data, we feel it is not justified to stop screening in age class 40-49. We are collecting detailed data and will evaluate the balance between goods and harms during the next years.

Some of the EU recommendations like double reading and making an appointment for mammography when inviting women will not be part of the organised programme in Tyrol. Thus, further investigation will be needed to prove whether mammography screening has an effect on BC mortality, even in the absence of these EU recommendations.

Conclusions

Up to now, in terms of BC mortality reduction our analysis shows that it is likely that the full potential of mammography screening has not yet been realised. In addition, available cancer registry data are not sufficient to assess the efficiency or efficacy of the current opportunistic screening programme. Therefore, to analyse surrogate indices like decrease in advanced stages and increase in early stages, interval cancer rates, and to investigate the cost efficiency of the established programme, it is absolutely necessary that a well-organised screening database be built up that contains all information needed to analyse the quality indicators suggested by the EU guidelines. In conclusion, we strongly advise that an organised mammography screening programme be introduced in Tyrol, namely one that will also allow a detailed analysis of the effects of mammography screening.

List of Abbreviations

CI: confidence interval; AP: age-period model; APC: ageperiod-cohort model; BC: breast cancer; HRT: hormone replacement therapy.

Acknowledgments

This work was supported by the ONCOTYROL Center for Personal Cancer Medicine. ONCOTYROL is a COMET Center funded by the Federal Ministry for Transport Innovation and Technology (BMVIT) and the Federal Ministry of Economics and Labour (BMWA), the Tyrolean Future Foundation (TZS) and the Land of Styria represented by the Styrian Business Promotion Agency (SFG) [and supported by UMIT - University for Health Sciences, Medical Informatics and Technology, Innsbruck Medical University, Tyrolean Health Insurance Fund and the Tyrolean Health Company.

Author details

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd, Cancer Registry of Tyrol, Innsbruck, Austria. ²Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Information Systems and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall iT, Austria. ³Oncotyrol - Center for Personalized Cancer Medicine, Innsbruck Medical University, Innsbruck, Austria. ⁴University Hospital, Innsbruck, Austria. ⁵Innsbruck Medical University, Department of Radiology, Kufstein, Austria. ⁷Innsbruck Medical University, Department of Radiology, Kufstein, Austria. ⁷Innsbruck Medical University, Department of Obstetrics and Gynecology, Innsbruck, Austria. ⁸Program in Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA. ⁹Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

Authors' contributions

WO designed the study, performed the analysis and wrote the paper. US contributed to study design and writing of the paper. WB assisted in writing the paper, especially contributions on ultrasound and screening at age below 50. TF und RK contributed to the discussion part, especially from the radiology point of view. CM contributed to the methods, report and discussion part, especially from the gynaecological point of view. All authors reviewed and agreed to the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 20 May 2009 Accepted: 20 February 2010 Published: 20 February 2010

References

- Parkin DM, Bray F, Ferlay J, Pisani P: Global cancer statistics, 2002. CA Cancer J Clin 2005, 55:74-108.
- Perry N, Broeders M, de Wolf C, (eds), *et al*: European guidelines for quality assurance in breast cancer screening and diagnosis. Luxembourg: Office for Official Publication of the European Communities, Fourth 2006.
- Österr. Gesundheitsbefragung 2006/07. Wien: Statistik Österreich 2008, (In German: Health Survey in Austria 2006/07).
- Cronin KA, Miglioretti DL, Krapcho M, Yu B, Geller BM, Carney PA, et al: Bias associated with self-report of prior screening mammography. Cancer Epid Biom Prev 2009, 18(6):1699-705.
- Frede TE: Opportunistic breast cancer early detection in Tyrol, Austria 1996-2004. Is a mammography-screening program necessary? Eur J Radiol 2005, 55(1):130-8.
- Vutuc C, Waldhoer T, Haidinger G: Breast cancer trends: opportunistic screening in Austria versus controlled screening in Finland and Sweden. Eur J Cancer Prev 2006. 15(4):343-6.
- Oberaigner W, Mühlböck H, Harrasser L: Tumorregister Tirol Bericht für die Diagnosejahre 2004/2006. Innsbruck: IET-Bericht 2008, (In German).
- 8. Leitner B: Qualitätsaspekte der Todesursachenstatistik. Statist Nachrichten 2005, 9:790-7, (In German).
- 9. Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B: Cancer Incidence in Five Continents. Volume VIII. Lyon: IARC 2002.
- Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue J, Boyle P: Cancer Incidence in Five Continents. (IARC Scientific Publications No. 160). Lyon: IARC 2007, IX.
- 11. Clayton D, Schifflers E: Models for temporal variation in cancer rates. I: Age-period and age-cohort models. *Stat Med* 1987, 6(4):449-67.
- Clayton D, Schifflers E: Models for temporal variation in cancer rates. II: Age-period-cohort models. Stat Med 1987, 6(4):469-81.
- Stata Statistical Software. Release 9.0. College Station, TX: StataCorp LP 2005.
- World Medical Association: World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects.http://www.wma.net/en/30publications/10policies/b3/index.html, (accessed 7 November 2009).
- 15. Oberaigner W, Siebert U: Are survival rates for Tyrol published in the Eurocare studies biased? *Act Oncol* 2009, **48**(7):984-91.
- Baré ML, Montes J, Florensa F, Sentís M, Donoso L: Factors related to nonparticipation in a population-based breast cancer screening programme. *Eur J Cancer Prev* 2003, 12(6):487-94.
- 17. Bray F, Moller B: Predicting the future burden of cancer. Nat Rev Cancer 2006, 6(1):63-74.
- 18. Holford TR: Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annu Rev Public Health* 1991, **12**:425-57.
- Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, *et al*: The Decrease in Breast-Cancer Incidence in 2003 in the United States. NEM 2007, 356(16):1670-4.
- 20. IARC Handbooks of Cancer Prevention: Breast Cancer Screening. Lyon: IARC-Publications 2002, 7.
- Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ: Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. *BMJ* 2000, 321(7262):665-9.
- Otto SJ, Fracheboud J, Looman CWN, Broeders MJM, Boer R, Hendriks JHCL, et al: Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality. A systematic review. Lancet 2003, 361:1411-17.
- Jonsson H, Nyström L, Törnberg S, Lenner P: Service screening with mammography of women aged 50-69 years in Sweden: effects on mortality from breast cancer. J Med Screen 2001, 8:152-60.
- 24. Duffy SW, Tabar L, Chen HH, Holmqvist M, Yen MF, Abdsalah S, *et al*: **The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties.** *Cancer* 2002, **95**:458-69.
- Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE: Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002, 359:909-19.

- Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen MB, *et al*: Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. *BMJ* 2005, 330(7485):220.
- 27. Gotzsche PC, Nielsen M: Screening for Breast cancer with Mammography. Cochrane Database of Systematic Reviews 2006, 4: CD001877.
- Anttila A, Koskela J, Hakama M: Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. J Med Screen 2002, 9(4):153-8.
- 29. Blanks RG, Moss SM, Patnick J: Results from the UK NHS breast screening programme 1994-1999. J Med Screen 2000, 7(4):195-8.
- Hakama M, Pukkala E, Heikkila M, Kallio M: Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. *BMJ* 1997, 314(7084):864-7.
- The Swedish Organised Service Screening Evaluation Group: Reduction in Breast Cancer Mortality from Organized Service Screening with Mammography: 1. Further Confirmation with Extended Data. Cancer Epidem Bio Prev 2006, 15(1):45-51.
- Törnberg S, Carstensen J, Hakulinen T, Lenner P, Hatschek T, Lundgren B: Evaluation of the effect on breast cancer mortality of population based mammography screening programmes. J Med Screen 1994, 1:184-7.
- Peto R, Boreham J, Clarke M, Davies C, Beral V: UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years. *Lancet* 2000, 355(9217):1822.
- Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al: Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer. NEJM 2005, 353(17):1784-92.
- Paci E, Duffy SW, Giorgi D, Zappa M, Crocetti E, Vezzosi V, et al: Quantification of the effect of mammographic screening on fatal breast cancers: The Florence Programme 1990-96. Br J Cancer 2002, 87(1):65-9.
- Baker S, Wall M, Bloomfield A: Breast cancer screening for women aged 40 to 49 years-what does the evidence mean for New Zealand? N Z Med J 2005, 118(1221):U1628.
- 37. Kolb TM, Lichy J, Newhouse JH: Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002, 225(1):165-75.
- Nothacker M, Duda V, Hahn M, Warm M, Degenhardt F, Madjar H, Weinbrenner S, Albert US: Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. BMC Cancer 2009, 9:335.
- Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dunser M: Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. Semin Ultrasound CT MR 2000, 21(4):325-36.
- Mandelblatt JS, Cronin KA, Bailey S, et al: Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009. 151(10):738-47.
- 41. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009, **151(10)**:716-26.

Pre-publication history

The pre-publication history for this paper can be accessed here: http://www. biomedcentral.com/1471-2458/10/86/prepub

doi:10.1186/1471-2458-10-86

Cite this article as: Oberaigner *et al.*: **Breast cancer incidence and** mortality in Tyrol/Austria after fifteen years of opportunistic mammography screening. *BMC Public Health* 2010 **10**:86.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit

BioMed Central

Appendix 42

pendix 42

B1-6 Appendix Full Papers

European Journal of Public Health, Vol. 21, No. 3, 387–391 © The Author 2010. Published by Oxford University Press on behalf of the European Public Health Association. All rights reserved. doi:10.1093/eurpub/ckq099 Advance Access published on 22 July 2010

Do women with cancer have better survival as compared to men after adjusting for staging distribution?

Willi Oberaigner^{1,2,3}, Uwe Siebert^{2,3,4,5}

- 1 Cancer Registry of Tyrol, Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd, Innsbruck, Austria
- 2 Department of Public Health, Information Systems and Health Technology Assessment, Institute of Public Health, Medical Decision Making and Health Technology Assessment, UMIT, University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria
- 3 Oncotyrol, Center for Personalized Cancer Medicine, Innsbruck, Austria
- 4 Department of Health Policy and Management, Center for Health Decision Science, Harvard School of Public Health, Boston, MA, USA
- 5 Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Correspondence: Willi Oberaigner, Cancer Registry of Tyrol, TILAK GmbH, Anichstrasse 35, Innsbruck, Austria, tel: +43-512-504-22310; fax: +43-512-504-22315, e-mail: willi.oberaigner@iet.at

Received 9 January 2009, accepted 28 June 2010

Background: Gender aspects in medicine are receiving increasing attention, namely also in oncology. For this reason, we decided to investigate whether for solid cancer sites women have better survival outcome than do men in the population of Tyrol, Austria. **Methods:** We conducted an observational population-based study in Tyrol. All solid cancer sites excluding non-melanoma skin cancer and sex-specific sites were analysed in total and all specific sites with more than 500 patients in the analysis. By the end of 2006, follow-up was ended. We applied a relative excess risk model, thus correcting for differences in life expectancy between women and men. **Results:** For all cancer sites combined, after adjusting for case mix, women had a relative excess risk of 0.95 (95% CI 0.91–0.99). For the following sites our analysis resulted in a relative excess risk statistically different from 1, namely for women as compared to men: head and neck without larynx 0.72 (95% CI 0.56–0.93), stomach 0.86 (95% CI 0.75–0.97) and lung 0.82 (95% CI 0.75–0.90). **Conclusion:** In a healthcare system with free access to diagnostics and therapy, after adjusting for staging distribution female cancer patients have a lesser excess mortality risk than do men for lung, stomach and head and neck cancer and also for all cancer sites combined after adjusting for case mix.

Keywords: cancer, gender, population based, prognosis, survival

.....

Introduction

G ender aspects in healthcare systems are receiving increasing attention, but are still not adequately recognized either in terms of public health or of diagnosis and treatment. Already in the 1990s evidence for poorer survival in women was demonstrated for cardiovascular diseases.¹ On the other hand, some evidence shows better survival of female oncology patients: for some cancer sites like lung cancer better survival in women is well established, while for other cancer sites the picture is still unclear. Also, for all cancer sites combined, investigations have shown better survival for women than for men. However, the relative risk depends on age and diminishes or even reverses for older age groups.^{2,3} Some authors conclude that a detailed investigation of gender aspects could lead to improvement in treatment.^{4,5}

Of course, in order to establish gender as a prognostic factor, we must take into consideration that cancer is a multi-factorial system. For some sites like stomach or colorectum, women are older at the time of diagnosis than are men. Furthermore, we know that in nearly all countries women have a longer life expectancy than do men.^{6,7}

Properly dealing with all these factors poses methodological challenges. We know that the method of relative survival

properly accounts for differences in age structure and in life expectancy.^{8–10} However, these models are not so well recognized, and many publications base their analysis on Cox models.

Our main goal was to analyse survival differences between women and men for the main solid cancer sites by applying a model that adjusts for the main factors registered in a cancer registry. The analysis was performed using the incidence data set for Tyrol for years of diagnosis from 1988 (when our cancer registry was started) to 2003.

Methods

Incidence data for the population of Tyrol are collected by the Cancer Registry of Tyrol. The cancer registry was established in 1986, and data have been registered on a population basis since 1988. Also since 1988, our registry data have been published data in Cancer Incidence in Five Continents,^{11–13} thus giving evidence of good completeness of the incidence data.

Assessment of patient life status is passive. We do a probabilistic record linkage between incidence data and the official mortality data set for Tyrol collected by Statistics Austria.^{14,15}

We analysed all patients with solid cancer cases in the incidence data set for Tyrol with year of diagnosis from 1988 to 2003, $N=43\,987$. DCO cases (N=1945), cases found at autopsy (N=494) and non-melanoma skin cancer (NMSC) cases (N=2969) were excluded. In addition, for eight cases we could not identify patient status; these cases were also excluded from the analysis. Analysis was restricted to adult patients (defined as age ≥ 20 years). For patients with multiple cancer sites, only the chronologically first cancer was included in the analysis. Finally, we excluded sexspecific cancer sites. We ended up with an analysis data set of 21 102 cases. Closure of this study was end of 2006.

We analysed all sites for which we had at least 500 cases. The number of 500 derives from applying rules of thumb proposed by Harrell et al.¹⁶ For an overview of all solid cancer sites, we aggregated all solid cancer sites without gender-specific sites (this means we excluded the female cancer sites ovary, cervix, corpus and other gynaecological sites and the male sites prostate, testis and other male genital system). In addition, because breast cancer is predominantly a female cancer and we observe in our population only up to three cases per year in males, we also excluded male breast cancer from the analysis. Gender-specific cancer sites contribute, of course, to overall cancer incidence and cancer survival. But in order to analyse whether women share better survival than men, it is in our opinion best to restrict the analysis to cancer sites occurring in both sexes in order to have comparable settings. In the analysis for all cancer sites combined in one group, we also adjusted for case mix.

Gender difference in survival was modelled using a relative excess risk model (RER).¹⁷ In a first step, relative survival rates are computed using a stata procedure strs provided by Paul Dickman.¹⁸ These relative survival rates are then modelled using a generalized linear model. In more formal terms: the hazard-function $\lambda(t,x)$ for a patient with characteristics x at time t is estimated as the sum of a baseline hazard $\lambda^*(t,x)$ and a so-called excess hazard $\nu(t,x)$. It is assumed that the excess hazard is a product of covariates x_1 to x_n , here written as $\exp(x\beta)=e^{x\beta}$. This means:

$$\lambda(t,x) = \lambda^*(t,x) + e^{x\beta} = \lambda^*(t,x) + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n}.$$
(1)

We applied the Hakulinen-Tenkanen model using the stata code proposed by Paul Dickman.¹⁸ For all models, we included follow-up time in the model, and the analysis was restricted to the first 5 years of follow-up because it is usually inappropriate to assume proportional hazard assumption on longer follow-up periods. All relative excess risks given by the respective RER model are for women compared to men as the reference group. For short, we use the notation RER for women. Models were built separately for every cancer site. We started with a kind of full model with terms for gender, year of follow-up, four age categories, two period categories, stage and histological verification. Cases with stage unknown remained in the analysis, whereby unknown stage was explicitly categorized. We then dropped terms if they were not statistically significant; significance was tested using the likelihood ratio test. Afterwards, if model fit was not good, we added interaction terms if the respective term had a statistically significant effect. Model fit was assessed by deviance and Pearson residuals, divided by degrees of freedom. Confidence intervals (CIs) for estimators were computed based on standard errors given by observed information matrix, the standard Stata option.

Data on life expectancy were provided in routine statistics published by Statistics Austria and the Department of Statistics in Tyrol.

All computations were performed with Stata Version 9.19

Results

In the following paragraphs, we describe in brief some basics of patient characteristics for every cancer site investigated. Details of patient characteristics are given in table 1 and univariate and multivariate relative risks with information on model fit in table 2.

We analysed a total of 941 head and neck cancer cases (without larynx), one-quarter of which were in women; see table 1. Mean age was 60 years and there were only small differences in age structure. There are distinct differences in staging distribution: the proportion of early Stage I was 18% for women and 13% for men and the proportion of Stage IV was 20% for women and 38% for men. RER for women was 0.57 in univariate analysis and 0.72 (95% CI 0.56–0.93) in multivariate analysis, see table 2.

We analysed a total of 2418 stomach cancer cases, 47% of which were in women; see table 1. Female cases were older in the mean (72 vs. 69). We found no differences in staging distribution. RER for women was 0.97 in univariate analysis and 0.86 (95% CI 0.75–0.97) in multivariate analysis, see table 2.

We analysed a total of 4519 colorectal cancer cases, half of which were in women (51%); details are shown in table 1. Female cases were older in the mean (71 vs. 67). RER for women was 1.16 in univariate analysis and 1.06 (95% CI 0.95–1.18) in multivariate analysis; see table 2.

We analysed a total of 951 pancreatic cancer cases, slightly more than half of which were in women (54%); details are shown in table 1. Female cases were older in the mean (73 vs. 67). About one-third of cases had no staging information; Stage IV accounted for 44% of men and 37% of women. RER for women was 1.06 in univariate analysis and 0.96 (95% CI 0.78–1.19) in multivariate analysis; see table 2.

We analysed a total of 3742 lung cancer cases, about one-quarter of which were in women (26%); details are shown in table 1. There were only minor differences in age structure. We observed only small differences in staging distribution, however for one-quarter of the cases stage was unknown. RER for women was 0.87 in univariate analysis and 0.82 (95% CI 0.75–0.90) in multivariate analysis; see table 2.

We analysed a total of 1670 bladder cancer cases, about one-quarter (28%) being in women; details are shown in table 1. There were some differences in age distribution; mean age was 71 years for women and 69 years for men. We observed distinct differences in staging distribution: the proportion of early Stage I was 49% for women and 61% for men, the proportion of Stages III and IV was 15% for women and 11% for men, and the proportion of cases whose stage was unknown was 19% for women and 13% for men. RER for women was 1.57 in univariate analysis and 1.13 (95% CI 0.88–1.46) in multivariate analysis; see table 2.

We analysed a total of 1264 kidney cancer cases, 42% of which were in women; see table 1. There were differences in age distribution: mean age was 68 years for women and 63 years for men. We observed no differences in staging distribution. RER for women was 1.18 in univariate analysis and 1.19 (95% CI 0.93–1.53) in multivariate analysis; see table 2.

We analysed a total of 1607 melanomas, with a slight predominance in women (54%); details are shown in table 1. There were no differences in age distribution and no differences in staging distribution: ~90% of cases were Stages I and II, 5–8% Stages III and IV and 4% had no staging information. RER for women was 0.92 in univariate analysis and 0.85 (95% CI 0.55–1.31) in multivariate analysis, see table 2.

We analysed a total of 752 thyroid cancer cases, three-quarters of which were in women; see table 1.

There were some differences in age structure: mean age was 53 years for women and 55 years for men. We observed differences in staging distribution: the proportion of Stages I and II was 69% in women and 59% in men and of Stages III and IV 22% in women and 28% in men. RER for women was 0.70 in univariate analysis and 0.74 (95% CI 0.42–1.30) in multivariate analysis; see table 2.

For all solid cancer sites combined, we analysed 21 102 cases, 42% of which were in women; details are shown in table 1. Women were slightly older, mean age being 67 years for women and 65 years for men. There were only slight differences in staging distribution for all sites combined. RER for women was 0.91 in the univariate analysis, 0.88 (95% CI 0.84-0.91) in the multivariate analysis without adjusting for case mix and 0.95 (95% CI 0.91-0.99) after adjusting for case mix. If the analysis was broken down by age group, univariate RER for women was 0.67, 0.81, 0.93 and 1.07, multivariate RER for women without adjusting for case mix was 0.81 (95% CI 0.76-0.89), 0.78 (95% CI 0.71-0.84), 0.91 (95% CI 0.84-0.99) and 1.07 (95% CI 0.96-1.19) and multivariate RER for women after adjusting for case mix 0.95 (95% CI 0.87-1.04), 0.84 (95% CI 0.77-0.92), 0.95 (95% CI 0.88-1.03) and 1.10 (95% CI 0.98-1.22) for age groups 20-59, 60-69, 70-79 and \geq 80, respectively. For details see table 3 and figure 1.

Table 1 Patient characteristics

Site	N		Percent	Mean age
Head and neck (without larynx) [C00–C14, C30–C31]	941	Women	26	60
		Men	74	59
Stomach [C16]	2418	Women	47	72
		Men	53	69
Colorectum [C18–C21]	4519	Women	51	71
		Men	49	67
Pancreas [C25]	951	Women	54	73
		Men	46	67
Lung [C33–C34]	3742	Women	26	66
		Men	74	65
Bladder [C67]	1670	Women	28	71
		Men	72	69
Kidney [C64–C66, C68]	1264	Women	42	68
		Men	58	63
Melanoma [C43]	1607	Women	54	54
		Men	46	54
Thyroid [C73]	752	Women	74	53
		Men	26	55
All solid sites (except NMSC) [C00–C80, except C44 and C50–C63]	21 102	Women	42	67
		Men	58	65

Discussion

Our main objective was to investigate whether survival differs between women and men in the population of Tyrol. Using the incidence data set for Tyrol, we applied a RER model that adjusts for main factors. The analysis was conducted for the main solid cancer sites and for the combination of all solid cancer sites in total and was split according to age class.

Most of our site-specific results are in line with published results, see for example.^{2–5,20–22} We observed poorer survival for women only for colorectal cancer, bladder cancer and kidney cancer, with none of these results being statistically significant. For colorectal cancer, our finding of a non-significant RER (1.06) does not stand in contradiction to published studies, see for example.^{4,5} A publication bias might have prevented non-significant results from being published, whereas we analysed and report results on all solid cancer sites with an adequate number of cases.

For bladder cancer, we obtained a multivariate RER of 1.13 for women without statistical significance. Recent analysis by Mungan *et al.*²⁰ showed poorer survival for women in the SEER (significant) and also in the Netherlands (non-significant) data set. However, the results differed for Stage I (better survival for women) and Stages II and IV (poorer survival for women). Micheli *et al.*^{2,3} also found poorer survival for women. Therefore, for bladder cancer, there seems to be some tendency towards poorer survival in women. However, there are well-recognized differences in the classification and registration of tumours that are recorded as malignant by some cancer registries and as non-malignant (benign) by others.²³

For aggregation of all cancer sites combined, we separately fitted a model instead of aggregating site-specific results. This procedure thus indirectly adjusted for differences in site mix between women and men.

For all cancer sites combined, the lesser excess mortality for females is 0.88 without adjusting for case mix and 0.95 after adjusting for case mix. This result is identical to a recent analysis by Micheli et al.3 on the large EUROCARE-4 data set. However, Micheli et al. observed a rather homogeneous gradient from a larger difference for younger age groups to a minor difference for older age groups. Our results do not show this homogeneous gradient and we observe a tendency towards worse survival for females in age group ≥ 80 years. Part of these differences can be explained by the distribution of sites by age groups in our data: whereas for younger females, the proportion of sites with better survival for females compared to males is larger then 55%, this proportion reduces to about one-third for women aged ≥ 80 years and in contrast the proportion of, for example, colorectal cancer with an insignificant RER of 1.06 increases from 17% to 32% (data not shown).

Table 2 RER estimators for solid cancer sites (univariate and multivariate relative risk and information on model fit)

Site	Univariate RER ^a with 95% CI	Multivariate RER ^a with 95% CI	Variables in model ^b	Model FIT ^c
Head and neck	0.57 (0.44–0.75)	0.72 (0.56–0.93)	Age, Stage	1.04/0.98
Stomach	0.97 (0.87-1.08)	0.86 (0.75-0.97)	Age, Stage, Period, Fup*Stage	1.12/1.01
Colorectum	1.16 (1.04–1.29)	1.06 (0.95–1.18)	Age, Stage, Period, HV, Fup*Stage	1.12/1.00
Pancreas	1.06 (0.92–1.23)	0.96 (0.78–1.19)	Age, Stage, Period, HV, Fup*Stage	1.23/1.15
Lung	0.87 (0.80-0.95)	0.82 (0.75-0.90)	Age, Stage, Period, HV, Fup*Stage	1.21/1.07
Bladder	1.57 (1.23–2.00)	1.13 (0.88–1.46)	Age, Stage, Period, HV, Fup*Stage	0.98/1.00
Kidney	1.18 (0.92–1.50)	1.19 (0.93–1.53)	Age, Stage. HV, Fup*Stage	1.20/1.13
Melanoma	0.92 (0.57-1.49)	0.85 (0.55–1.31)	Age, Stage, Fup*Stage	0.99/0.90
Thyroid	0.70 (0.40–1.23)	0.74 (0.42–1.30)	Age	1.06/0.81

a: For women compared to men

b: Year of follow-up (Fup) is always in model, also in univariate model; Fup*Stage, interaction term for follow-up and stage; HV, histological verification; Period, year of diagnosis

c: Deviance divided by degrees of freedom and Pearson divided by degrees of freedom

Factors that explain the observed differences between women and men under discussion are differences in tumour cell biology, which could be influenced by reproductive hormones,²¹ differences in anatomical situation, for example, for bladder cancer²⁰ and melanoma,²² and possibly most importantly differences in risk factors, especially comorbidity combined with smoking-related cancers. It is well known that about 3 of 10 cancer cases can be attributed to smoking and that smoking increases general mortality.²⁴ Relative survival adjusts for differences in background mortality, however, does not adjust for differences in mortality between smokers and non-smokers. Cancer registries usually contain no information on smoking habits of patients. Therefore, it would be very interesting to estimate the effect that differences in smoking prevalence between women and men have on the survival difference we observed.

Strengths and limitations

The following paragraphs will deal with the strengths and limitations of our study. One of the strengths of our study is that it employs a population-based data set, thus analysing all cancer

 Table 3 RER estimators for all solid cancer sites except NMSC combined (univariate and multivariate relative risk)

Age (years)	<i>N</i> (% female)	Univariate RER ^a	Multivariate ^b RER ^a without adjusting for case mix with 95% Cl	Multivariate ^b RER ^a after adjusting for case mix with 95% Cl
All	21 102 (42%)	0.91	0.88 (0.84–0.91)	0.95 (0.91–0.99)
20–59	6547 (39%)	0.67	0.81 (0.76–0.89)	0.95 (0.87–1.04)
60–69	5395 (34%)	0.81	0.78 (0.71–0.84)	0.84 (0.77–0.92)
70–79	5680 (45%)	0.93	0.91 (0.84–0.99)	0.95 (0.88–1.03)
\geq 80	3480 (57%)	1.07	1.07 (0.96–1.19)	1.10 (0.98–1.22)

a: For women compared to men

b: Adjusted for age, stage, year of diagnosis, histological verification and interaction terms for follow-up and stage and follow-up and age; year of follow-up (Fup) is always in model, also in univariate model

patients in the whole population, because we know that trial patients are often a prognostically favourable subset of all patients.²⁵ The incidence data set for Tyrol has been published in Cancer Incidence in Five Continents since 1988, which is an indirect measure of good completeness.

The next question is whether the model we applied is well-suited to answer our question. First, in survival analysis for oncological patients, it is state of the art to adjust for baseline mortality by applying relative survival. Therefore, relative survival is applied, for example, to compare survival figures in various countries.^{8,9} It is also well known that women and men have distinct life expectancy and mortality rates. Therefore, in order for a comparison of survival between women and men to be valid it is essential to adjust for differences in life expectancy. The model we applied is based on relative survival rates and, as such, should adequately adjust for that difference.

Summarizing, the main strengths are that the model we applied seems to be appropriate and the survival data are valid.

We are, of course, faced with some limitations. Cancer registries usually have only limited data for controlling information biases. If we compare survival for women and men, it is necessary to adjust for factors influencing survival. We noted in the 'Results section' that for some cancer sites like stomach, colorectum, pancreas and kidney, women are older than men at time of diagnosis. Also, some sites show clear differences in staging distribution. Our model adjusts for these few factors. However, residual confounding could be a limiting factor. Whether gender has a direct effect on survival, whether the effect is confounded in a classical way by, for example, tumour stage or whether the effect is influenced by some unknown factor interacting with tumour stage and with survival needs to be discussed; see for example Cole and Hernan.²⁶

It is also possible that a change occurred over time in factors influencing survival differences. Access to the medical system in Austria was already free of charge in the 1980s and 1990s. However, the social situation of women has changed greatly in the last three decades with a transition occurring from a very traditional female role to women holding a position in modern society. We cannot rule out the possibility that these changes influenced survival figures.



Multivariate without case mix adjustment

Multivariate with case mix adjustment

Figure 1 Gender effects for solid cancer sites and all sites combined by age class: gender RER with 95% Cls. RER estimates are shown for women compared to men as the reference group

Funding

This work was supported by the ONCOTYROL Center for Personalized Cancer Medicine. ONCOTYROL is a K1-COMET Center and funded by the Federal Ministry for Transport Innovation and Technology (BMVIT) and the Federal Ministry of Economics and Labour/the Federal Ministry of Economy, Family and Youth (BMWA/BMWFJ), the Tyrolean Future Foundation (TZS) and the State of Styria represented by the Styrian Business Promotion Agency (SFG) and supported by UMIT, University for Health Sciences, Medical Informatics and Technology and TILAK—Hospital Holding Company.

Conflicts of interest: None declared.

Key points

- In a healthcare system with free access to diagnostics and therapy, female cancer patients have a lesser excess mortality risk than do men for lung, stomach and head and neck cancer sites after adjusting for staging distribution and for all sites combined after also adjusting for case mix.
- Every cancer registry's report should routinely break down all results for gender.
- When analysing gender differences in survival, differences in life expectancy must be considered.

References

- Dick P, Sherif C, Sabeti S, et al. Gender differences in outcome of conservatively treated patients with asymptomatic high grade carotid stenosis. *Stroke* 2005;36:1178–83.
- 2 Micheli A, Mariotto A, Giorgi Rossi A, et al. The prognostic role of gender in survival of adult cancer patients. EUROCARE Working Group. *Eur J Cancer* 1998;34(14 Spec No): 2271–8.
- 3 Micheli A, Ciampichini R, Oberaigner W, et al. The advance of womenin cancer survival: An analysis of EUROCARE-4 data. Eur J Cancer 2009;45:1017–27.
- 4 Wichmann MW, Müller C, Hornung HM, et al. Gender differences in longterm survival of patients with colorectal cancer. Br J Surg 2001;88:1092–8.
- 5 Karner-Hanusch J, Marian B. Aspects of gender in colorectal tumors. Wien Med Wochenschr 2006;156:541-4.
- 6 Barford A, Dorling D, Smith GD, Shaw M. Life expectancy: women now on top everywhere. *Br Med J* 2006;332:808.

- 7 World Health Organisation. World Health Report 2006: Working Together for Health. Available at: http://www.who.int/whr/2006/en/. (6 May 2009, date last accessed).
- 8 Brenner H, Hakulinen T. On crude and age-adjusted relative survival rates. J Clin Epidemiol 2003;56:1185–91.
- 9 Brenner H, Hakulinen T. Age adjustment of cancer survival rates: methods, point estimates and standard errors. Br J Cancer 2005;93:372–5.
- 10 Brenner H, Gefeller O. Deriving more up-to-date estimates of long-term patient survival. J Clin Epidemiol 1997;50:211–6.
- Parkin DM, Whelan SL, Ferlay J, et al. *Cancer Incidence in Five Continents*, Vol. VII (IARC Scientific Publications No. 143). Lyon: IARC, 1997.
- 12 Parkin DM, Whelan S, Ferlay J, et al. *Cancer Incidence in Five Continents*, Vol. VIII (IARC Scientific Publications No. 155). Lyon: IARC, 2002.
- 13 Curado MP, Edwards B, Shin HR, et al. *Cancer Incidence in Five Continents*, Vol. IX (IARC Scientific Publications No. 160). Lyon: IARC, 2007.
- 14 Oberaigner W, Stuehlinger W. Record linkage in the Cancer Registry of Tyrol, Austria. *Methods Inf Med* 2005;44:626–30.
- 15 Oberaigner W. Errors in survival rates caused by routinely used deterministic record linkage methods. *Methods Inf Med* 2007;46:420–4.
- 16 Harrell FE, Lee KL, Califf RM, et al. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;3:143–52.
- 17 Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;23:51–64.
- 18 Dickman PW. strs Version 1.2.5. Available at: http://www.pauldickman .com/survival/strs. (2 July 2007, date last accessed).
- 19 Stata Statistical Software: Release 9.0. College Station, TX: StataCorp LP, 2005.
- 20 Mungan NA, Aben KK, Schoenberg MP, et al. Gender differences in stage-adjusted bladder cancer survival. *Urology* 2000;55:876–80.
- 21 Chatkin JM, Abreu CM, Fritscher CC, et al. Is there a gender difference in non-small cell lung cancer survival? *Gend Med* 2004;1:41–7.
- 22 Scoggins CR, Ross MI, Reintgen DS, et al. Gender-related differences in outcome for melanoma patients. *Ann Surg* 2006;243:693–8; discussion 698–700.
- 23 Quinn MJ, Wood H, Rowan S, Cooper N. Cancer Atlas of the United Kingdom and Ireland 1991–2000 (Studies on Medical and Population Subjects No. 68). Houndmills: Office for National Statistics (ONS), 2005.
- 24 Danaei G, Vander Hoorn S, Lopez AD, et al. Comparative Risk Assessment collaborating group (Cancers). Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet* 2005;366:1784–93.
- 25 Antmann K, Amato D, Wood W, et al. Selection bias in clinical trials. J Clin Oncol 1985;3:1142–7.
- 26 Cole SR, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;31:163–5.

Appendix 48

pendix 48

B1-7 Appendix Full Papers

Oberaigner et al. BMC Public Health 2011, **11**:91 http://www.biomedcentral.com/1471-2458/11/91

RESEARCH ARTICLE



Open Access

Introduction of organised mammography screening in tyrol: results of a one-year pilot phase

Willi Oberaigner^{1,2,3*}, Wolfgang Buchberger^{2,4†}, Thomas Frede^{5†}, Martin Daniaux^{5†}, Rudolf Knapp^{6†}, Christian Marth^{7†}, Uwe Siebert^{2,3,8,9†}

Abstract

Background: Efficiency and efficacy of organised mammography screening programs have been proven in large randomised trials. But every local implementation of mammography screening has to check whether the well established quality standards are met. Therefore it was the aim of this study to analyse the most common quality indices after introducing organised mammography screening in Tyrol, Austria, in a smooth transition from the existing system of opportunistic screening.

Methods: In June 2007, the system of opportunistic mammography screening in Tyrol was changed to an organised system by introducing a personal invitation system, a training program, a quality assurance program and by setting up a screening database. All procedures are noted in a written protocol. Most EU recommendations for organised mammography screening were followed, except double reading. All women living in Tyrol and covered by social insurance are now invited for a mammography, in age group 40-59 annually and in age group 60-69 biannually. Screening mammography is offered mainly by radiologists in private practice. We report on the results of the first year of piloting organised mammography screening in two counties in Tyrol.

Results: 56,432 women were invited. Estimated participation rate was 34.5% at one year of follow-up (and 55.5% at the second year of follow-up); 3.4% of screened women were recalled for further assessment or intermediate screening within six months. Per 1000 mammograms nine biopsies were performed and four breast cancer cases detected (N = 68). Of invasive breast cancer cases 34.4% were \leq 10 mm in size and 65.6% were node-negative. In total, six interval cancer cases were detected during one year of follow-up; this is 19% of the background incidence rate.

Conclusions: In the Tyrolean breast cancer screening program, a smooth transition from a spontaneous to an organised mammography screening system was achieved in a short time and with minimal additional resources. One year after introduction of the screening program, most of the quality indicators recommended by the European guidelines had been reached.

However, it will be necessary to introduce double reading, to change the rule for BI-RADS 3, and to concentrate on actions toward improving the participation rate.

Background

Breast cancer is the leading cause of female cancer death in all industrialised countries (and also worldwide), and the breast is also the leading incident cancer site for females [1]. Therefore, screening methods for breast cancer are of greatest public health importance.

* Correspondence: willi.oberaigner@tilak.at

Efficiency and efficacy of organised mammography screening programs have been proven in large randomised trials conducted in Europe and North America [1-9]. For several years already, organised mammography screening programs have been recommended in the EU [2]. Austria is one of the European countries where up to 2006 no organised programs were implemented, but where coverage by spontaneous mammography screening has been reported to be rather high: in a health survey conducted in Austria in 2006-2007 more than 80% of women aged 40+ answered that they had



© 2011 Oberaigner et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

⁺ Contributed equally

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Cancer Registry of Tyrol, Innsbruck, Austria

Full list of author information is available at the end of the article

had at least one mammography (ever) and more than 40% had had one in the past year [10]. However, it is known that self-reporting of screening usage overestimates true coverage [11], and our results are in line with this observation comparing the 40% coverage reported in the above-mentioned health survey to 34.5% in our study. In 2006, the Austrian health minister declared mammography to be one of the top health agendas, and in July 2006 a decision was made to implement organised mammography screening programs, namely in a first step in pilot regions, of which Tyrol is the largest.

Two questions now arise. The first question is whether it is really necessary to change the existing spontaneous mammography system, and severe doubts have been raised [12]. Up to now, our knowledge about the performance of the existing spontaneous program is minimal. We have only very limited information from routine reports from the Cancer Registry of Tyrol giving data on stage distribution of breast cancer cases on the population level and on the time trend of breast cancer incidence and mortality [13]. But it is well agreed that this information is by far not sufficient to assess the quality of a mammography program. To date most quality indices recommended by the EU guidelines cannot be calculated because we do not have the necessary information. It is our opinion that it is not justified to offer a mammography screening system to healthy women without having at least profound information on commonly agreed quality measures.

After that, the second question is how to change the existing screening system in the most efficient manner. It is well accepted that there is no uniform solution for implementing an organised screening system in a country, but, instead, when setting up the mammography system the health system conditions in the respective country must be given consideration. The outstanding challenge in introducing a mammography screening program in a country where a spontaneous screening system was conducted for some time is whether to make a smooth transition to an organised system or to completely redesign the existing screening system as was done, for example, in Germany [14]. In Tyrol, a clear decision was made to set up the new program while making best possible use of the existing experience and mammography screening network, which was established over the last fifteen years. Based on the latter strategy, it was possible to establish a country-wide mammography screening program in very short time. The price to be paid was the risk of potential quality problems, because some EU guidelines concerning the structure of the screening system were not fully adhered to.

To our knowledge, also in Europe a number of countries still have no organised mammography programs [15],

and therefore the experience in Tyrol can make an important contribution to the question how to switch a health system with spontaneous mammography screening to an organised system that meets well-accepted quality guidelines.

In June 2007, an organised mammography screening program was introduced in two central counties of Tyrol accounting for forty percent of that state's population. It was the aim of this study to analyse whether a smooth change from a spontaneous to an organised mammography system can meet the quality indicators recommended by the EU guidelines.

Methods

Study population, invitation

The target population includes all women aged 40 to 69 living in two counties of Tyrol (i.e., the capital of Innsbruck, and the surrounding area) and covered by compulsory social insurance. More than 97% of the population of Tyrol are covered by compulsory social insurance (personal communication). All women in the target population are sent a personal invitation letter: women aged 40-59 annually, and women aged 60-69 biannually. All women are invited regardless of their screening history and their individual cancer status. Invitations are addressed directly to the women; the invitation is to consult a screening unit. Women invited to screening receive detailed information on the screening programme and must sign an informed consent before screening. Mammography screening is offered by 12 screening units, nine of them run by radiologists in private practice and three by hospital outpatient departments at the two public hospitals in the study area. The mammogram is read by only one radiologist; ultrasound (US) is offered to women at the radiologist's decision. The mammography result is coded according to the BI-RADS [16] scheme, and the participating women are informed of the screening result immediately after the screening test. Women with BI-RADS 1-2 go back to screening, women with BI-RADS 3 are invited for intermediate screening in six months and women with BI-RADS 4 or 5 are invited for further assessment. Assessment is offered by three hospital radiology units in the study area and includes clinical inspection, mammography, US, MRI and biopsy as needed. The one large assessment unit at Innsbruck University Hospital works closely with a breast cancer centre. As the system is open, women are free to contact the assessment unit irrespective of the mammogram result, and a number of the women go to assessment even if the mammography result is BI-RADS 1,2,3. All program activities were planned carefully and documented in a written protocol. The program is directed by a screening group that meets monthly. A subgroup of the project team is

responsible for quality assurance, which is based on quarterly analysis of screening data. According to social insurance regulations, women must first visit their general practitioner who refers them to a screening unit. All radiologists participating in the program had to undergo a training program and need ÖRG (Austrian Radiology Association) certification. In the median, private radiologists performed 2450 mammograms and the three hospital units 1134, 1379 and 4620 mammograms per year. There is no waiting time for mammography. Of the women who were referred to assessment, 64% waited less than five working days, 18% between six and ten working days and 19% more than ten working days.

Data collection

All mammography units register basic information in a database. It is noteworthy that all mammograms are registered, not only those for women belonging to the target population. Due to data privacy restrictions, women must sign a written consent to permit data transfer to the screening database; if a woman refuses consent, an empty dataset marked simply "data transfer declined" is sent to the mammography database. Screening information is transferred to the screening database after pseudonymising the woman's social insurance number. The pseudonymisation process permits linkage of data for a specific woman coming from different units and guarantees data confidentiality, because the pseudonymisation process can be reversed only within the screening unit. An analogous procedure was established for assessment units. Finally, data on tumour characteristics are collected by the Cancer Registry of Tyrol. The Cancer Registry has developed a network of various data sources that guarantees a high degree of completeness and validity of cancer data in the population of Tyrol. Details on registry procedures and figures on completeness have been reported elsewhere [17]. The Cancer Registry also collects each patient's social insurance number and, consequently, Cancer Registry data can be linked to the screening database by applying the same pseudonymisation process. Cancer Registry data enable us to analyse data on tumour characteristics (e.g., tumour size, lymph node status, information on surgery) and to assess interval cancer cases.

Participation rate should reflect all women undergoing a screening mammography. Part of them refuse consent to permit data transfer. In order to account for these women, we estimated the proportion of all women denying consent and belonging to age group 40-69 as being 56% of 5.8%, namely 3.2%. Thus, the estimated participation rate is equal to the observed participation rate plus 3.2%. Finally, invitation data are provided by the invitation system run by the social insurance carrier. Invitation data are transferred to the screening database as aggregated numbers of invited women per month of invitation, age group (five-year age classes) and county. Recall is defined as call for further assessment or invitation to intermediate screening within six months.

Statistical analysis

Plausibility checks are implemented both at the mammography and the assessment unit level and at the central screening database level.

The screening database is realised as STATA datasets. Linkage between screening data, assessment data and Cancer Registry data is based on the pseudonym number. We report numbers and proportions as defined in the EU guidelines. For some indices, population-based rates are computed using the official population data supplied by Statistics Austria. No statistical testing is applied. All reporting is done with STATA Version 9 [18].

Spontaneous mammography screening was introduced in Tyrol already in the early 1990 s. Thus, the underlying background incidence is defined by years 1988-1990, see Table 1.

This study was conducted in conformity to the Helsinki Declaration [19]. The project was approved by the Ethics Committee of Medical University Innsbruck.

Results

From June 2007 to May 2008, 56,432 women in the target population were invited, 40.3% in age group 40-49, 31.1% in age group 50-59 and 28.6% in age group 60-69, see Table 2. A total of 1,188 invitation letters were returned as undeliverable. Because of their small number, we did not take the returned invitations into consideration for analysis. After deleting N = 80 cases without a BI-RADS classification, we ended up with **17,645 screening cases** in the analysis dataset. Of the screening cases 40.4% were in age group 40-49, 32.2% in age group 50-59 and 27.4% in age group 60-69.

Table 1 Breast cancer incidence rate in Tyrol beforespontaneous screening program 1988-1990 (backgroundincidence) and before changing the screening system2005-2007

	40-49 ¹⁾	50-59 ¹⁾	60-69 ¹⁾	Total (40-69) ¹⁾
1988-1990				
Breast Cancer	52 (128.8)	54 (176.0)	74 (244.0)	180 (177.5)
Invasive	50 (123.8)	53 (172.7)	72 (238.5)	175 (172.9)
In situ	2 (5.0)	1 (3.3)	2 (5.5)	5 (4.6)
2005-2007				
Breast Cancer	82 (146.2)	87 (206.3)	96 (263.0)	265 (196.5)
Invasive	74 (132.0)	79 (187.3)	89 (245.6)	242 (179.9)
In situ	8 (14.2)	8 (19.0)	7 (17.4)	23 (16.6)

¹⁾Average number of incident cases per year and age-specific rate per 100,000.

40-49	50-59	60-69	Total (40-69)	
22739	17531	16162	56432	
7124	5689	4832	17645	
31.3%	32.5%			
34.5%	35.7%			
54.3%	53.5%	48.2%	52.3%	
57.5%	56.7%	51.4%	55.5%	
	40-49 22739 7124 31.3% 34.5% 54.3% 57.5%	40-49 50-59 22739 17531 7124 5689 31.3% 32.5% 34.5% 35.7% 54.3% 53.5% 57.5% 56.7%	40-4950-5960-6922739175311616271245689483231.3%32.5%34.5%34.5%35.7%54.3%54.3%53.5%48.2%57.5%56.7%51.4%	40-4950-5960-69Total (40-69)227391753116162564327124568948321764531.3%32.5%

Table 2 Invitation system: Number of women invited, number of mammograms performed (numbers and age percentages)

¹⁾Also a part of women age 40-59 go to mammography screening at a two year interval.

The observed participation rate in the first year of follow-up was 31.3%; after correcting for women declining data transfer to the screening database, the overall participation rate was 34.5% (34.5%, 35.7% and 33.1% in age groups 40-49, 50-59 and 60-69, respectively). After completing a second year of follow-up (end of observation was 31 May 2009), 55.5% of the invited women had undergone at least one screening examination, with a higher participation rate in younger age classes (57.5%, 56.7% and 51.4% for age groups 40-49, 50-59 and 60-69, respectively), see Table 2.

Screening outcome was negative for 96.6%; 1.6% of cases were recommended for intermediate mammography (six months) after screening (1.2% in age group 60-69), 1.8% were referred for further assessment (1.4% in age group 60-69) and in 10 cases the screening outcome was unknown (Table 3).

According to screening policy, the screening radiologists were free to perform additional US. The reason for additional US was recorded as breast density grades according to the American College of Radiology (ACR) 3 or 4 in 37.6%, equivocal findings on mammography in 16.9%, and other nonspecified reasons in 45.5% of cases. Overall, 83.3% of women had an additional US examination, with clear differences between the age groups: the proportions were 89.2%, 81.3% and 77.1% in age groups 40-49, 50-59 and 60-69, respectively (Table 4).

Of 315 assessments performed, 38.4% had a core biopsy (45.5% in age group 60-69), 7.6% a fine needle biopsy and 1.9% (six cases) an open biopsy. This means that of 1000 women screened, nine had a biopsy (8 in

age group 60-69) and 0.3 underwent an open biopsy (Table 5). Of all assessments 75.6% were negative, 2.9% were recommended for intermediate screening and in 68 (21.6%) screening cases breast cancer was diagnosed (12.9%, 20.2% and 42.4% in age groups 40-49, 50-59 and 60-69 respectively), see Table 6.

For all fine needle biopsies, the diagnosis was benign. The positive predictive value of core biopsy was 50.4% in total and 35.3%, 47.5% and 80.0% in age groups 40-49, 50-59 and 60-69, respectively. The six open biopsies showed two benign and four malignant cancer cases.

Of the 68 breast cancer cases, 61 were invasive, and seven were ductal carcinoma in situ. The cancer detection rate was 3.9 per 1000 mammograms in total and 2.5, 3.9 and 5.8 in age groups 40-49, 50-59 and 60-69 respectively, see Table 6.

Of all invasive cancers detected, 34.4% were less then 10 mm in size and 65.6% were node-negative, without differences for age group, see Table 7.

After linking screening data and Cancer Registry data (only depseudonymised data), we observed a total of six interval cancer cases within one year after screening, four in age group 40-49, one in age group 50-59 and one in age group 60-69. The interval cancer rate was 34 per 100,000 mammograms; this is 19% of the underlying background incidence rate (defined by years of diagnosis 1988-1990). Details are shown in Table 8, which shows the most important quality indicators defined in the EU guidelines restricted to age group 50-69, because EU recommendations are given for that age group.

Table 3 Screening outcome

	40-49	50-59	60-69	Total
Negative	6853 (96.2%)	5482 (96.4%)	4704 (97.4%)	17039 (96.6%)
Intermediate screening test following screening	125 (1.8%)	95 (1.7%)	57 (1.2%)	277 (1.6%)
Assessment				
Recommended	142 (2.0%)	109 (1.9%)	68 (1.4%)	319 (1.8%)
Performed	140 (2.0%)	109 (1.9%)	66 (1.4%)	315 (1.8%)
Unknown ¹⁾	4	3	3	10
Total	7124	5689	4832	17645

¹⁾BI-RADS 0 without assessment was treated as unknown.

Table 4 Addition	onal ultrasour	d imaging	at	screening
------------------	----------------	-----------	----	-----------

55				
	40-49	50-59	60-69	Total
Ultrasound added to mammography screening	6358 (89.2%)	4624 (81.3%)	3723 (77.1%)	14705 (83.3%)
Reason for ultrasound:				
Breast density (ACR 3/4)	2855 (44.9%)	1638 (35.4%)	1041 (28.0%)	5534 (37.6%)
Equivocal finding	1037 (16.3%)	810 (17.5%)	635 (17.1%)	2482 (16.9%)
Other	2466 (38.8%)	2176 (47.1%)	2047 (55.0%)	6689 (45.5%)

Discussion

We analysed the situation after a first year of introducing organised mammography screening in two counties in Tyrol accounting for 40% of that state's population. The organised program was established in a smooth transition from the existing spontaneous mammography screening system, namely by introducing a written protocol, a personal invitation system, a training program, and by setting up a screening database allowing us to investigate performance and outcome parameters in detail. Although not all EU recommendations have been followed to date, most quality indicators are in the range of accepted and/or desired levels given by the EU guidelines: the proportion of cases called for further assessment (20 per 1000 mammograms), the biopsy rate (9 per 1000 mammograms), the proportion of invasive screening-detected cancer (89.7%), the proportion of invasive cancer ≤ 10 mm in size (34.4%) and the interval cancer rate expressed as a multiple of the background incidence rate (19%). The average number of screens read by a radiologist (about 2400) does not meet the EU recommendation of 5000. However, in about four of five women an additional US is done, which is known to improve the sensitivity of the screening test, see for example [20,21].

Some of the parameters are near the value expected in subsequent screening rounds and not in a first round (for example, recall rate and breast cancer detection rate). However, we must remember that the organised program was introduced after fifteen years of the spontaneous screening program. Some indicators like proportion of stage II+ cancers and node-negative cancer are slightly outside the EU-accepted levels. Some of the observations could be due to small numbers (we observed a total of 68 cancer cases). So it is too early to come to final conclusions on the mammography screening model we describe here.

Only one indicator clearly misses the EU recommendations, namely the participation rate. However, participation rate depends not only on program organisation, but also on cultural background. A look at neighbouring German-speaking countries, which should have a similar culture, shows participation rates of 54% in Germany [14] and 25.9% to 65.9% in five cantons of Switzerland [22]. Thus, a cumulative participation rate of 55% in the first two years would seem to be rather successful by comparison to that of countries with a similar cultural background, albeit not the goal we aimed for.

Among the strengths of the Tyrolean breast cancer screening program is its implementation: within a short time and with minimal additional resources it was possible to set up an organised population-based screening program that - at least at evaluation after one year - met all of the EU quality indicators except participation rate. It was not necessary to set up extra screening units; instead, the program used the existing network of screening and assessment units. In terms of epidemiology, another of the program's strengths is a complete pseudonymised database

	40-49	50-59	60-69	Total
Additional imaging				
Ultrasound	140 (100%)	106 (97.2%)	61 (92.4%)	307 (97.5%)
MRI	24 (17.1%)	21 (19.3%)	21 (31.8%)	66 (21.0%)
Biopsy	63 (45.0%)	51 (46.8%)	37 (56.1%)	151 (47.9%)
Core biopsy	51 (36.4%)	40 (36.7%)	30 (45.5%)	121 (38.4%)
Fine needle biopsy	10 (7.1%)	10 (9.2%)	4 (6.1%)	24 (7.6%)
Open biopsy	2 (1.4%)	1 (0.9%)	3 (4.5%)	6 (1.9%)
Biopsy rate per 1000 mammograms	8.8	9.0	7.7	8.6
PPV for core biopsy ¹⁾	35.3.4%	47.5%	80.0%	50.4%
PPV for open biopsy ²⁾				66.7%
Total	140	109	66	315

Table 5 Assessment procedure

¹⁾For all fine needle biopsies the result was benign.

²⁾For six open biopsies two were benign and four malignant. Because of small numbers, PPV is shown only for the total.

Table 6 Assessment outcome						
	40-49	50-59	60-69	Total		
Negative	117 (83.6%)	85 (78.0%)	36 (54.5%)	238 (75.6%)		
Recommendation for intermediate screening after assessment	5 (3.6%)	2 (1.8%)	2 (3.0%)	9 (2.9%)		
Breast cancer	18 (12.9%)	22 (20.2%)	28 (42.4%)	68 (21.6%)		

Total	140	109
Ratio screening breast cancer detection rate vs. background incidence rate ¹⁾	1.9	2.2
Breast cancer detection rate per 1000 mammograms	2.5	3.9
Invasive	14 (10.0%)	20 (18.3%)
In situ	4 (2.9%)	2 (1.8%)
Breast cancer	18 (12.9%)	22 (20.2%)

¹⁾Background incidence rate: see Table 1.

that covers all mammography exams and can be linked to the Cancer Registry data. The database was set up in a very short time, it is up to date within one month and serves as a very important tool for monitoring the program quality indicators or, more generally speaking, for all kinds of quality assurance tasks.

Nevertheless, the program differs from many organised programs in the EU in three aspects. Firstly, we also included women aged 40-49. During previous spontaneous screening campaigns women aged 40-49 were officially invited to participate in mammography screening. The discussion in the USA after publishing the revised recommendation by USPSTF [23,24] shows that it is very challenging and difficult for women to understand why a service is cancelled that was recommended for several years. Without well founded data, we feel it is not justified to discontinue screening in age class

Table 7 Chave stavistics of investive several

	40-49	50-59	60-69	Total		
Tumour size (mm):	13.5; 3-50	14; 2-60	12; 1-49	13; 1-60		
Median; range						
Tumour size (mm):						
<= 10 mm	6 (42.9%)	6 (30.0%)	10 (37.0%)	21 (34.4%)		
11-20 mm ¹⁾	6 (42.9%)	8 (40.0%)	12 (44.4%)	26 (42.6%)		
>20 mm	3 (21.4%)	6 (30.0%)	5 (18.5%)	14 (23.0%)		
Lymph node involvement	5 (35.7%)	7 (35.0%)	9 (33.3%)	21 (34.4%)		
Total	14	20	27	61		
Tumor stage according to TNM						
pT1	12 (85.7%)	14 (70.0%)	22 (81.5%)	48 (78.7%)		
pT2	2 (14.3%)	5 (25.0%)	5 (18.5%)	12 (19.7%)		
pT3		1 (5.0%)		1 (1.6%)		
pN0	9 (64.3%)	13 (65.0%)	18 (66.7%)	40 (65.6%)		
pN1	4 (28.6%)	4 (20.0%)	6 (22.2%)	14 (23.0%)		
pN2	1 (7.1%)	2 (10.0%)	2 (7.4%)	5 (8.2%)		
pN3		1 (5.0%)	1 (3.7%)	2 (3.3%)		
MO	13 (92.9%)	19 (95.0%)	26 (96.3%)	58 (95.1%)		
M1	1 (7.1%)	1 (5.0%)	1 (3.7%)	3 (4.9%)		

7 (2.2%)

3.9

22

315

61 (19.4%)

40-49. We are collecting detailed data and will evaluate the goods and harms [23,25] in coming years. In addition, the analysis of breast cancer mortality in Tyrol in the past decade shows that mortality decrease was greatest in women aged 40-49, an effect that can at least partly be attributed to spontaneous screening [13].

1 (1.5%)

5.8

24

66

27 (40.9%)

Secondly, we offer breast US as an additional diagnostic tool in screening. In opportunistic screening in Tyrol, US was offered to women with dense breasts (ACR density grades 3 and 4) and with inconclusive findings on mammography [26,27]. The line of reasoning concerning screening in age group 40-49 also holds for adjunct US, bearing in mind that adjunct US was offered during the last decade. It has been shown by various authors that the additional use of US can improve cancer detection rates, especially in younger women and women with dense breasts [20,21,26]. The relative percentage of carcinomas found in supplemental breast US examinations as a fraction of the total number of detected cancers was reported by four studies with a mean percentage of 22.5% (15%-34%) [19]. In the second year of the organised program we collected detailed data allowing us to analyse the contribution of US to sensitivity and specificity outside the framework of studies, namely in a population-based setting.

And thirdly, we were not able to implement double reading during the piloting phase. Interestingly, performance parameters, especially interval cancer rate, showed that also without double reading an acceptable quality level was achieved. The question of the completeness of interval cancer rate depends on the completeness of the Cancer Registry of Tyrol, which covers the target population. We analysed quality measures for the Cancer Registry in detail, also completeness, and found a high degree of completeness, both for all cancer sites and for breast cancer [17]. Therefore, we think it is unlikely that we missed interval cancers. However, the numbers are small and we cannot exclude the possibility that we missed one or two cases, which would mean a 33% increase in the interval cancer rate (from 6 to 8 cases). Clearly, a longer observation period is necessary before

	Tyrol 50-69	EU-accepted	EU-desired
Attendance proportion	54.2	>70	>75
Recall rate ¹⁾	3.1 (327/10521)	<7%/ <5%	<5%/ <3%
Interval cancer rate/Background incidence rate (BIR) 0-11 months	9% (two interval cancers in age 50-69)	30%	<30%
Breast cancer detection rate	2.3 * BIR	3*BIR	>3*BIR
Stage II+/Total cancers screen-detected	38.0% (19/50)	NA	<30%
Invasive cancers \leq 10 mm/Total invasive cancers	34.0% (16/47)	NA	>= 25%
Proportion of invasive cancers that are < 15 mm in size	51.1% (24/47)	50%	>50%
Invasive cancer/Total cancers screen-detected	94% (47/50)	90%	80%-90%
Node-negative cancer/Total invasive cancers screen-detected	66.0% (31/47)	NA	>70%
Benign to malignant open surgical biopsy ratio	0:1 (0:4)	<=1:2	<=1:4
Benign to malignant core biopsy ratio	1:0.9 (60:64)		

¹⁾Recall is defined as call for further assessment or invitation to intermediate screening within six months.

coming to a final conclusion on this parameter. This observation could in part be explained by short screening intervals in the age group 40-59 and by the use of additional US. Nevertheless, double reading will be introduced in the regular screening program after concluding the piloting phase.

Also, as we do not have a scheduling system, women are invited to consult the screening unit at any time that is convenient to them. The time when the invitation is sent is independent of a woman's individual screening history. The latter point is corrected in the next invitation round, which required a change in the written consent for reasons of data privacy.

Many countries have had a mammography screening program running for decades or for a shorter time. On the other hand, six EU member states still have no organised nationwide breast cancer screening program, and in seven member states a nationwide rollout was in 2007 [15]. What can be learned from our experience by countries that are thinking of introducing or are already in the process of planning to introduce a mammography screening program? In our opinion, the greatest difference between our approach and many other approaches is the smooth transition made from the existing spontaneous program to organized population-based screening. We made use of the network of screening and assessment units that had already been set up during spontaneous screening. What we added was an invitation system covering the entire population of Tyrol, a screening database that allows quality indices to be monitored and a well-defined training program for both screening and assessment units. With this strategy we were able to meet most EU quality indices in a very short time.

Conclusions

A smooth transition from a spontaneous to an organised mammography screening system that uses the existing

screening units, introduces an invitation system and a quality assurance program (including a screening database) can meet the quality indicators recommended by European guidelines in a short time and with minimal additional resources. However, it will be necessary to introduce double reading, to change the rule for BI-RADS 3 and to concentrate on actions toward improving the participation rate.

List of Abbreviations

US: ultrasound; ÖRG: Austrian Radiology Association; ACR: American College of Radiology.

Acknowledgements

This work was supported by the ONCOTYROL Center for Personal Cancer Medicine. ONCOTYROL is a COMET Center funded by the Federal Ministry for Transport Innovation and Technology (BMVIT) and the Federal Ministry of Economics and Labour (BMWA), the Tyrolean Future Foundation (TZS) and the Land of Styria represented by the Styrian Business Promotion Agency (SFG) and supported by UMIT - University for Health Sciences, Medical Informatics and Technology, Innsbruck Medical University, Tyrolean Health Insurance Fund and Tyrolean Health Company.

Author details

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Cancer Registry of Tyrol, Innsbruck, Austria. ²Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Information Systems and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria. ³ONCOTYROL - Center for Personalized Cancer Medicine, Innsbruck, Austria. ⁴Medical Director, TILAK, Innsbruck, Austria. ⁶Innsbruck Medical University, Department of Radiology, Kufstein, Austria. ⁶Innsbruck Medical University, Department of Obstetrics and Gynecology, Innsbruck, Austria. ⁸Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA. ⁹Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

Authors' contributions

WO designed the study, performed the analysis and wrote the paper. WB contributed to study design and assisted in writing the paper, especially contributions on ultrasound and screening at age below 50. TF, MD und RK contributed to the discussion part, especially from the radiology point of view. CM contributed to the methods, report and discussion part, especially

from the gynaecological point of view. US contributed to writing of the paper. All Authors reviewed and agreed to the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 29 March 2010 Accepted: 9 February 2011 Published: 9 February 2011

References

- Anttila A, Koskela J, Hakama M: Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. J Med Screen 2002, 9(4):153-8.
- Hakama M, Pukkala E, Heikkila M, Kallio M: Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. BMJ 1997, 314(7084):864-7.
- Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE: Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002, 359(9310):909-19.
- Tabar L, Yen MF, Vitak B, Chen HH, Smith RA, Duffy SW: Mammography service screening and mortality in breast cancer patients: 20-year followup before and after introduction of screening. *Lancet* 2003, 361(9367):1405-10.
- Tornberg S, Kemetli L, Lynge E, et al: Breast cancer incidence and mortality in the Nordic capitals, 1970-1998. Trends related to mammography screening programmes. Acta Oncol 2006, 45(5):528-35.
- Gotzsche PC, Nielsen M: Screening for breast cancer with mammography. Cochrane Database Syst Rev 2006, , 4: CD001877.
- Gotzsche PC, Nielsen M: Screening for breast cancer with mammography. Cochrane Database Syst Rev 2009, , 4: CD001877.
- Reduction in breast cancer mortality from organized service screening with mammography: 1. Further confirmation with extended data. *Cancer Epidemiol Biomarkers Prev* 2006, 15(1):45-51.
- Reduction in breast cancer mortality from the organised service screening with mammography: 2. Validation with alternative analytic methods. *Cancer Epidemiol Biomarkers Prev* 2006, 15(1):52-6.
- Klimont J, Kytir J: Österreichische Gesundheitsbefragung 2006/2007. Hauptergebnisse und methodische Dokumentation. Wien: Statistik Austria; 2007.
- Cronin KA, Miglioretti DL, Krapcho M, et al: Bias associated with self-report of prior screening mammography. *Cancer Epidemiol Biomarkers Prev* 2009, 18(6):1699-705.
- 12. Vutuc C, Waldhoer T, Haidinger G: Breast cancer trends: opportunistic screening in Austria versus controlled screening in Finland and Sweden. *Eur J Cancer Prev* 2006, **15(4)**:343-6.
- Oberaigner W, Buchberger W, Frede T, Knapp R, Marth C, Siebert U: Breast cancer incidence and mortality in Tyrol/Austria after fifteen years of opportunistic mammography screening. *BMC Public Health* 2010, 10(1):86.
- 14. Evaluationsbericht 2005-2007. Ergebnisse des Mammographie-Screening-Programms in Deutschland. Köln: Kooperationsgemeinschaft Mammographie 2009.
- Karsa L, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M: Cancer Screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening. First Report. Brussels: European Commission; 2008.
- Breast Imaging Reporting and Data System Atlas (BI-RADS[®]Atlas). American College of Radiology (ACR); 2003.
- 17. Oberaigner W, Siebert U: Are survival rates for Tyrol published in the Eurocare studies biased? Acta Oncol 2009, 48(7):984-91.
- Stata Statistical Software: Release 9.0. College Station, Tx, StataCorp LP; 2005.
- World Medical Association: World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. [http://www.wma.net/en/30publications/10policies/b3/index.html], (accessed 7 November 2009).
- Nothacker M, Duda V, Hahn M, et al: Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. BMC Cancer 2009, 9:335.

- 21. Kolb TM, Lichy J, Newhouse JH: Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002, **225**(1):165-75.
- 22. Breast cancer screening by mammography: International evidence and the situation in Switzerland. Bern: Swiss Cancer League/Oncosuisse; 2007.
- Mandelblatt JS, Cronin KA, Bailey S, et al: Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009, 151(10):738-47.
- 24. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009, 151(10):716-26, W-236.
- 25. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L: Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009, **151(10)**:727-37, W237-42.
- Buchberger W, Niehoff A, Obrist P, DeKoekkoek-Doll P, Dunser M: Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. Semin Ultrasound CT MR 2000, 21(4):325-36.
- 27. Frede TE: Opportunistic breast cancer early detection in Tyrol, Austria 1996-2004. Is a mammography-screening program necessary? *Eur J Radiol* 2005, **55**(1):130-8.

Pre-publication history

The pre-publication history for this paper can be accessed here: http://www.biomedcentral.com/1471-2458/11/91/prepub

doi:10.1186/1471-2458-11-91

Cite this article as: Oberaigner *et al*: **Introduction of organised** mammography screening in tyrol: results of a one-year pilot phase. *BMC Public Health* 2011 **11**:91.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit

B1-8 Appendix Full Papers

DOI 10.1007/s00038-011-0266-4

ORIGINAL ARTICLE

Prostate-specific antigen testing in Tyrol, Austria: prostate cancer mortality reduction was supported by an update with mortality data up to 2008

Willi Oberaigner · Uwe Siebert · Wolfgang Horninger · Helmut Klocker · Jasmin Bektic · Georg Schäfer · Ferdinand Frauscher · Harald Schennach · Georg Bartsch

Received: 5 May 2010/Revised: 30 April 2011/Accepted: 30 May 2011 © The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Objectives The objective of this study was to update an in-depth analysis of the time trend for prostate cancer (PCA) mortality in the population of Tyrol by 5 years, namely to 2008. In Tyrol, prostate-specific antigen (PSA) tests were introduced in 1988/89; more than three-quarters of all men in the age group 45–74 had at least one PSA test in the past decade.

Methods We applied the same model as in a previous publication, i.e., an age-period-cohort model using Poisson regression, to the mortality data covering more than three decades from 1970 to 2008.

Results For Tyrol from 2004 to 2008 in the age group 60+ period terms show a significant reduction in prostate

W. Oberaigner (⊠) Cancer Registry of Tyrol, TILAK GmbH, Anichstrasse 35, Innsbruck, Austria e-mail: willi.oberaigner@iet.at

W. Oberaigner · U. Siebert Department of Public Health and Health Technology Assessment, University for Health Sciences, Medical Informatics and Technology, Hall i.T, Austria

W. Oberaigner · U. Siebert

Oncotyrol Division of Public Health Decision Modelling, Health Technology Assessment and Health Economics, Oncotyrol Center for Personalized Cancer Medicine, Innsbruck, Austria

U. Siebert

Cardiovascular Research Program, Department of Radiology, Institute for Technology Assessment, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

U. Siebert

Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA

Published online: 17 June 2011

cancer mortality with a risk ratio of 0.70 (95% confidence interval 0.57, 0.87) for Tyrol, and for Austria excluding Tyrol a moderate reduction with a risk ratio of 0.92 (95% confidence interval 0.87, 0.97), each compared to the mortality rate in the period 1989–1993.

Conclusions This update strengthens our previously published results, namely that PSA testing offered to a population at no charge can reduce prostate cancer mortality. The extent of mortality reduction is in line with that reported in the other recent publications. However, our data do not permit us to fully assess the harms associated with PCA screening, and no recommendation for PSA screening can be made without a careful evaluation of overdiagnosis and overtreatment.

W. Horninger \cdot H. Klocker \cdot J. Bektic \cdot G. Bartsch Department of Urology, Innsbruck Medical University, Innsbruck, Austria

G. Schäfer Department of Pathology, Innsbruck Medical University, Innsbruck, Austria

F. Frauscher Department of Radiology, Innsbruck Medical University, Innsbruck, Austria

H. Schennach Department of Immunology and Transfusion Medicine, Innsbruck Medical University, Innsbruck, Austria **Keywords** Prostate-specific antigen · Prostate cancer · Mortality · Screening · Early detection

Introduction

Prostate cancer (PCA) is the second-leading cause of male cancer death in most industrialized countries (Boyle et al. 2008). One out of six men in the United States and in most industrialized countries will be affected by PCA during his lifetime (SEER 2011). These facts explain why there is an exceptional interest in scientifically proven evidence on whether organized prostate-specific antigen (PSA) screening reduces PCA mortality and what harms are associated with such screening. Until now, screening healthy men for PCA has shown to be feasible and acceptable in the large studies (De Koning et al. 2002). The results of two large randomized studies with more than 75,000 cases and controls in each study were recently published (Schröder et al. 2009; Andriole et al. 2009), one showing a 20% reduction in the PCA death rate with a p value of 0.04, and the other showing no reduction.

Tyrol is one of the few countries where PSA testing was introduced already as early as 1988/89, and since 1993 it has been offered to all men aged 45–74 (Bartsch et al. 2001). In 2006, we published an analysis of PCA mortality in Tyrol until the year of death 2003 and concluded that PSA testing offered free of charge to a population can reduce PCA mortality (Oberaigner et al. 2006). Critical questions were raised about the short observation period and too small numbers and therefore unstable results. Recently, the mortality data for Austria to year of death 2008 were released, which enables us to update our analysis by adding 5 years to the observation period, namely 2004–2008.

Therefore, in order to determine whether PSA testing might be able to reduce PCA mortality in Tyrol, our goal was to update our analysis of PCA mortality in Tyrol by a 5-year period, namely by years of death 2004–2008.

Methods

Mortality data were analyzed for Tyrol, and for Austria excluding Tyrol. In Austria, mortality data are collected by Statistics Austria (Hansluwka 1989); details are described elsewhere (Oberaigner et al. 2006). We analyzed all cases coded for cause of death PCA as described above.

Population data are also collected by Statistics Austria. Census data are available for the years 1971, 1981, 1991 and 2001; for intercensus, years population figures are extrapolated based on births, deaths and migration information.

🖉 Springer

The male population of Tyrol in the last year of observation, namely 2008, was 343,340, and of Austria excluding Tyrol 3,715,295.

Analysis of mortality time trends was based on ageperiod-cohort (APC) modeling (Clayton and Schifflers 1987a, b). APC models allow separate effects to be estimated for age (A), period or year of death (P) and cohort (C) by means of Poisson regression. The method is identical to that used in our previous publication and was described there (Oberaigner et al. 2006).

We fitted separate models for PCA mortality for Tyrol, and for Austria excluding Tyrol. For Tyrol, the APC model reaches 24 degrees of freedom (DF) and deviance 30.4, which seems to be reasonably good. For Austria excluding Tyrol, the APC model reaches 24 DF and deviance 103.0. For every step in model extension, the likelihood ratio test shows that the parameter effect is different from a zero effect. Thus, it is justified to add each parameter step by step.

For statistical analysis, the number of PCA deaths was aggregated in 5-year age groups, 5-year period groups and consequently 5-year cohort groups. There are a very few PCA deaths for age below 60 in Tyrol (3.3% of all PCA deaths). We thus decided to build the model for age groups beginning with age 60–64 and continue in 5-year age groups. Our hypothesis was that the mortality rate decreases following PSA testing, so the reference category for period was 1989–1993.

In order to rule out the possibility that the choice of the age of PCA deaths analyzed by us, namely 60+, had a predominant influence on the result, we performed a sensitivity analysis by applying the same model to age groups 50+, 50-74 and 60-74.

The analysis was performed with Stata Version 9, using poisson procedure for Poisson regression (Stata Statistical Software 2005).

Results

Basic numbers of PCA deaths and age-standardized rates are given in Table 1; Fig. 1. Effects of the APC model are described in Table 2; reference category for age is the age group 60–64, for the period group 1989–1993 and for cohort the cohort group 1882–1886. Figure 2 shows the period effects for Tyrol and for Austria excluding Tyrol.

Age effects are of similar size for Tyrol and for Austria excluding Tyrol: as compared to age group 60–64, the effects are about 2, 5, 8, 14 and 22 for age groups 65–69, 70–74, 75–79, 80–84 and +85, respectively.

Period effects, each compared to years of death 1989–1993, are about 0.7–0.8 for the 1970s and 1980s in Tyrol, and about 0.9 for each of the two decades for Austria excluding Tyrol; details are shown in Table 2; Fig. 2. In

Year of death	Tyrol		Austria excluding Tyrol		
	Number	ASR ^a	Number	ASR ^a	
1970	56	17.1	677	13.5	
1971	41	12.6	695	13.5	
1972	39	11.0	700	13.8	
1973	56	15.9	779	14.9	
1974	51	13.8	777	15.1	
1975	47	13.3	771	14.7	
1976	52	14.7	809	15.4	
1977	50	13.7	779	14.2	
1978	57	15.2	865	16.3	
1979	68	18.0	783	14.4	
1980	90	22.9	845	15.3	
1981	52	13.4	849	15.3	
1982	69	17.1	874	16.0	
1983	61	15.2	837	15.1	
1984	68	17.1	835	15.0	
1985	76	17.7	905	15.5	
1986	70	16.7	925	15.8	
1987	87	20.9	984	16.9	
1988	71	15.7	941	16.2	
1989	72	15.3	986	17.0	
1990	96	20.3	1,014	16.8	
1991	96	21.0	1,110	18.3	
1992	91	18.3	1,048	17.1	
1993	96	20.4	1,081	17.7	
1994	95	19.5	993	16.1	
1995	93	19.2	1,109	17.4	
1996	91	17.8	1,079	16.9	
1997	88	15.9	1,096	16.9	
1998	60	11.3	1,079	16.1	
1999	79	14.2	1,143	16.9	
2000	79	13.7	1,150	16.5	
2001	85	14.8	1,099	16.1	
2002	79	14.0	1,059	15.2	
2003	68	11.6	1,092	15.5	
2004	59	10.2	1,080	14.8	
2005	62	9.7	1,035	13.7	
2006	68	10.5	1,015	12.8	
2007	74	11.2	992	12.1	
2008	63	9.0	1,121	12.9	

^a Age-standardized rate per 100,000 using SEGI weights

the time period after optional PSA testing was introduced for all men in Tyrol, for Tyrol we observe effects of 0.97 (95% confidence interval (CI) 0.84, 1.12), 0.86 (95% CI 0.72, 1.04) and 0.70 (95% CI 0.57, 0.87) for time periods 1994–1998, 1999–2003 and 2004–2008, respectively. For Austria excluding Tyrol, we observe effects of 1 (95% CI



Fig. 1 Prostate cancer mortality: age-standardized rate in Tyrol and in Austria excluding Tyrol for years of death 1970–2008 (*vertical line* is end of previous publication)

0.96, 1.04), 1.03 (95% CI 0.98, 1.08) and 0.92 (95% CI 0.87, 0.97), again for time periods 1994–1998, 1999–2003 and 2004–2008, respectively. The confidence intervals for the risk ratios for Tyrol and for Austria excluding Tyrol for 2004–2008 overlap only at the border.

For Tyrol, the cohort effects are about 1.5 until 1916, after which we see a decrease to about 0.8 for cohorts born after 1930. For Austria excluding Tyrol, the cohort effects are rather stable with estimators between 1.20 and 1.40 and about 1 for cohorts born after 1927.

By extending and/or subsetting the age of PCA deaths in the analysis for the time period 2004–2008, we observe estimators of 0.70 (95% CI 0.51, 0.95), 0.47 (95% CI 0.30, 0.75) and 0.40 (95% CI 0.26, 0.61) for age groups 50+, 50-74 and 60-74, respectively, each compared to 1989–1993. Details of period estimators are shown in Table 3.

Discussion

Our analysis is based on an observational study of PCA mortality data conducted in the population of Tyrol, where PSA testing has been offered to men free of charge since it was introduced in the early 1990s. For Tyrol, we observed a mortality reduction of 14% for 1999–2003 and 30% for 2004–2008, and for Austria excluding Tyrol a mortality reduction of 8% for 2004–2008. The reduction for 2004–2008 is statistically significant for both Tyrol and Austria excluding Tyrol.

Are observations for Tyrol and for Austria excluding Tyrol plausible?

Our main study design is a historical comparison within Tyrol. The second comparison, although not formally

🖄 Springer

 Table 2 Results of the age-period-cohort model, drift in cohort

	Tyrol		Austria excluding Tyrol	
	Estimator	95% CI	Estimator	95% CI
Age				
60–64	1.00	Reference	1.00	Reference
65–69	2.05	1.65, 2.57	2.24	2.12, 2.38
70–74	4.63	3.73, 5.74	4.38	4.14, 4.64
75–79	7.86	6.23, 9.92	8.43	7.93, 8.97
80-84	13.29	10.23, 17.27	14.44	13.46, 15.48
+85	21.40	16.00, 28.62	23.65	21.88, 25.57
Period				
1970–1973	0.66	0.50, 0.86	0.86	0.80, 0.93
1974–1978	0.62	0.49, 0.79	0.88	0.83, 0.94
1979–1983	0.78	0.65, 0.94	0.86	0.82, 0.91
1984–1988	0.81	0.70, 0.95	0.91	0.87, 0.95
1989–1993	1.00	Reference	1.00	Reference
1994–1998	0.97	0.84, 1.12	1.00	0.96, 1.04
1999-2003	0.86	0.72, 1.04	1.03	0.98, 1.08
2004-2008	0.70	0.57, 0.87	0.92	0.87, 0.97
Cohort				
1882–1886	1.00	Reference	1.00	Reference
1887–1891	1.49	0.93, 2.38	1.24	1.10, 1.40
1892–1896	1.59	1.05, 2.42	1.27	1.14, 1.41
1897-1901	1.74	1.18, 2.55	1.35	1.22, 1.49
1902-1906	1.55	1.08, 2.24	1.41	1.28, 1.55
1907–1911	1.40	0.98, 2.00	1.44	1.31, 1.58
1912–1916	1.42	0.99, 2.03	1.43	1.30, 1.58
1917-1921	1.21	0.84, 1.74	1.29	1.17, 1.42
1922-1926	0.97	0.66, 1.42	1.14	1.03, 1.27
1927-1931	0.79	0.53, 1.20	1.10	0.98, 1.23
1932-1936	0.79	0.50, 1.25	1.05	0.92, 1.18
1937–1941	0.80	0.48, 1.34	0.89	0.77, 1.03
1942–1946 ^a				

^a Because there is drift in cohort, there is no estimator for the last cohort 1942–1946

included in the model, is between Tyrol and Austria excluding Tyrol. In the absence of PSA testing, we have good reasons to assume parallel time trends in PCA mortality in both regions, because health services in general as well as the diagnosis and therapy for cancer patients are uniform throughout Austria. Therefore, it is likely that the reduction in PCA mortality in Tyrol is mostly due to PSA testing, which was the main difference in PCA management between Tyrol and Austria excluding Tyrol up to 2000. In Tyrol, PSA cutoff levels equalled to 2.5, 3.5, 4.5 and 6.5 in age groups 40–49, 50–59, 60–69 and 70–79, respectively, up to October 1996 and were cut in half to 1.25, 1.75, 2.25 and 3.25 in the same age groups mentioned above (Bartsch 2008). Nevertheless, some smaller part of



Fig. 2 Prostate cancer mortality in Tyrol and in Austria excluding Tyrol for years of death 1970–2008, age 60+: estimated period effects

the reduction could be a consequence of other unknown factors. Whereas in Tyrol, PSA testing was introduced already in the early 1990s, in Austria excluding Tyrol PSA testing began some 7–10 years later. A look at PCA incidence (Statistics Austria 2008) supports this interpretation (see Fig. 3), because to our knowledge the introduction of PSA testing is the only possible explanation for the sharp increase in PCA incidence in Austria excluding Tyrol between 1995 and 2000. Hence, the 8% PCA mortality reduction in Austria excluding Tyrol could also be due to PSA testing.

Compared to the literature, what this study adds to known facts?

It is well known that observational studies are prone to a number of possible biases (e.g. confounding), and we expect much stronger evidence from randomized studies. To date, we know of results from four randomized studies on PSA screening and mortality, see (Boyle and Brawley 2009) for a discussion of the randomized studies. However, due to contamination and attendance, the true value of screening could have been underestimated (van Leeuwen et al. 2010). Therefore, efforts have been taken to control for contamination and attendance in the ERSPC studies, and after correcting for these biases our result is in line with other results, see for example (Robool et al. 2009, van Leeuwen et al. 2010, Kerkhoff et al. 2010, Hugosson et al. 2010), bearing in mind that our approach is rather conservative because we have no age limit on PCA mortality cases.

The number needed to treat was reduced from 1,410 in the ERSPC study (Schröder et al. 2009) to 293 in the Göteburg study (Hugosson et al. 2010). Therefore, we feel it is time to revise the conclusion drawn by Boyle and Brawley, which was based on the original ERSPC results.

Table 3 Sensitivity analysis of the age-period-cohort model by subsetting age groups, period estimators

	Age 50+		Age 50–74		Age 60–74	
	Estimator	95% CI	Estimator	95% CI	Estimator	95% CI
Period						
1970–1973	0.66	0.44, 0.99	0.62	0.34, 1.13	0.59	0.37, 0.96
1974–1978	0.63	0.45, 0.87	0.62	0.3, 1.03	0.59	0.37, 0.93
1979–1983	0.78	0.61, 0.99	0.73	0.49, 1.11	0.70	0.46, 1.04
1984–1988	0.82	0.70, 0.98	0.99	0.73, 1.36	0.96	0.69, 1.34
1989–1993	1.00	Reference	1.00	Reference	1.00	Reference
1994–1998	0.95	0.80, 1.12	0.86	0.64, 1.17	0.90	0.66, 1.23
1999–2003	0.86	0.68, 1.09	0.62	0.42, 0.92	0.55	0.37, 0.80
2004-2008	0.70	0.51, 0.95	0.47	0.30, 0.75	0.40	0.26, 0.61



Fig. 3 Prostate cancer incidence: age-standardized rate in Tyrol and in Austria excluding Tyrol for years of diagnosis 1983–2007

The advantage of our study is that it permits us to investigate the effect of PSA screening in a real-life situation. While randomised studies concentrate on the efficacy of screening, we analyze the effectiveness of a PSA testing program conducted in a well-defined population. But, of course, we cannot overcome problems that are inherent to nonrandomised studies.

In order to rule out the possibility that the model choice had a predominant effect, we compared our results with results obtained with a joinpoint regression model. We applied the SEER package (Joinpoint 2010, Kim et al. 2000), see Fig. 4. The results fit to the results obtained with the APC model, with a reduction in Tyrol beginning at about 1990 and in Austria excluding Tyrol at about 2000. The size of the mortality decrease given by the joinpoint regression model is even greater, meaning our estimates are rather conservative. In addition, we conducted a sensitivity analysis of the influence of age by extending and/or subsetting the analysis to age groups 50+, 50-74 and 60-74 (see Table 3) and obtained stable estimates.



Fig. 4 Prostate cancer mortality in Tyrol and in Austria excluding Tyrol for years of death 1970–2008: result of joinpoint regression for age-specific rate age 60+. Annual percentage change for Tyrol 2.8* (1970–1992), -5.3* (1992–2008) and for Austria excluding Tyrol 3.6* (1970–1978), 1.4* (1978–1991), -0.3 (1991–1999), -3.8* (1999–2006), 2.4 (2006–2008), *denoting statistical significance

Limitations

The facts that PSA testing was already introduced in Tyrol around 1990 and that about three-fourths of all men aged 45–74 have undergone at least one PSA test for screening purposes seem to be good reasons for conducting this analysis. However, we are faced with severe limitations.

First, and this is probably the most severe limitation, our analysis is based on an observational design, which does not allow for any control of confounders. We conducted a comparison with a historical control group and attributed the main effect to PSA testing. Of course, we cannot rule out the possibility that other factors contributed to the reduction in PCA mortality.

Second, the outcome measure we analyzed is PCA mortality, and we have no validation of PCA as cause of death. We know that mortality statistics in Austria has been of high quality for decades (Hansluwka 1989). Linking

🖄 Springer

mortality data with the Cancer Registry data gives an estimate of false-positive PCA deaths: about 95% of those PCA deaths showed a PCA diagnosis in the Cancer Registry database. This proportion does not change over time (data not shown).

Third, we have no detailed knowledge of the volume of PSA testing. For Tyrol, we collected data from all PSA labs and estimated the PSA testing rate. After 9 years of intense PSA testing, we estimate that 75.1% of all men aged 45–74 in Tyrol have had at least one PSA screening test (Oberaigner et al. 2006).

Fourth, we have only a very limited data on harms caused by PSA testing. Pelzer analyzed 1,445 consecutive patients undergoing radical prostatectomy at the Department of Urology of Innsbruck Medical University and concluded that the rate of overdiagnosis is small (between 8 and 17%), but also noticed underdiagnosis (Pelzer 2008).

After weighing all limitations and strengths, our personal estimate is that although these limitations exist and cannot be formally ruled out, it is unlikely that all possible biases could have caused a 50% reduction in PCA mortality in men aged 50–74, albeit some part of this reduction can be due to a combination of biases.

Our study concerns a well-defined population of Tyrol, where we have some knowledge of PSA testing rates and information on therapy offered to the population. The APC model fits well for Tyrol, and when compared to Austria excluding Tyrol the PSA testing rate seems to be the main factor which is able to explain the difference in time trends between Tyrol and Austria excluding Tyrol. Of course, our analysis cannot overcome the problems of nonrandomized studies, but it can provide further information on the potential benefits of PSA testing or screening.

Acknowledgments This work was supported by the Oncotyrol Center for Personal Cancer Medicine. Oncotyrol is a Comet Center funded by the Federal Ministry for Transport Innovation and Technology (BMVIT) and the Federal Ministry of Economics and Labor (BMWA), the Tyrolean Future Foundation (TZS) and the Land of Styria represented by the Styrian Business Promotion Agency (SFG) [and supported by UMIT—University for Health Sciences, Medical Informatics and Technology, Innsbruck Medical University, Tyrolean Health Insurance Fund and Tyrolean Health Company.

Conflict of interest None.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Andriole GL, Grubb RL, Buys SS et al (2009) Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 360:1310–1319

- Bartsch G, Horninger W, Klocker H et al (2001) Prostate cancer mortality after introduction of prostate-specific antigen mass screening in the Federal State of Tyrol, Austria. Urology 58:417–424
- Bartsch G, Horninger W, Klocker H et al (2008) Tyrol prostate cancer demonstration project: early detection, treatment, outcome, incidence and mortality. BJU 101:809–816
- Boyle P, Brawley O (2009) Prostate Cancer: current evidence weighs against population screening. CA Cancer J Clin 59:220–224
- Boyle P, Levin B (eds) (2008) World cancer report 2008. IARC, Lyon
- Clayton D, Schifflers E (1987a) Models for temporal variation in cancer rates. I. Age-period and age-cohort models. Stat Med 6:449–467
- Clayton D, Schifflers E (1987b) Models for temporal variation in cancer rates. II. Age-period-cohort models. Stat Med 6:469–481
- De Koning HJ, Auvinen A, Berenguer SA et al (2002) Large-scale randomized prostate cancer screening trials: program performances in the European randomized screening for prostate cancer trial and the prostate, lung, colorectal and ovary cancer trial. Int J Cancer 97:237–244
- Hansluwka H, Cancer Mortality Statistics (1989) Availability and aspects of quality of mortality data worldwide. Springer, Berlin
- Hugosson J, Carlsson S, Aus G et al (2010) Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. Lancet Oncol 11:725–732
- Joinpoint Regression Program, Version 3.4.3 (2010) Statistical Research and Applications Branch. National Cancer Institute, Bethesda
- Kerkhof M, Roobol MJ, Cuzick J et al (2010) Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trail (ERSPC section Rotterdam). Int J Cancer 127:2639–2644
- Kim HJ, Fay MP, Feuer EJ, Midthune DN (2000) Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 19:335–551. (correction: 2001; 20:655)
- Oberaigner W, Horninger W, Klocker H et al (2006) Reduction of prostate cancer mortality in Tyrol, Austria after introduction of PSA testing. Am J Epidemiol 164(4):376–384
- Pelzer AE, Colleselli D, Bektic J et al (2008) Over-diagnosis and under-diagnosis of screen- vs non-screen-detected prostate cancers with in men with prostate-specific antigen levels of 2.0–10.0 ng/mL. BJU Int 101:1223–1226
- Robool MJ, Kerkhof M, Schröder FH et al (2009) Prostate cancer mortality reduction by prostate-specific antigen-based screening adjusted for nonattendance and contamination in the European randomised study of screening for prostate cancer (ERSPC). Eur Urol 56:584–591
- Schröder FH, Hugosson J, Roobol MJ et al (2009) Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 360:1320–1328
- SEER (2011) http://www.surveillance.cancer.gov/statistics/types/ lifetime_risk.html. (Accessed 19 April 2011)
- Stata Statistical Software (2005) Release 9.0. StataCorp LP, Stata-Corp, College Station, Texas
- Statistik Austria (2008) Jahrbuch der Gesundheitsstatistik 2007. Wien
- van Leeuwen PJ, Connolly D, Gavin A et al (2010) Prostate cancer mortality in screen and clinically detected prostate cancer: estimating the screening benefit. Eur J Cancer 46:377–383

B1-9 Appendix Full Papers

Oberaigner et al. BMC Public Health 2011, 11:673 http://www.biomedcentral.com/1471-2458/11/673

RESEARCH ARTICLE



Open Access

Introduction of organised mammography screening in Tyrol: results following first year of complete rollout

Willi Oberaigner^{1,2,3*}, Martin Daniaux⁴, Sabine Geiger-Gritsch¹, Rudolf Knapp⁵, Uwe Siebert^{2,3,6,7} and Wolfgang Buchberger^{2,8}

Abstract

Background: In Tyrol, Austria, the existing system of spontaneous mammography screening was switched in 2007 to an organised program by smoothly changing the established framework. This process followed most EU recommendations for organised mammography screening with the following exceptions: women aged 40-49 are part of the target population, screening is offered annually to the age group 40-59, breast ultrasound is available as an additional diagnostic tool, and double reading has not yet been implemented. After a pilot phase the program was rolled out to all of Tyrol in June 2008. The aim of this study was to analyse the performance of the organised screening system by comparing quality indices and recommended levels given in the well-established EU guidelines.

Methods: Working from the results of the pilot phase, we extended the organised mammography system to all counties in Tyrol. All women living in Tyrol and covered by compulsory social insurance were invited for a mammography, in the age group 40-59 annually and in the age group 60-69 biennially. Screening mammography was offered mainly by radiologists in private practice, with further assessment performed at hospitals. Using the screening database, all well-established performance indicators were analysed and compared with accepted/desired levels as per the EU guidelines.

Results: From June 2008 to May 2009, 120,440 women were invited. Per 1000 mammograms, 14 women were recalled for further assessment, nine underwent biopsy and four cancer cases were detected. Of invasive breast cancer cases, 32.3% and 68.4% were \leq 10 mm and \leq 15 mm in size, respectively, and 79.2% were node-negative. The positive predictive value for further assessment and for biopsy was 25.9% and 39.9%, respectively. Estimated two-year participation rate was 57.0%. In total, 14 interval cancer cases were detected during one year of follow-up; this is 18.4% of the background incidence rate.

Conclusions: In Tyrol, Austria, an organised mammography screening program was implemented in a smooth transition from an existing spontaneous screening system and was completely rolled out within a short time. The high level of performance already seen in the pilot phase was maintained after rollout, and improvements resulting from the pilot phase were affirmed after one year of complete rollout.

Background

Breast cancer is the leading cause of female cancer death in all industrialised countries (and also worldwide), and the breast is also the leading incident cancer site for females [1]. Therefore, screening methods for breast cancer are of greatest public health importance.



In 2006, in Tyrol, Austria, the decision was made to change the existing spontaneous mammography screening system to an organised program while, on the one hand, making best possible use of the mammography screening network established over the previous fifteen years and, on the other hand, following most EU recommendations for



© 2011 Oberaigner et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

^{*} Correspondence: willi.oberaigner@iet.at

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Cancer Registry of Tyrol, Innsbruck, Austria

Full list of author information is available at the end of the article

organised mammography screening. After a pilot phase conducted in two central counties of Tyrol covering 40% of the population from June 2007 to May 2008 [3], the organised system was completely rolled out to all of Tyrol in June 2008. It was possible to establish a country-wide mammography screening program in a very short time, which differs only in the following aspects from the EU guidelines [4]: women aged 40-49 are part of the target population, screening is offered annually in the age group 40-59, breast ultrasound is available as an additional diagnostic tool, and double reading has not yet been implemented.

To our knowledge, some European countries still have no organised mammography screening program or are in the process of planning to set up such a system [5,6]. Therefore, the Tyrolean experience can make an important contribution to deciding how to switch a health system with spontaneous mammography screening to an organised screening program that meets well-accepted quality guidelines.

It was the aim of this study to analyse the performance of the organised mammography screening system after complete rollout to all counties in Tyrol by measuring the quality indicators recommended by the EU guidelines [4] and to determine whether the high quality observed in the pilot phase could be affirmed after rollout.

Methods

Study population, invitation

The target population in the first year of complete rollout from June 2008 to May 2009 included all women aged 40 to 69 living in Tyrol and covered by compulsory social insurance, which is more than 97% of the population (personal communication). The main health insurance carrier sent out personal invitation letters to all the women in the target population in the month in which the women had their birthday: women aged 40-59 annually, and women aged 60-69 biennially. As women aged 60-69 and living in the two central counties of Tyrol where the pilot phase was conducted had already been invited in the pilot year, this group of women was not invited again in the first year of rollout (Figure 1). Mammography screening was offered by 22 screening units; thirteen of them were run by radiologists in private practice and nine by hospital outpatient departments. The mammogram was read by only one radiologist; ultrasound (US) was offered to women at the radiologist's discretion. Assessment was offered by nine hospital radiology units in the study area and included clinical inspection, mammography, US, magnetic resonance imaging (MRI) and biopsy as needed. Women were recalled for assessment either directly by the screening unit or by the general practitioner. The one large assessment unit at Innsbruck Medical University Hospital works closely with a breast cancer centre that was EUSOMA-certified in March 2010 [7]. All radiologists participating in the program underwent training and received ÖRG (Austrian Radiology Association) certification. In the median, private radiologists and hospital units performed 3234 and 1639 mammograms per year, respectively. The mammography screening system has been described in more detail elsewhere [3].

Data collection

All mammography units registered basic information in a database. Screening and assessment information was transferred to a central database after pseudonymising the woman's social insurance number [3]. In addition, data on tumour characteristics were collected by the Cancer Registry of Tyrol.

Statistical analysis

The screening and assessment data were realised as STATA datasets. Linkage between screening data, assessment data and Cancer Registry data is based on the pseudonym number. We reported numbers and proportions as defined in the EU guidelines [4]. For some indices, population-based rates were computed using the official population data supplied by Statistics Austria. No statistical testing was applied. All reporting was done with STATA Version 11 [8].

Performance indicators were reported from all screens in women aged 40-69 between June 2008 and May 2009.

Participation rate was calculated following a cohort approach: we counted every woman only once in the observation period, which was either one year or two years. Due to the fact that nearly half of women aged 40 to



59, who attend screening regularly, do not return for screening in the first year although they are invited annually, we computed for that age group also a two-year participation rate, meaning an observation period of two years.

Data on all mammography investigations performed in Tyrol are transferred to the screening database. A small portion (5.9%) of the women refused consent for data transfer to the screening database and we therefore receive only an empty dataset. Of all other mammography data, 76% belong to the screening population. By assuming this same proportion of 76% for the empty dataset, we calculated a proportion of 4.5% to be added to the observed participation rate accounting for empty datasets describing real numbers of mammography screening investigations.

As spontaneous mammography screening was already introduced to Tyrol in the early 1990s, the underlying background incidence rate (BIR) was defined by years of diagnosis 1988-1990.

This study was conducted in conformity with the Helsinki Declaration [9]. The project was approved by the Ethics Committee of Innsbruck Medical University.

Results

From June 2008 to May 2009 120,440 women in the target population were invited; this excluded women aged 60-69 and living in the two central counties where the pilot phase was conducted, who fell in the biennial screening interval for that age group and were thus not invited again in this first year of complete rollout (Figure 1). The observed one- and two-year participation rates were 31.6% and 52.5%, respectively (Table 1). Participation was higher in younger women. For example, the two-year observed participation rate was 55.1% in women aged 40-49 versus 50.3% in women aged 50-69.

Performance indicators were analysed for all screens performed in the first year of rollout, namely 42,834 screens. Of the women 75.5% underwent additional US (80.9% in women aged 40-49). Breast density (ACR 3/4) was the reason for additional US in 52.7% and 39.6% of women aged 40-49 and 50-69, respectively (Table 2).

Per 1000 screens, 14 women were recalled for further assessment. Screening result was unknown for a total of

Table 1 Invitation system: Number of women invited and participation rates

	40-49	50-69	Total (40-69)
Women invited	56,888	63,552	120,440
Observed one-year participation rate	32.6%	30.6%	31.6%
Estimated one-year participation rate	37.1%	35.1%	36.1%
Observed two-year participation rate	55.1%	50.3%	52.5%
Estimated two-year participation rate	59.6%	54.8%	57.0%

98 cases (0.2% of screens). Per 1000 screens, nine underwent biopsy. Of all biopsies, 86% were core biopsies and 3% open biopsies (13 cases). We observed 3.6 screendetected cancers per 1000 screens or a total of 153 breast cancer cases, of which 9.2% were diagnosed as in situ cancers. The positive predictive value (PPV) was 25.9% for further assessment, 39.9% for total biopsy and 45.8% for core biopsy. PPV was lower in age group 40-49 (18.7%, 31.3% and 34.9% for further assessment, total biopsy and core biopsy, respectively). Performance parameters are summarised in Table 3.

Of 139 invasive cancers diagnosed in screening, four changed to "in situ cancer" after final diagnosis. Two invasive cancer cases did not undergo surgery because of metastatic disease. Finally, three invasive cancer cases underwent neoadjuvant therapy and it was not possible to identify preoperative staging.

Of all invasive cancers detected and finally proven, 32.3% and 68.4% showed tumour size $\leq 10 \text{ mm}$ and $\leq 15 \text{ mm}$, respectively. Lymph node involvement was observed in 20.8% of invasive cancer cases (Table 4).

For invasive cancers, 90.6% of further assessments were carried out within five working days after screening and 87.1% and 90.1% underwent surgery within ten days and 15 days after decision to operate, respectively. For all cases except invasive cancers, 73.7% underwent assessment within five working days and 17.1% after ten or more working days (Table 5).

We observed a total of 14 interval cancer cases within one year after screening in all of Tyrol, five in age group 40-49, giving an interval cancer rate of 20.8% and 17.8% of the background incidence rate for age groups 40-49 and 50-69, respectively (Table 6). Table 7 shows the results for the most important quality indicators of the EU guidelines [4] restricted to age group 50-69. Given that the organised system was introduced after more than a decade of spontaneous mammography screening in Tyrol, as reference values we chose the accepted and desired ranges of EU quality indicators for subsequent rounds. Most of the indicators were within the EU range, except the participation rate (54.8% vs. the limit of 75%), the proportion of II+ cancers (33.3% vs. the limit of 25%), the proportion of invasive cancers (91.2%, this is slightly above the limit of 90%) and the proportion of cases that underwent surgery within \leq 15 working days after decision to operate (87.3%, this is slightly below the limit of 90%).

Discussion

We analysed performance after one year of rolling out an organised mammography screening program to all counties in Tyrol. The organised program was established in a smooth transition from an existing spontaneous mammography screening system, instead of setting up a completely new screening system, and was

Table 2	Additional	ultrasound	at	screening
---------	------------	------------	----	-----------

	40-49	50-69	Total
Ultrasound following mammography screening	15,126 (80.9%)	17,196 (71.2%)	32,322 (75.5%)
Reason for ultrasound:			
Breast density (ACR 3/4)	7,971 (52.7%)	6,806 (39.6%)	14,777 (45.7%)
Equivocal finding	1,801 (11.9%)	2,318 (13.5%)	4,119 (12.7%)
Other	5,354 (35.4%)	8,072 (46.9%)	13,426 (41.5%)

Table 3 Performance parameters

	40-49 ¹⁾	50-69 ¹⁾	Total ¹⁾
Recall for further assessment rate [per 1000 screens] and number of recalls ²)	14.6 (273)	13.2 (318)	13.9 (591)
Intermediate screening test recommended in six months	18.9 (354)	12.1 (292)	15.1 (646)
Screening result unknown ³⁾	2.8 (52)	1.9 (46)	2.3 (98)
Biopsy rate [per 1000 screens]	8.7 (163)	9.1 (220)	8.9 (383)
Cancer detection rate [per 1000 screens]	2.7 (51)	4.2 (102)	3.6 (153)
Invasive	2.5 (46)	3.9 (93)	3.3 (139)
In situ	0.3 (5)	0.4 (9)	0.3 (14)
Proportion of in situ cases	9.8%	8.8%	9.2%
Ratio screening breast cancer detection rate vs. background incidence rate ⁴⁾	2.1	2.0	2.0
PPV assessment	18.7% (51/273)	32.1% (102/318)	25.9% (153/591)
PPV biopsy	31.3% (51/163)	46.4% (102/220)	39.9% (153/383)

¹⁾ Rates were rounded to one decimal; numbers in brackets are numbers of cases (i.e. recall rate, biopsy rate, cancer detection rate). PPV is rounded to one decimal; numbers in brackets are detailed numbers for computing PPV.

²⁾ For one case assessment was recommended, but performed at an institution outside the screening system.

³⁾ Cases with BI-RADS 0 without assessment were treated as unknown.

⁴⁾ Background incidence rate defined by years of diagnosis 1988-1992.

previously tested in a pilot phase comprising 40% of the target population [3]. Although not all EU recommendations were followed, most quality indicators are in the range of accepted/desired levels given by the EU guidelines [4]. The only parameter that clearly missed the EU guidelines was the participation rate: the two-year participation rate was 57% as compared to the 75% recommended by the EU guidelines. In our opinion, a cumulative participation rate of 57% after two years of observation looks successful when compared to neighbouring countries [10-12]. Nevertheless, it is not the goal we aimed for.

Table 4 Characteristics of invasive cancer cases

	40-49	50-69	Total
Tumour size (mm): N = 133	13; 4-25	12; 1-35	13; 1-35
Median; range			
Tumour size (mm):			
< = 10 mm	14 (32.6%)	29 (32.2%)	43 (32.3%)
< = 15 mm	28 (65.1%)	63 (70.0%)	91 (68.4%)
11-20 mm ¹⁾	23 (53.5%)	42 (46.7%)	65 (48.9%)
> 20 mm	6 (14.0%)	19 (21.1%)	25 (18.8%)
Lymph node involvement	8 (18.6%)	19 (21.8%) ¹⁾	27 (20.8%)
Staging according to UICC			
I	30 (69.8%)	55 (61.8%)	85 (64.4%)
II	13 (30.2%)	31 (34.8%)	44 (33.3%)
III		3 (3.4%)	3 (2.3%)

Notes: Of 139 invasive cancer cases, four cases were finally "in situ"; two invasive cancer cases did not undergo surgery because of metastatic status. Three cases were without lymph node status because of neoadjuvant therapy and because we could not identify pretherapeutic TNM stage.
Table 5 Waiting times

Invasive cancers				
	40-49	50-69	Total	
Screening to assessment				
≤ 5 wd	40 (87.0%)	86 (92.5%)	126 (90.6%)	
6-10 wd	3 (6.5%)	3 (3.2%)	6 (4.3%)	
> 10 wd	3 (6.5%)	4 (4.3%)	7 (5.0%)	
Decision to operate to date of therapy	1			
≤ 10 wd	42 (93.3%)	73 (83.9%)	115 (87.1%)	
11-15 wd	1 (2.2%)	3 (3.4%)	4 (3.0%)	
16-30 wd	1 (2.2%)	4 (4.6%)	5 (3.8%)	
> 30 wd	1 (2.2%)	7 (8.0%)	8 (6.1%)	
All screens except those ending in	invasive cancers			
	40-49	50-69	Total	
Screening to assessment				
≤ 5 wd	162 (72.3%)	166 (75.1%)	328 (73.7%)	
6-10 wd	17 (7.6%)	24 (10.9%)	41 (9.2%)	
> 10 wd	45 (20.1%)	31 (14.0%)	76 (17.1%)	

wd: working days

The **strengths** of the Tyrolean breast cancer screening program are its implementation and performance: we were able to set up an organised population-based screening program within a short time with minimal additional resources that shows good performance. In summary, the recall for further assessment rate and the biopsy rate are fairly low, PPV was good as compared to other programs, only few open biopsies were performed, and despite the lack of double reading the interval cancer rate of 20% of the underlying BIR is rather good as compared to other programs [10,13-15].

However, this study has several weaknesses. First, up to now we have not implemented double reading as recommended in the EU guidelines. Interestingly, performance parameters and especially interval cancer rate showed that also without double reading an acceptable quality level was achieved. One reason could be the extensive use of additional US, about three of four women underwent additional US. The real benefit of US in a population-based mammography screening program is currently under discussion and has to be further evaluated [16,17]. Calculation of the interval cancer rate is reliant on the completeness of the Cancer Registry of Tyrol, which covers the target population. Completeness of incidence data in general has been shown to be very good [3,18]. In order to be able to analyse interval cancer rates for the screening program the timeliness of registration of breast cancer was improved, and linkage between cancer registry data and screening data is based on pseudonymising the social insurance number, which is read electronically. In the meantime, we have also assessed interval cancer in the time window 12 to 23 months for the pilot phase of the Tyrol program, see [3], and found five interval cancer cases in age group 40-49 (55% of BIR) and seven interval cancer cases in age group 50-69 (33% of BIR), data not shown.

Second, the average number of screens read by a radiologist in Tyrol per year (about 3200) does not meet the EU recommendation of 5000. A recent publication [19] showed that annual numbers below 5000 can still provide good sensitivity and acceptable false-positive rates.

Third, we used BI-RADS categories instead of a single yes/no rule for recall for further assessment. Some radiologists still use BI-RADS 0 (meaning unclear result) in a small number of cases (0.2% of all screens), and 15 per 1000 screens were invited to an intermediate screening test six months following a BI-RADS 3 screening result. Due to this inconsistency, the current program includes the following modifications: BI-RADS 0 is no longer allowed and BI-RADS 3 is strictly associated with recall for further assessment.

Many countries have run a mammography screening program for decades or for a shorter time. On the other hand, there are still some countries with no organised

Table 6 Interval cancer rate within first year

	40-49	50-69	Total
Interval cancer rate per 100,000 screens (number of cases in brackets)	26.7 (5)	37.3 (9)	32.7 (14)
Proportion of background incidence rate ¹⁾ (in percent)	20.8%	17.8%	18.4%

¹⁾ based on years of diagnosis 1988-1990

	Tyrol 50-69	EU-accepted ¹⁾	EU-desired ¹⁾
Participation rate (after two years of observation)	54.8%	> 70	> 75
Recall for further assessment rate ¹⁾	1.3 (318)	< 5%	< 3%
Breast cancer detection rate	2.0 * BIR	1.5*BIR	> 1.5*BIR
Interval cancer rate/Background incidence rate (BIR) 0-11 months	18% (9)	30%	< 30%
Proportion of screen-detected cancers that were invasive	91.2% (93/102)	90%	80%-90%
Proportion of screen-detected cancers that were stage II+	33.3% (34/102) ²⁾	25	< 25%
Node-negative cancer/Total invasive cancers screen-detected	78.2%	75%	> 75%
Invasive cancers ≤ 10 mm/Total invasive cancers	32.2%	≥25%	≥30%
Proportion of invasive cancers that were \leq 15 mm in size	70.0%	50%	> 50%
Time between results of screening and assessment $< = 5 \text{ wd}^{3}$	92.5%	90%	> 90%
Time between decision to operate and surgery $< = 15 \text{ wd}^{4)}$	87.3%	90%	> 90%

Table 7 EU Guidelines, quality indicators (with accepted and desired levels)

¹⁾ We took reference values for subsequent rounds, because we had a decade of spontaneous screening before beginning the organised program.

²⁾ One of 150 cases without UICC staging (no staging because of neoadjuvant therapy).

³⁾ We show time between screening and performed assessment, not offered assessment.

⁴⁾ We show time between decision to operate and surgery performed (not between decision to operate and date offered for surgery).

breast cancer screening program. For those countries thinking of or already in the process of introducing a mammography screening program, our manner of introducing an organised program can serve as one how-to example. In our opinion, the greatest difference between our approach and other approaches, especially compared to Germany, is the smooth transition made from an existing spontaneous program to an organised population-based screening. We made use of the network of screening and assessment units that had already been set up during spontaneous screening and added an invitation system covering the entire population of Tyrol, a screening database that allows quality indices to be monitored and a well-defined training program for both screening and assessment units. With this strategy we were able to meet most EU quality indices within a very short time.

Conclusions

In Tyrol, Austria, an organised mammography screening system realised in a smooth transition from an existing spontaneous screening system was rolled out in a short time. The high level of performance already observed in the pilot phase has not changed after the first year of complete rollout. Improvements suggested during the pilot phase were affirmed after rollout: it will be necessary to concentrate on efforts to improve the participation rate, introduce double reading, change the rule for BI-RADS 3, and reduce the number of additional ultrasound exams.

List Of Abbreviations

US: ultrasound, ÖRG: Austrian Radiology Association, ACR: American College of Radiology, MRI: magnetic resonance imaging, BIR: background incidence rate, PPV: posivite predictive value.

Acknowledgements

We thank Patricia Gscheidlinger for secretarial support. Mammography Screening Project Tyrol Working Group: Radiologists: B. Bischof, Innsbruck; J. Bliem, Wörgl; M. Daniaux, Innsbruck; A. Dessl, Zams; G. Egender, Hall/Tirol; W. Hönlinger, Innsbruck; W. Kastlunger, Schwaz; R. Knapp, Kufstein; A. Lederer, Lienz; C. Lottersberger, Schwaz; D. Lungenschmid, Innsbruck; A. Oberhauser, Innsbruck; E. Partl, Kitzbühel; T. Penz, Innsbruck; C. Rapf, Lienz; E. Schmaranzer, St. Johann; G. Schön, Telfs; R. Schöpf, Landeck; F. Weber, Hall in Tirol; K. Wicke, Innsbruck; C. Wolf, Reutte **Project Team**: W. Buchberger, M. Daniaux, C. Greil-Thum, R. Knapp, G. Nestler, W. Oberaigner, B. Schreiner, K. Wicke

This work was supported by the ONCOTYROL Center for Personal Cancer Medicine. ONCOTYROL is a COMET Center funded by the Austrian Federal Ministry of Transport, Innovation and Technology (BMVIT) and the Austrian Federal Ministry of Economics, Family and Youth (BMWFJ), the Tyrolean Future Foundation (TZS) and the State of Styria represented by the Styrian Business Promotion Agency (SFG) and supported by UMIT - University for Health Sciences, Medical Informatics and Technology, Innsbruck Medical University, Tyrolean Health Insurance Fund and Tyrolean Health Company.

Author details

¹Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Cancer Registry of Tyrol, Innsbruck, Austria. ²Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria. ³ONCOTYROL - Center for Personalized Cancer Medicine, Innsbruck, Austria. ⁴Innsbruck Medical University, Department of Radiology, Innsbruck, Austria. ⁵Kufstein County Hospital, Department of Radiology, Kufstein, Austria. ⁶Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA. ⁷Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. ⁸Tyrolean State Hospitals Ltd., Medical Department, Innsbruck, Austria.

Authors' contributions

WO and WB designed the study. WO performed the analysis and wrote the paper. SGG contributed to writing the paper and to critically reviewing the draft. WB, MD und RK contributed to the Discussion section, especially from the radiology point of view. US contributed to writing the paper. All authors reviewed and agreed to the final version of the manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 20 May 2011 Accepted: 30 August 2011 Published: 30 August 2011

References

1. Boyle P, Levin B, Eds: World Cancer Report 2008 Lyon: IARC Press; 2008.

- 2. Gotzsche PC, Nielsen M: Screening for breast cancer with mammography. Cochrane Database Syst Rev 2011, CD001877.
- Oberaigner W, Buchberger W, Frede T, Daniaux M, Knapp R, Marth C, Siebert U: Introduction of organised mammography screening in tyrol: results of a one-year pilot phase. *BMC Public Health* 2011, 11:91-99.
- Perry N, Broeders M, de Wolf C, (Eds.): European guidelines for quality assurance in breast cancer screening and diagnosis. Luxembourg: Office for Official Publications of the European Communities; 2006.
- Karsa L, Anttila A, Ronco G, Ponti A, Malila N, Arbyn M: Cancer Screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening. First Report Brussels: European Commission; 2008.
- Dowling EC, Klabunde C, Patnick J, Ballard-Barbash R: Breast and cervical cancer screening programme implementation in 16 countries. J Med Screen 2010, 17:139-146.
- Eusoma-Breast Unit-Certification Process. [http://www.eusoma.org/Engx/ BreastUnits/AccreditationProcess.aspx?cont=ap_accredited], accessed 09.05.2011.
- 8. Stata Statistical Software: Release 11 College Station, Tx, StataCorp LP; 2009.
- World Medical Association Declaration of Helsinki: ethical principles for medical research involvin human subjects. [http://www.wma.net/en/ 30publications/10policies/b3/index.html], (acessed 7 November 2009).
- Kooperationsgemeinschaft Mammographie, (Ed.): Evaluationsbericht 2005-2007. Ergebnisse des Mammographie-Screening-Programms in Deutschland. Köln: 2009.
- Bulliard JL, Ducros C, Jemelin C, Arzel B, Fioretta G, Levi F: Effectiveness of organised versus opportunistic mammography screening. *Ann Oncol* 2009, 20:1199-1202.
- Giordano L, Giorgi D, Ventura L, Stefanini V, Senore C, Castagno R, Paci E, Segnan N: Time trends of process and impact indicators in Italian breast screening programmes: 1998-2008. *Epidemiol Prev* 2010, 34:27-34.
- Hofvind S, Vacek PM, Skelly J, Weaver DL, Geller BM: Comparing screening mammography for early breast cancer detection in Vermont and Norway. J Natl Cancer Inst 2008, 100:1082-1091.
- Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ: Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. *BMJ* 2000, 321:665-669.
- Smith-Bindman R, Chu PW, Miglioretti DL, Sickles EA, Blanks R, Ballard-Barbash R, Bobo JK, Lee NC, Wallis MG, Patnick J, Kerlikowske K: Comparison of screening mammography in the United States and the United kingdom. JAMA 2003, 290:2129-2137.
- Corsetti V, Houssami N, Ghirardi M, Ferrari A, Speziani M, Bellarosa S, Remida G, Gasparotti C, Galligioni E, Ciatto S: Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1year follow-up. Eur J Cancer 2011. 47:1021-1026.
- 17. Husby JA, Espeland A, Kalyanpur A, Brocker C, Haldorsen IS: Double reading of radiological examinations in Norway. Acta Radiol 2011.
- 18. Oberaigner W, Siebert U: Are survival rates for Tyrol published in the Eurocare studies biased? *Acta Oncol* 2009, 1-8.
- Buist DS, Anderson ML, Haneuse SJ, Sickles EA, Smith RA, Carney PA, Taplin SH, Rosenberg RD, Geller BM, Onega TL, et al: Influence of annual interpretive volume on screening mammography performance in the United States. Radiology 2011, 259:72-84.

Pre-publication history

The pre-publication history for this paper can be accessed here: http://www.biomedcentral.com/1471-2458/11/673/prepub

doi:10.1186/1471-2458-11-673

Cite this article as: Oberaigner *et al.*: **Introduction of organised** mammography screening in Tyrol: results following first year of complete rollout. *BMC Public Health* 2011 **11**:673.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit

BioMed Central

420 © 2007

Schattauer GmbH

Errors in Survival Rates Caused by Routinely Used Deterministic Record Linkage Methods

W. Oberaigner

Cancer Registry of Tyrol, Department of Clinical Epidemiology of the Tyrolean State Hospitals Ltd., Innsbruck, Austria

Summary

Objective: It was the objective of this study to assess the impact of applying various record linkage methods to one of the most important outcome measures in oncological epidemiology, namely survival rates. Methods: To assess the life status of patients, incidence data published by the Cancer Registry of Tyrol were analyzed with three routinely used methods of record linkage for incidence and mortality data. Of these methods, two were deterministic and the third a probabilistic method developed by the Cancer Registry of Tyrol. We studied the impact of record linkage methods on a simple measure (mortality rate) and a more complex measure (relative survival rate). The analysis was based on the published incidence data for Tyrol for the years 1992 to 1996. Results of deterministic record linkage methods were simulated. **Results:** The error rates for simple mortality rate and relative survival rate are considerable. For the first deterministic record linkage method, relative differences in mortality rate range from 11.9% to 14.8% (men) and 24.5% to 28.2% (women) and relative differences in relative five-year survival from 11.4% to 16.3% (men) and from 19.3% to 26.4% (women). For the second deterministic record linkage method, relative differences in mortality rate range from 4.8% to 5.9% (men) and from 4.9% to 7.4% (women), while relative differences in relative five-year survival range from 5.1% to 7.0% (men) and from 4.4% to 6.1% (women).

Conclusions: Our study shows that in order to calculate valid mortality and survival rates a probabilistic method of record linkage must be applied.

Keywords

Record linkage, survival analysis, relative survival rate, mortality rate

Methods Inf Med 2007; 46: 420-424

doi:10.1160/ME0299

Introduction

Survival rates are some of the most important general outcome parameters in oncology, reflecting improvements in early detection, diagnosis and therapy of cancer patients [1]. Most cancer registries routinely report survival rates, and many publications have reported survival rates of populationbased cancer registries, on both a national and an international level. See for example [1-4].

Computation of survival rates calls for a method to assess the life status of every cancer case at time of analysis. This can be done either in an active way (this means the registry must actively ask for the status of every patient) or in a passive way (in most cases by linking mortality data to incidence data), see [2, 5]. Two kinds of error can occur when using the passive method:

- Homonym and synonym errors caused by the record linkage method. This means that a) date of death is not linked to an incident case that in fact has died, b) a date of death is linked to a case that in fact is still alive, or c) a wrong date of death is linked to a truly fatal case [6].
- 2) Errors caused by out-migrants. Usually, at least in Austria, mortality data are available only for the population covered by the particular cancer registry. If a cancer case leaves the population between date of incidence and date of death, then she/he never dies from the point of view of passive follow-up because the case is not included in the mortality file of the population under study.

Whereas some countries use unique person identifiers in various data sources and record linkage can thus be based on such an identifier (for example in Scandinavian countries), many European countries do not use a unique person identifier. Thus, record linkage must be based on components identifying the patient, like last name, first name, date of birth and so on.

In this situation, record linkage is a nontrivial problem, see for example [7-13]. We can distinguish between deterministic record linkage, meaning that two data records are linked only when certain components are completely identical in both datasets, and probabilistic record linkage procedures, which compute a probability of identity and can also take into account errors in registration, documentation or data input [14]. For this reason, probabilistic methods can link data even when there are differences in certain components.

Various studies have been conducted on homonym and synonym rates in connection with record linkage studies; see for example [6]. Our focus, however, was to assess the direct impact of record linkage methods on epidemiological measures. In epidemiological studies, at least in countries where there is no unique person identifier, assessment of patient life status is often based on record linkage and correctness of record linkage affects survival rates. Hence, in addition to homonym and synonym rates, the size of error caused by improper record linkage methods is of great importance to the epidemiological community.

It was the aim of this study to analyze the impact by various routinely used linkage methods on important epidemiological outcome measures used in survival analysis.

Methods

We compared the effects of three methods of record linkage for incidence and mortality data. The study is based on the dataset of

Downloaded from www.methods-online.com on 2011-11-30 | ID: 1000466760 | IP: 193.171.69.65

For personal or educational use only. No other uses without permission. All rights reserved Received: January 12, 2005; accepted: May 16, 2006

cancer incidence for Tyrol, which is one of nine Austrian states. In Austria, cancer incidence data are collected on a national level by Statistics Austria, a government agency entrusted with cancer registration by a law passed in 1968. Because incidence data on the national level did not reach sufficient completeness and in order to make data available at the regional level, so-called local Cancer Registries were established in some of the Austrian states between 1980 and 1990. Thus, in the state of Tyrol, cancer incidence data are collected by the Cancer Registry of Tyrol, which began work in 1986. Cancer data for the population of Tyrol have been registered on a population basis since 1988. Also since 1988, data have been published in Cancer Incidence in Five Continents [5, 15, 16], thus giving evidence of the good completeness of our incidence data.

In Tyrol, information on cancer cases is obtained by means of a standardized questionnaire requesting sex, age, cancer site and histology, date of diagnosis and stage. Strict rules ensure that these variables are collected in accordance with international guidelines. Either the questionnaire is completed by a physician, or a Cancer Registry clerk collects data directly from clinical records in the treating hospital. Two independent databases are built up, one incidence database and one so-called investigation database that includes all information on possible cancer diagnoses (mainly pathology reports, but also information from radiotherapy units and various other data sources), thus allowing the registry to check completeness.

The life status of cases is assessed passively by linking the incidence dataset and the complete mortality file for the population of Tyrol (collected by Statistics Austria); record linkage is routinely done with a probabilistic method described in detail in the next paragraph. In order to compare the effects of applying three methods of record linkage, we not only used the routine method, but also simulated patient life status according to two deterministic record linkage methods widely used by Austrian registries. Our main goal was to study the impact of applying various record linkage methods to important epidemiological measures. For this purpose we chose the mortality and relative survival rates, which are among the most important oncological outcome measures.

Relative survival rates are used because observed survival rates are influenced by noncancer mortality. Hence, relative survival rates, derived as the ratio of observed survival rates divided by the expected survival rates of subjects of corresponding age and sex in the general population, reflect the "net survival" related to the cancer of interest. In our analysis, the relative survival rates were estimated according to the method of Hakulinen [17] and were computed using the program Surv3 developed by the Finnish Cancer Registry [18]. This program uses two datasets, one describing the general population mortality files and the second describing patient data. By applying life table methods, both observed and relative survival figures are obtained. Parameter files allow adjustment for various conditions.

The first deterministic record linkage method (in the following denoted Det1) links persons if and only if last name, first name, date of birth and sex are completely identical, see Table 1. Date of death is added to a case if and only if the components listed above are identical in both datasets; in all other situations the case is treated as alive.

The second method (denoted Det2) has the same definition with the only exception that component first name is not taken into consideration.

The third method (denoted Prob) is a method developed by the Cancer Registry of Tyrol [11]. This method is based on probabilistic record linkage theory [14]. Using the components last name, birth surname, first name, date of birth, sex and municipality code or zip code, a probability of identity

is computed for every pair of persons (denoted p-val), also taking into account phonetic translations and documentation and typing errors. The p-value is defined as the weighted sum of the probability of individual components, where the weight for a component is defined by the logarithm of the probability that the component is equal for identical persons, divided by the probability that the component is equal for nonidentical persons. Detailed formulas can be found in [11]. If p-val is greater than 0.95, we assume without further checks that the components describe the same person; for a p-val smaller than 0.75 we assume, again without further checks, that the components describe different persons. A p-val between 0.75 and 0.95 calls for a decision on a caseby-case basis. In general, this means further information is needed to describe the persons more precisely. No blocking techniques were implemented, because the computing times were short enough for our typical problems.

The analysis was based on the incidence dataset of the Cancer Registry of Tyrol. For organizational reasons, we decided to use incidence data from the years 1992 to 1996. Incidence data were published in Cancer Incidence in Five Continents [16], and relative survival rates were analyzed in the framework of the EUROCARE study [1]. Incidence rates are in the range observed in Central European countries, except for prostate cancer. Since the introduction of PSA testing beginning around 1990, prostate cancer incidence rates are among the highest in Europe. With few exceptions relative survival rates are among the best survival rates seen in Europe [1]. Survival rates are influenced by several parameters, of which therapy, diagnostic procedures and availability of screening programs are important. Also, structural conditions, for

Definition of record linkage methods analyzed

k-	Det1	Det2	Prob
	Equality of last name first name date of birth sex	Equality of last name date of birth sex	Probability based on last name birth surname first name date of birth sex municipality code

Oberaigner

			м	en		Women						
Year of Diagnosis	N	Prob	Det 1		Det2		N	Prob	Det1		De	et2
		Rate	Rate	Diff ¹⁾	Rate	Diff ¹⁾		Rate	Rate	Diff ¹⁾	Rate	Diff ¹⁾
1992	1317	66.0	57.0	13.6%	62.8	4.8%	1332	55.6	40.9	26.4%	51.4	7.4%
1993	1386	62.0	52.9	14.8%	58.4	5.8%	1267	55.3	41.3	25.4%	52.6	4.9%
1994	1413	58.5	51.5	12.0%	55.2	5.7%	1212	54.7	39.3	28.2%	51.7	5.4%
1995	1379	54.0	46.8	13.3%	51.0	5.5%	1175	47.3	35.7	24.5%	44.3	6.5%
1996	1399	49.7	43.7	11.9%	46.7	5.9%	1266	46.9	34.4	26.8%	43.7	6.9%
¹⁾ relative diffe	erence based o	on Prob										

 Table 2
 Mortality rates for incident cancer cases in Tyrol, classified by sex and year of diagnosis

example volume size and academic status of hospitals, influence survival rates at the population level.

Results

Table 2 shows mortality rates classified by sex and year of diagnosis. For men, mortality rates range from 49.7% to 66.0% (Prob), 43.7% to 57.0% (Det1) and 46.7% to 62.8% (Det2) and for women from 46.9% to 55.6% (Prob), 34.4% to 41.3% (Det1) and 43.7% to 52.6% (Det2). The closure date for this study was December 31, 2002. Patients with earlier year of diagnosis are subject to longer observation periods, and thus the simple mortality rate is higher for the earlier years of diagnosis. Mortality rates according to Prob are higher than the rates obtained with a deterministic method, because Prob

assigns a date of death even if single components differ. Mortality rates according to method Det2 are higher than for Det1, because Det2 has a less strict rule for adding date of death than does Det1.

We also compute relative differences when comparing Det1 or Det2 to Prob. For Det1 relative differences range from 11.9% to 14.8% for men and from 24.5% to 28.2% for women, and for Det2 from 4.8% to 5.9% for men and from 4.9% to 7.4% for women.

If we assume that Prob represents the true survival rate, on average Det1 would underestimate the mortality rate by approx. 13% (men) and approx. 26% (women), while Det2 would underestimate the mortality rate for both sexes by approx. 5%-6%.

Results for relative five-year survival rates are shown in Table 3. For men, relative five-year survival rates range from 45.7% to 52.0% (Prob), from 53.0% to 58.1% (Det1) and from 48.5% to 54.9% (Det2). For

women, relative five-year survival rates range from 52.0% to 56.4% (Prob), from 65.8% to 67.3% (Det1) and from 54.5% to 59.3% (Det2). Relative differences for Det1 as compared to Prob range from 11.4% to 16.3% for men and from 19.3% to 26.4% for women, while relative differences for Det2 as compared to Prob range from 5.1% to 7.0% for men and from 4.4% to 6.1% for women. If we assume that Prob represents the true survival rate, Det1 would overestimate the relative five-year survival rate by approx. 14% (men) and 22% (women) and Det2 by 5-6% for both sexes.

Discussion

It was the aim of this study to assess the size of error in mortality rate and relative survival rate caused by using deterministic

Table 3 Relative five-year survival rate for incident cancer cases in Tyrol, classified by sex and year of diagnosis

			М	en			Women						
Year of Diagnosis	N	Prob		Det1		t1 Det2		Prob	Det 1		Det2		
		Rate	Rate	Diff ¹⁾	Rate	Diff ¹⁾		Rate	Rate	Diff ¹⁾	Rate	Diff ¹⁾	
1992	1317	45.7	53.0	15.9%	48.5	6.0%	1332	54.5	66.2	21.4%	57.8	6.1%	
1993	1386	46.9	54.6	16.3%	50.2	7.0%	1267	54.1	66.4	22.8%	56.5	4.4%	
1994	1413	49.2	55.6	12.9%	52.4	6.5%	1212	52.0	65.8	26.4%	54.5	4.8%	
1995	1379	51.6	58.1	12.6%	54.2	5.1%	1175	56.4	67.3	19.3%	59.3	5.0%	
1996	1399	52.0	57.9	11.4%	54.9	5.7%	1266	55.0	67.1	22.0%	58.3	5.9%	

¹⁾ relative difference based on Prol

423

record linkage procedures. Correctness of record linkage is indeed a relevant problem, because many studies in various medical fields add status of patient or date of death passively by linking patient data with mortality files, see for example [19–22]. In this situation, errors in record linkage bias survival and mortality rates.

Prob denotes a probabilistic record linkage procedure developed by the Cancer Registry of Tyrol. Possible errors occurring when using Prob were systematically checked during implementation of the method, resulting in negligible homonym and synonym rates [11]. Consequently, it is justified to measure error rates by comparing results from deterministic methods with those given by Prob.

Det1 and Det2 are record linkage methods routinely used in various registries (personal communication). It is well known that deterministic record linkage procedures cause errors, namely false-positive and false-negative results, which raises the question why deterministic methods are still in use. The main reason for applying deterministic record linkage is that this method needs no further interaction. In contrast, when using probabilistic record linkage unclear cases must be decided on a case-bycase basis, which calls for additional time or financial resources. Also, methods for deterministic record linkage can be programmed with very simple tools, which may contribute to their rather wide use.

For the cancer registry dataset, results obtained with deterministic methods were reconstructed based on the criteria defining the deterministic methods. The error rates for Det1 and Det2 are plausible, because the deterministic methods use strict rules for equality of components like name and date of birth, and deterministic methods cannot deal with documentation and typing errors that occur in practical situations.

Likewise, it is plausible that Det1 generates much higher error rates than does Det2, because Det2 does not require equality of first names. This is the reason why error rates between men and women are of comparable size for Det2, but Det1 shows huge gender differences. When looking at these results for the first time, we hypothesized that this discrepancy is due to problems with first name. Indeed, a detailed look at pairs of persons linked by Det2 and not by Det1 showed that in the majority of cases the first names differed.

Our results depend on the specific deterministic procedures used for comparison. It should be noted that both methods analyzed are frequently used in practice. The literature shows some rather sophisticated procedures for deterministic record linkage used mostly in English-speaking countries [23], and some reports [24-26] conclude that deterministic record linkage procedures can provide results with sufficient accuracy. It is self-evident that all such comparisons strongly depend on the method used for deterministic record linkage and to some extent on the language the names originate in. Also, it should be remembered that our results depend on first names, and, in fact, the structure of first names in the German language is more complicated that in English. Also, some degree of the problem could be specific to Austria, because when building up the mortality data file, the first name is taken from official documents, whereas most hospital records, which provide the basis for medical data, use the person's habitual first name (in Austria's rural regions it is customary for people, especially older people, to have and use a nickname in place of their official first name; official mortality data use the official first name).

One might ask whether these results also hold for other languages. The discriminating power of first names and last names depends to some extent on the language. The majority of last names are German-language names, so our results hold only for German-language names or German-speaking countries. Since the structure of Ger-

Table 4Out-migrant rate for the population of Tyrolfor the year 1999, classified by age and sex

Age group	Men	Women
0—19	0.91%	1.16%
20—44	2.67%	2.17%
45—64	1.17%	0.76%
65—84	0.45%	0.36%
≤85	0.33%	0.28%
Total	1.63%	1.33%

man-language names is comparable in all German-speaking countries, we can assume a similar magnitude of error for all datasets with mainly German-language names when using these deterministic linkage methods. Also, our probabilistic method implements transformations specific for the German language.

We would like to address a further Austria-specific aspect, namely immigrants who came to Austria in recent decades mainly from the former Yugoslavia and Turkey. This group now accounts for about 10% of the population of Tyrol. Up to now, the specific age structure of the immigrant group means that cancer rates are still low for immigrants. In the future, however, we expect more and more cancer patients with non-German-language names, which will cause severe problems for two reasons. First, we expect a different discriminating power, especially for Turkish names, and second, a large part of the secretarial staff in hospitals and registries has no knowledge of the structure of non-German languages and names, which means that typing and "listening" errors are more likely to occur.

Finally, we would like to stress that conclusions depend very much on the overall targets of record linkage projects. For example, for administrative purposes [27] or when adding medical records to existing cancer registry data [28], a sensitivity of about 95% or less is usually accepted. However, for epidemiological studies like survival studies, synonym rates can have clear consequences on study outcome, as shown by our analysis.

The Introduction addressed the outmigrant rate as the second error source for the passive follow-up procedure. The outmigrant rate for Tyrol is assessed periodically by the Department of Statistics for Tyrol. Table 4 shows the results for the year 1999, classified by age and sex. We see that, in general, out-migrant rates for the population of Tyrol were smaller than 2% for both sexes. In addition, in the age classes relevant for cancer mortality, namely below 19 and above 65, out-migrant rates are 1% and lower. We thus conclude that outmigrants play a minor role in cancer survival rates in Tyrol. However, these figures are specific for Tyrol and possibly do not hold for other Austrian states.

Conclusion

We analyzed the effect of two deterministic methods of record linkage on mortality rates and relative survival rates. The differences as compared to a well-established probabilistic record linkage method were considerable. Overestimation of survival rate was up to 15% for men and 25% for women when using Det1 and 5-6% when using Det2, each as compared to Prob. This analysis clearly shows that in order to calculate valid survival results, a probabilistic method of record linkage must be applied. Outmigrants in Tyrol have minimal influence on mortality rates and survival rates and can thus be neglected.

References

- Coleman MP, Gatta G, Verdecchia A, Esteve J, Sant M, Storm H, Allemani C, Ciccolallo L, Santaquilani M, Berrino F. EUROCARE-3 summary: cancer survival in Europe at the end of the 20th century. Ann Oncol 2003;V128-V149.
- Coebergh JW, Sant M, Berrino F, Verdecchia A. Survival of Adult Cancer Patients in Europe Diagnosed from 1978–1989: The Eurocare II Study. Eur J Cancer 1998.
- Brenner H, Hakulinen T. Very-long-term survival rates of patients with cancer. J Clin Oncol 2002; 4405-9.
- SEER Program version (2001): SEER*Stat 4.0.9, Data Incidence – SEER 9, Registries Public Use (1973-1998). 2000.
- Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas B. Cancer Incidence in Five Continents. Volume VIII. Lyon: 2002.

- Brenner H, Schmidtmann I. Determinants of homonym and synonym rates of record linkage in disease registration. Methods Inf Med 1996; 35: 19-24.
- Newman TB, Brown AN. Use of commercial record linkage software and vital statistics to identify patient deaths. J Am Med Inform Assoc 1997; 233-7.
- Quantin C, Bouzelat H, Allaert FA, Benhamiche AM, Faivre J, Dusserre L. How to ensure data security of an epidemiological follow-up: quality assessment of an anonymous record linkage procedure. Int J Med Inf 1998; 117-22.
- Gomatam S, Carter R, Ariet M, Mitchell G. An empirical comparison of record linkage procedures. Stat Med 2002; 1485-96.
- van den Brandt P, Schouten LJ. Development of a Record Linkage Protocol for Use in the Dutch Cancer Registry for Epidemiological Research. Int J Epid 1990;553–8.
- Oberaigner W, Stühlinger W. Record Linkage in the Cancer Registry of Tyrol, Austria. Methods Inf Med 2005; 44 (5): 626-30.
- Verykios VS, Moustakides GV, Elfeky MG. A Bayesian decision model for cost optimal record matching. Vldb Journal 2003; 28-40.
- Dal Maso L, Braga C, Franceschi S. Methodology used for "Software for automated linkage in Italy" (SALI)(1). Journal of Biomedical Informatics 2001; 387-95.
- Jaro MA. Probabilistic Linkage of Large Public Health Data Files. Stat Med 1995; 14: 491-8.
- Oberaigner W, Mühlböck H, Harrasser L. Tumorregister Tirol – Bericht für die Diagnosejahre 1997/98. Innsbruck: 2003.
- Parkin DM, Whelan SL, Ferlay J, Raymond L, Yuen J. Cancer Incidence in Five Continents. Volume VII. Lyon: IARC, 1997.
- Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comp Progr in Biomed 1985; 197-207.
- Voutilainen ET, Dickman PW, Hakulinen T. SURV 3 – Relative Survival Analysis. Program Version 3.00b1. Helsinki: Finish Cancer Registry, 2001.
- Muhlberger V, Greil V, Stuhlinger W, Pachinger O. Mortality analysis 2 years after invasive cardiologic interventions in Innsbruck in the year of 1998 by record linkage of the "Statistik Österreich". Herz 2001; 495-500.

- Bopp M, Minder CE. Mortality by education in German speaking Switzerland, 1990-1997: results from the Swiss National Cohort. International Journal of Epidemiology 2003; 346-54.
- 21. Garcia M, Schiaffino A, Fernandez E, Marti M, Salto E, Perez G, Peris M, Borrell C, Nieto FJ, Borras JM. The Cornella Health Interview Survey Follow-Up (CHIS. FU) study: design, methods, and response rate. Bmc Public Health 2003; 12.
- Morgan CL, Kerr MP. Epilepsy and mortality: A record linkage study in a UK population. Epilepsia 2002; 1251-5.
- Grannis SJ, Overhage JM, Hui S, McDonald CJ. Analysis of a Probabilistic Record Linkage Technique without Human Review. Proc AMIA Symp 2003; 259-63.
- Adams MM, Wilson HG, Casto DL, Berg CJ, McDermott JM, Gaudino JA, McCarthy BJ. Constructing reproductive histories by linking vital records. Am J Epidemiol 1997; 339-48.
- Alsop JC, Langley JD. Determining first admissions in a hospital discharge file via record linkage. Methods Inf Med 1998; 37: 32-7.
- 26. Liu S. Development of record linkage of hospital discharge data for the study of neonatal readmission. Chronic Dis Can 1999; 77-81.
- Kendrick SW, Douglas MM, Gardner D, Hucker D. Best-link matching of Scottish health data sets. Methods Inf Med 1998; 37: 64-8.
- Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare-tumor registry database. Med Care 1993; 732-48.

Correspondence to:

Willi Oberaigner Cancer Registry of Tyrol Department of Clinical Epidemiology of the Tyrolian State Hospitals Ltd. Anichstrasse 35 Innsbruck Austria E-mail: willi.oberaigner@iet.at Eur J Haematol 2001: 66: 43–49 Printed in UK. All rights reserved

Allogeneic or autologous stem cell transplantation (SCT) for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma: a single-centre experience

Nachbaur D, Oberaigner W, Fritsch E, Nussbaumer W, Gastl G. Allogeneic or autologous stem cell transplantation (SCT) for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma: a single-centre experience.

Eur J Haematol 2001: 66: 43–49. © Munksgaard 2001.

Abstract: Purpose of the study: The aim of the study was to evaluate which patient might benefit most from allogeneic stem cell transplantation (SCT) in the treatment of relapsed and/or refractory lymphoma. Patients and methods: Thirty-eight consecutive lymphoma patients receiving either autologous (n = 24) or allogeneic (n = 14) stem cell grafts at our institution from 1986 to 1998 were retrospectively analysed regarding overall survival (OS), disease-free survival (DFS), transplantrelated mortality (TRM), and relapse incidence (RI). Uni- and multivariate analyses were performed to identify patient characteristics predictive for outcome after SCT. Results: The probabilities of OS, DFS, TRM, and relapse were 57%, 51%, 29%, and 30% following autologous and 43%, 43%, 29%, and 38% following allogeneic SCT. Disease status (sensitive versus refractory) and the time interval between diagnosis and SCT were the most powerful predictive parameters for OS and TRM, whereas elevated serum LDH levels were significant in determining relapse. Conclusions: In patients with elevated serum LDH levels and bone marrow involvement at the time of transplantation allogeneic was superior to autologous SCT and resulted in better outcome due to a lower relapse incidence strongly suggesting the existence of a graftversus-lymphoma effect.

High-dose chemo/radiotherapy with autologous stem cell rescue offers a reliable chance of longterm disease-free survival for patients with relapsing or refractory Hodgkin's disease (HD) or non-Hodgkin's lymphoma (NHL) who otherwise have a poor prognosis (1–4). During the past decade several studies have demonstrated the relative effectiveness of this approach and its advantages over conventional-dose salvage regimens in selected patients (5–9). In contrast, allogeneic SCT is used only sporadically because of its higher TRM and the risk of developing acute and/or chronic graftversus-host disease (GVHD) (10–17). Potential graft-versus-lymphoma (GVL) effects associated

David Nachbaur¹, Wilhelm Oberaigner², Esther Fritsch³, Walter Nussbaumer⁴, Günther Gastl¹

¹BMT Unit, Division of Hematology and Oncology, Department of Internal Medicine, University Hospital, Innsbruck, ²Cancer Registry of Tyrol, Innsbruck, ³Department of Radiation of Therapy and ⁴Department of Transfusion Medicine, Universitiy Hospital, Innsbruck, Austria

Key words: stem cell transplantation (SCT); graftversus-lymphoma (GVL) effect; Hodgkin's disease (HD); non-Hodgkin's lymphoma (NHL)

Correspondence: David Nachbaur, MD, Division of Hematology & Oncology, Universitiy Hospital of Internal Medicine, Anichstrasse 35, A-6020 Innsbruck, Austria Tel: +43 512 504 3384 Fax: +43 512 504 3338 e-mail: david.nachbaur@uibk.ac.at

Accepted for publication 13 September 2000

with allogeneic SCT are supposed in some series but warrant confirmation by prospective, randomized trials (18–21). To define better the role of allogeneic SCT in the management of poor-risk lymphoma we here report on our experience obtained in 38 consecutive patients transplanted at our institution from 1986 to 1998.

Patients and methods

Patients

Between 1986 and 1998, 38 patients (HD, n=14; NHL, n=24) received either autologous (n=24) or

Nachbaur et al.

allogeneic, HLA-identical (n = 14) stem cell grafts. Preference was generally given to an allograft when patients were candidates for both types of graft. All patients were classified as having poor prognosis defined as either primary refractory or achieving only partial recovery (PR) after initial standard treatment, or as disease relapsing within one year from diagnosis after achieving complete remission (CR) following initial chemotherapy, or as patients with second or subsequent relapse. Patients with NHL were classified according to the updated Kiel classification and patients with HD according to the Rye classification (22, 23). Patient characteristics are detailed in Tables 1 and 2.

Preparative regimens and stem cell infusion

For patients with previous dose-limiting radiation therapy the standard preparative regimen was highdose cyclophosphamide, carmustine, and etoposide (CBV, consisting of 100 mg/kg CY, 15 mg/kg BCNU, and 60 mg/kg etoposide) followed by either autologous (n=15) or allogeneic (n=2) SC infusion (24). Patients without prior dose-limiting radiotherapy received a combination of high-dose CY (100 mg/kg), etoposide (60 mg/kg), and fractionated TBI (fTBI, 12 Gy, given in six fractions over three consecutive days) followed by either allogeneic (n=6) or autologous (n=9) SCT. Standard highdose CY(120 mg/kg) plus fTBI (12 Gy) was given prior to five allogeneic and one autologous SCT.

Table 1. Patient characteristics I

One patient received a combination of BU (8 mg/kg), CY (100 mg/kg), and single-dose TBI (10 Gy) followed by allogeneic stem cell infusion.

Allografted patients received a median of 3.41 (range 2.1-6.2) × 10⁸/kg BM nucleated cells from their HLA-identical sibling donors and autografted patients were reinfused with a median of 0.37 (range 0.19-0.47) × 10⁸/kg BM plus a median of 2.52 (range 0.18-4.34) × 10⁸/kg PBMC obtained by means of steady-state leukapheresis (n=9). Since 1995 patients undergoing autologous SCT were reinfused with either $\ge 4.0 \times 10^8$ /kg unmodified PBMC or $\ge 2.0 \times 10^6$ /kg immunoselected (Cell-Pro) CD34 + hematopoietic stem cells mobilised with either high-dose CY (4–7 g/m², n=12) or DHAP (n=3) plus G-CSF according to recently published standard procedures (25).

Relapse prophylaxis

Patients with radiological (CT scan) evidence of localised disease at the time of SCT underwent involved-field irradiation (20–30 Gy) starting as soon as possible within the first three months after hematopoietic regeneration.

Supportive care including graft-versus-host disease prophylaxis

Patients receiving autologous SCT were treated under strict reverse isolation without laminar airflow. Patients receiving allografts were nursed in

	Allogeneic SCT ($n = 14$)	Autologous SCT ($n=24$)
Median age (yr, range)	35 (16–48)	41 (16–55)
Female:male ratio	6:8	7:17
Primary diagnosis		
Hodgkin's disease	5 (36%)	9 (37%)
Non-Hodgkin's lymphoma	9 (64%)	15 (63%)
Low-grade	2 (22%)	8 (53%)
High-grade	7 (78%)	7 (47%)
Median time from diagnosis to SCT (months, range)	17 (3–62)	26 (6–117)
Disease status at SCT		
Sensitive	6 (43%)	15 (62%)
Refractory/progressive	8 (57%)	9 (38%)
Number previous lines of treatment		
1–2	8 (57%)	9 (38%)
≥3	6 (43%)	15 (62%)
Time from diagnosis to SCT		
≼3 yr	11 (79%)	15 (62%)
>3 yr	3 (21%)	9 (38%)
Serum LDH levels at SCT		
≤240 IU/L	11 (79%)	19 (79%)
>240 IU/L	3 (21%)	5 (21%)
Bone marrow involvement at SCT		
No	10 (71%)	17 (71%)
Yes	4 (29%)	7 (29%)
Age at SCT		
≪40 yr	11 (79%)	11 (46%)
> 40 yr	3 (21%)	13 (54%)

44

Table 2. Patient characteristics II

	Allogeneic SCT ($n = 14$)	Autologous SCT ($n = 24$)
Median observation time following SCT (months, range)	18 (0–132)	25 (0–154)
Median follow-up for patients alive (months, range)	106 (39–132)	45 (11–154)
Conditioning regimen		
CBV	2	15
CY/fTBI	5	1
BUCY/TBI	1	0
CY/VP-16/fTBI	6	8
TBI-containing	12 (86%)	9 (38%)
G-CSF after Tx	3 (21%)	22 (92%)
Median days until leukocytes > 1.0 G/L (range)	16 (12–29)	15 (9–37)
GVHD prophylaxis		
CsA	11	_
CsA/MTX	3	-

laminar airflow rooms from the beginning of the conditioning regimen until hematopoietic regeneration. No prophylactic systemic antibiotics were administered. All patients underwent a non-absorbable oral gut decontamination with vancomycin, gentamycin, and nystatin. Pneumocystis carinii prophylaxis was performed with trimethoprimsulfamethoxazole given in a 10-d course before transplantation and after the take. Cytomegalovirus (CMV) pneumonia prophylaxis consisted of infusions of CMV hyperimmunoglobulin (Cytotect, Cutter, 1 ml/kg) every other week until day +100. Irradiated (25 Gy), leukocyte-depleted and CMVnegative red cells and platelet transfusions from single donors were administered when hemoglobin levels were 7.0 g/dL or less and platelets were 20 G/L or less. To accelerate hematopoietic regeneration G-CSF (5 μ g/kg/d) was given to 3/14 (21%) allografted and to 22/24 (92%) autografted patients starting on the day after SC infusion.

GVHD prophylaxis consisted of cyclosporine A (CsA) alone (n=11) or in combination with shortcourse methotrexate (MTX) according to the Seattle protocol (n=3). Grading and treatment of acute and chronic GVHD was performed according to the standard Seattle criteria and protocols.

Statistics

Survival analyses were performed according to the method of Kaplan and Meier (26). Overall survival (OS) was calculated from the date of SCT to the date of death from any cause or day of last follow-up. Disease/progression free survival (DFS) was calculated from the date of SCT to the date of documented disease relapse/progression. Transplant-related mortality was defined as the probability of death without relapse or disease progression. For two patients receiving a second graft (one autologous and one allogeneic) because of disease progression/relapse after autologous SCT the unit studied was the patient and the censored data correspond to the date of the last contact for each patient according to the recently published European Group for Blood and Marrow Transplantation (EBMT) statistical guidelines (27).

Univariate analysis of the following parameters was performed using the log-rank test and SPSS software to identify patient characteristics predictive for outcome after SCT: diagnosis (HD vs. NHL), stem cell source (allogeneic vs. autologous), disease status at the time of SCT (SD vs. RD/PD), conditioning regimen (TBI-containing vs. chemotherapy alone), number of previous lines of treatement (1–2 vs. \geq 3), BM involvement at the time of SCT (yes vs. no), and serum LDH levels at the time of SCT (≤ 240 IU/L vs. > 240 IU/L). For the variable "age" the median age (40 yr) of the whole study population was chosen as cut-off. For the variable "time interval between diagnosis and SCT" an interval of 3 yr was chosen because the greatest difference in overall survival was observed between patients receiving a transplant within or beyond this time frame.

Multivariate analysis was performed using Cox's proportional hazards model. The factors examined were the same as those included in the univariate analysis.

Results

Hematological engraftment, overall survival (OS), and disease-free survival (DFS)

All but four patients dying too early because of regimen-related toxicity and/or infection engrafted (defined as the first day with a persistent leukocyte count > 1.0 G/L) after a median of 16 d (range 12–29) following allogeneic and a median of 15 d (range 9–37) following autologous SCT.



Fig. 1. Probabilities of overall survival (OS), relapse/progression and transplant-related mortality following autologous (A) or allogeneic (B) stem cell transplantation for poor-risk lymphoma.

The probability of OS $\pm 95\%$ confidence interval (CI) for all patients was $51\pm8\%$ with an actuarial survival at 3 yr of $57\pm11\%$ vs. $43\pm13\%$ following autologous vs. allogeneic SCT (Fig. 1) and a disease/progression-free survival of $51\pm10\%$ vs. $43\pm13\%$.

The most powerful predictive parameters for survival following SCT were disease status and time interval from diagnosis to SCT ($p \le 0.05$, log-rank test, Fig. 2 and Table 3). Younger age, non-TBI-containing conditioning, normal serum LDH levels, and absence of bone marrow involvement were also associated with improved survival but did not reach statistical significance (Table 3).

Allogeneic SCT was superior to autologous SCT only in patients with elevated serum LDH levels >240 IU/L (overall survival $33 \pm 27\%$ vs. $20 \pm 18\%$, difference not significant) and BM involvement at the time of SCT (overall survival $75 \pm 22\%$ vs. 0%, p=0.091) mainly due to a lower RI.

For HD patients, OS following allogeneic vs. autologous SCT was $40 \pm 22\%$ vs. $67 \pm 15\%$. For patients with NHL, survival following allogeneic and autologous SCT was $44 \pm 17\%$ and $47 \pm 15\%$, respectively. Overall survival following autologous

SCT was $42\pm21\%$ for patients with low-grade lymphoma vs. 57+16% for high-grade lymphoma.

Survival for patients with sensitive disease receiving allografts and autografts was $67 \pm 21\%$ and $78 \pm 18\%$, respectively, whereas survival in patients with refractory disease was only $25 \pm 13\%$ following allo- and $22 \pm 12\%$ following autotransplantation (differences not significant). Also for all other variables listed in Table 3 survival was not different between auto- and allotransplantation.

Relapse incidence (RI) and graft-versus-host disease (GVHD)

The probability of relapse/progression ($\pm 95\%$ CI) at 3 yr for the entire study cohort was $34 \pm 9\%$, with a RI of $38 \pm 15\%$ vs. $31 \pm 11\%$ following allogeneic vs. autologous SCT (Fig. 1).

The only factor significantly determining RI was serum LDH level with a significantly higher RI in patients with elevated serum LDH levels ($71 \pm 23\%$ vs. $25 \pm 9\%$, p = 0.048, Table 3). Also patients with refractory disease had a higher RI ($51 \pm 17\%$ vs. $25 \pm 10\%$), but this did not reach statistical significance (Fig. 2, Table 3).



Fig. 2. Probabilities of overall survival (OS), relapse/progression and transplant-related mortality following stem cell transplantation for poor-risk lymphoma with sensitive (A) or refractory/progressive disease (B).

46

Stem cell transplantation for malignant lymphoma

Table 3. Uni-(log-rank test) and multivariate (Cox model) analysis of prognostic variables for overall survival, relapse incidence, and transplant-related mortality following SCT for poor-risk lymphoma (*n*=38)

		Overall survival (OS)			Relapse incidence (RI)				Transplant-related mortality (TRM)				
Prognostic variables	Patients at risk	(% ± SD)	Log-rank	Cox model	Odd's ratio	$(\% \pm SD)$	Log-rank	Cox model	Odd's ratio	(% ± SD)	Log-rank	Cox model	Odd's ratio
Type of SCT													
Autologous	24	56 <u>+</u> 11				31 ± 11				29 ± 11			
Allogeneic	14	43 <u>+</u> 13	0.42	0.4814	1.70	38 ± 15	0.7652	0.7392	1.40	29 <u>+</u> 12	0.8138	0.1882	4.70
Age at SCT													
≪40 yr	22	59 ± 10				34 ± 11				23 ± 9			
>40 yr	16	36 ± 14	0.361	0.8742	0.90	34 ± 14	0.8779	0.9848	1.00	44 ± 18	0.5154	0.6121	0.70
Diagnosis													
Hodgkin's disease	14	57 <u>+</u> 13				39 ± 15				29 <u>+</u> 12			
Non-Hodgkin's lymphoma	24	48 <u>+</u> 11	0.6499	0.6351	1.40	31 ± 10	0.7989	0.7858	1.30	30 ± 11	0.8649	0.2611	0.20
Conditioning													
Chemotherapy	17	64 ± 11				29 ± 12				18 ± 9			
TBI-containing	21	42 ± 11	0.2439	0.4392	1.60	38 ± 12	0.639	0.9757	1.00	37 <u>+</u> 12	0.3191	0.3043	2.60
Disease status at SCT													
Sensitive	21	75 ± 10				25 ± 10				12 ± 8			
Refractory/progressive	17	24 ± 10	0.0003	0.0099	7.80	51 ± 17	0.189	0.4205	2.30	49±12	0.0033	0.034	19.00
Previous lines of treatment													
1–2	17	50 ± 13				33 ± 12				30±14			
≥3	21	51 ± 11	0.8839	0.424	0.50	34 ± 13	0.8014	0.738	0.70	29±10	0.687	0.08	0.10
Time from diagnosis to SCT													
≼3 yr	26	60 ± 10				30 ± 9				19 ± 8			
>3 yr	12	33 ± 14	0.0438	0.0724	3.90	49 ± 20	0.8014	0.925	1.10	50 ± 14	0.013	0.016	34.80
Serum LDH levels at SCT													
≪240 IU/L	30	58 ± 10				25 ± 9				26±9			
>240 IU/L	8	25 ± 15	0.0714	0.8949	0.90	71 ± 23	0.048	0.1896	3.60	43 <u>+</u> 19	0.456	0.3029	0.50
BM involvement at SCT													
No	27	59 ± 10				36 ± 10				19±8			
Yes	11	34 ± 15	0.2108	0.6096	0.70	27 <u>+</u> 17	0.5689	0.1923	0.20	52 <u>+</u> 18	0.0966	0.4976	2.60

There was a trend for a lower RI following allotransplantation in patients with elevated serum LDH levels $(33 \pm 15\% \text{ vs. } 73 \pm 22\% \text{ for autologous})$ transplantation, difference not significant) and in patients with BM involvement (0% vs. $47 \pm 15\%$ for autologous transplantation, difference not significant). We also observed a trend for a lower RI in HD patients following allotransplantation (33 + 27%) vs. $43 \pm 19\%$ following autologous SCT), whereas for NHL patients the incidence of relapse following allogeneic SCT was even higher than following autologous SCT $(43 \pm 19\% \text{ vs. } 24 \pm 12\%)$, but there were more high-grade lymphomas in the allogroup (78% vs. 47%). For none of the variables listed in Table 3 was relapse incidence significantly different between allo- and autotransplantation.

Acute GVHD (aGVHD) grades I–II developed in 7/14 allografted patients after a median of 22 (1– 27) d. None of the patients developed aGVHD grades III–IV or died due to GVHD, and only one patient had laboratory signs of limited chronic GVHD of the liver requiring prolonged administration of CsA. Interestingly, 4/7 allografted patients without aGVHD died of relapse/progression, whereas none of the allografted patients with evidence of aGVHD relapsed within the study period. Transplant-related mortality (TRM) and causes of death

Eighteen patients died within the observation period. The causes of death in the allogroup were relapse/progression (n=4), viral interstitial pneumonitis (n=3), and septical multiorgan failure (n=1), whereas in the autogroup three patients died of relapse/progression, six patients due to infectious complications and one patient died of cardiac toxicity.

Sensitivity of disease and disease duration were the only variables associated with a significantly lower TRM following SCT (Table 3).

As shown in Fig. 1, TRM was similar following allogeneic and autologous SCT. There was a trend for a higher TRM in patients with HD receiving allografts ($40 \pm 22\%$ vs. $22 \pm 14\%$ following autologous SCT, difference not significant).

For all other variables listed in Table 3 no significant differences in TRM between allogeneic and autologous SCT were observed.

Discussion

In accordance with other reports, sensitive disease was the most powerful predictive parameter for survival after SCT for poor-risk lymphoma (2–5, 14, 15, 21). Additionally, our data demonstrate that a shorter disease duration was associated with favourable outcome, indicating that high-dose chemotherapy with stem-cell support should be offered early to patients with relapsing and/or refractory lymphoma. As suspected, for both variables the better outcome resulted from a significantly lower transplant-related mortality not only following autologous but also allogeneic SCT.

To date there is no answer to the question of which lymphoma patient should receive and would benefit most from allogeneic SCT if both treatment modalities are available. Although there is some evidence supporting the concept of a graft-versuslymphoma (GVL) effect at least in selected subgroups, so far no prospective study has proved any survival advantage of allogeneic over autologous SCT because of a higher procedure-related mortality (10, 18–21).

Although in our study differences between autoand allotransplantation with regard to any of the variables tested by uni- and multivariate analysis were statistically not significant because of the small patient numbers in the individual subgroups, the improved survival due to a lower relapse incidence following allotransplants in patients with elevated serum LDH levels and bone marrow involvement at the time of SCT supports the hypothesis of a GVL effect. Whether such an effect also exists in Hodgkin's disease as supposed from the lower RI found in HD patients receiving allografts remains to be shown. Irrespectively, the fact that none of the allografted patients developing clinical GVHD relapsed compared with 4/7 patients without GVHD strongly argues in favour of the existence of either tumor-specific (tumor antigen-directed) or unspecific (minor antigen-directed) antitumor effects mediated by the graft. Interestingly, in contrast to other published series we did not find a higher treatment-related mortality following allografting with the exception of HD patients, but this difference did not reach statistical significance (18, 20, 21). Nonetheless, treatmentrelated mortality remains high and accounted for 50% of all deaths in the allogroup and for 70% of all deaths in the autogroup, and was definitively only acceptable in patients with sensitive disease. Whether graft manipulation by either T-cell depletion or CD34⁺ selection with or without graded T-cell add-back can overcome these shortcomings at least in the allogeneic setting of SCT without any negative effect on relapse/progression or infectious complications remains to be seen (28).

Disease relapse/progression is the second major factor contributing to treatment failure following both types of SCT. All but one relapsing patient receiving a second autologous transplant died within weeks to months due to progressive disease, requiring alternative treatment strategies for those patients. Donor lymphocyte infusions (DLI) alone or in combination with chemotherapy might work after allogeneic SCT as shown by some recent reports but are far from being successful in all relapsing patients (29, 30). Recurrence after autologous SCT might be effectively salvaged by allogeneic or a second autologous SCT with acceptable toxicity, but long-term outcome remains poor (31, 32).

In conclusion, our results support the existence of graft-versus-lymphoma effect and identify а patients with elevated serum LDH levels and bone marrow involvement as those patients who might benefit most from allogeneic SCT when using standard procedures. The results, however, are preleminary and must be interpreted with caution due to the small patient number unless they have been confirmed by larger prospective, randomized studies. Recent advances especially the introduction of alternative, less toxic treatment modalities such as non-myeloablative conditioning will help to reduce toxicity and to clarify the mechanisms involved in graft-versus lymphoma effects and will inevitably increase the number of allogeneic transplants in the treatment of malignant lymphoma (33, 34).

References

- 1. ARMITAGE JO. Bone marrow transplantation in the treatment of patients with lymphoma. Blood 1989;**73**:1749–1758.
- PHILIP T, ARMITAGE JO, SPITZER G, et al. High-dose therapy and autologous bone marrow transplantation after failure of conventional chemotherapy in adults with intermediategrade or high-grade non-Hodgkin's lymphoma. N Engl J Med 1987;**316**:1493–1498.
- 3. GRIBBEN JG, GOLDSTONE AH, LINCH DC, *et al.* Effectiveness of high-dose combination chemotherapy and autologous bone marrow transplantation for patients with non-Hodgkin's lymphoma who are still responsive to conventional-dose chemotherapy. J Clin Oncol 1989;7:1621–1629.
- 4. APPELBAUM FR, SULLIVAN KM, BUCKNER CD, *et al.* Treatment of malignant lymphoma in 100 patients with chemotherapy, total body irradiation, and marrow transplantation. J. Clin Oncol 1987;**5**:1340–1347.
- PHILIP T, GUGLIELMI C, HAGENBEEK A, et al. Autologous bone marrow transplantation as compared with salvage chemotherapy in relapses of chemotherapy-sensitive non-Hodgkin's lymphoma. N Engl J Med 1995;333:1540–1545.
- TAKVORIAN T, CANELLOS GP, RITZ J, et al. Prolonged diseasefree survival after autologous bone marrow transplantation in patients with non-Hodgkin's lymphoma with a poor prognosis. N Engl J Med 1987;316:1499–1505.
- 7. PETERSON FB, APPELBAUM FR, HILL R, *et al.* Autologous marrow transplantation for malignant lymphoma: A report of 101 cases from Seattle. J Clin Oncol 1990;**8**:638–647.
- 8. NADEMANEE A, MOLINA A, O'DONENELL MR, et al. Results of high-dose therapy and autolgous bone marrow/stem cell

48

Stem cell transplantation for malignant lymphoma

transplantation during remission in poor-risk intermediateand high-grade lymphoma: International Index high and high-intermediate risk group. Blood 1997;**90**:3844–3852.

- HAIOUN C, LEPAGE E, GISSELBRECHT C, et al. Benefit of autologous bone marrow transplantation over sequential chemotherapy in poor-risk aggressive non-Hodgkin's lymphoma: Updated results of the prospective study LNH87–2. Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol 1997;15:1131–1137.
- VERDONCK LF. Allogeneic versus autologous bone marrow transplantation for refractory and recurrent low-grade non-Hodgkin's lymphoma: updated results of the Utrecht experience. Leuk Lymphoma 1999;34:129–136.
- 11. APPELBAUM FR, SULLIVAN KM, THOMAS ED, *et al.* Allogeneic marrow transplantation in the treatment of MOPP-resistant Hodgkin's disease. J Clin Oncol 1985;**3**:1490–1494.
- VAN BESIEN KW, KHOURI IF, GIRALT SA, et al. Allogeneic bone marrow transplantation for refractory and recurrent low-grade lymphoma: The case for aggressive management. J Clin Oncol 1995;13:1096–1102.
- SHEPHERD JD, BARNETT MJ, CONNORS JM, et al. Allogeneic bone marrow transplantation for poor-prognosis non-Hodgkin's lymphoma. Bone Marrow Transplant 1993;12:591–596.
- 14. STEIN RS, GREER JP, GOODMAN S, KALLIANPUR A, AHMED MS, WOLFF SN. High-dose therapy with autologous or allogeneic transplantation as salvage therapy for small cleaved cell lymphoma of follicular center cell origin. Bone Marrow Transplant 1999;23:227–233.
- MENDOZA E, TERRITO M, SCHILLER G, LILL M, KUNKEL L, WOLIN M. Allogeneic bone marrow transplantation for Hodgkin's and non-Hodgkin's lymphoma. Bone Marrow Transplant 1995;15:299–303.
- DANN EJ, DAUGHERTY CK, LARSON RA. Allogeneic bone marrow transplantation for relapsed and refractory Hodgkin's disease and non-Hodgkin's lymphoma. Bone Marrow Transplant 1997;20:369–374.
- 17. DEMIRER T, WEAVER CH, BUCKNER CD, *et al.* High-dose cyclophosphamide, carmustine, and etoposide followed by allogeneic bone marrow transplantation in patients with lymphoid malignancies who had received prior dose-limiting radiation therapy. J Clin Oncol 1995;**13**:596–602.
- RATANATHARATHORN V, UBERTI J, KARANES C, et al. Prospective comparative trial of autologous versus allogeneic marrow transplantation in patients with non-Hodgkin's lymphoma. Blood 1994;84:1050–1055.
- VERDONCK LF, DEKKER AW, LOKHORST HM, PETERSEN EJ, NIEUWENHUIS HK. Allogeneic versus autologous bone marrow transplantation for refractory and recurrent lowgrade non-Hodgkin's lymphoma. Blood 1997;90:4201– 4205.
- 20. CHOPRA R, GOLDSTONE AH, PEARCE R, et al. Autologous versus allogeneic bone marrow transplantation for non-Hodgkin's lymphoma: a case-controlled analysis of the

European Bone Marrow Transplant Group Registry Data. J Clin Oncol 1992;**10**:1690–1695.

- JONES RJ, AMBINDER RF, PIANATDOSI S, SANTOS GW. Evidence for a graft-versus-lymphoma effect associated with allogeneic bone marrow transplantation. Blood 1991;77:649–653.
- 22. STANSFELD AG, DIEBOLD J, NOEL H, *et al.* Updated Kiel classification for lymphomas. Lancet 1988;**i**:292–293.
- LUKES RJ, BUTLER JJ. The pathology and nomenclature of Hodgkin's disease. Cancer Res 1966;26:1063–1081.
- 24. WEAVER CH, APPELBAUM FR, PETERSEN FB, *et al.* High-dose cyclophosphamide, carmustine, and etoposide followed by autologous bone marrow transplantation in patients with lymphoid malignancies who have received dose-limiting radiation therapy. J Clin Oncol 1993;**11**:1329–1335.
- 25. NACHBAUR D, FINK FM, NUSSBAUMER W, et al. CD34+ selected autologous peripheral blood stem cell transplantation (PBSCT) in patients with poor-risk hematological malignancies and solid tumors. A single centre experience. Bone Marrow Transplant 1997;20:827–834.
- KAPLAN E, MEIER P. Non-parametric estimation for incomplete observations. J Am Stat Assoc 1958;53:457–481.
- 27. LABOPIN M. EBMT Statistical guidelines. EBMT News 1999;9:4–5.
- JUCKETT M, ROWLINGS P, HESSNER M, *et al.* T cell-depleted allogeneic bone marrow transplantation for high-risk non-Hodgkin's lymphoma: clinical and molecular follow-up. Bone Marrow Transplant 1998;21:893–899.
- BERNARD M, DAURIAC C, DRENOU B, et al. Long-term followup of allogeneic bone marrow transplantation in patients with poor prognosis non-Hodgkin's lymphoma. Bone Marrow Transplant 1999;23:329–333.
- 30. VAN BESIEN KW, DE LIMA M, GIRALT SA, *et al.* Management of lymphoma recurrence after allogeneic transplantation: the relevance of graft-versus-lymphoma effect. Bone Marrow Transplant 1997;**19**:977–982.
- DE LIMA M, VAN BESIEN KW, GIRALT SA, et al. Bone marrow transplantation after failure of autologous transplant for non-Hodgkin's lymphoma. Bone Marrow Transplant 1997;19:121–127.
- 32. VOSE JM, BIERMAN PJ, ANDERSON JR, *et al.* Progressive disease after high-dose therapy and autologoustransplantation for lymphoid malignanciy: clinical cours and patient follow-up. Blood 1992;**15**:2142–2148.
- 33. SLAVIN S, NAGLER A, NAPARSTECK E, et al. Nonmyeloablative stem cell transplantation and cell therapy as an alternative to conventional bone marrow transplantation with lethal cytoreduction for the treatment of malignant and nonmalignant hematologic diseases. Blood 1998;91:756–763.
- 34. KHOURI IF, KEATING M, KÖRBLING M, et al. Transplant-lite: induction of graft-versus-malignancy using fludarabinebased nonablative chemotherapy and allogeneic blood progenitor-cell transplantation as treatment for lymphoid malignancies. J Clin Oncol 1998;16:2817–2824.

Appendix 84

pendix 84

B2-2 Appendix Full Papers

Nephrol Dial Transplant (2003) 18: 1562–1567 DOI: 10.1093/ndt/gfg184

Original Article

Regional variability in the incidence of end-stage renal disease: an epidemiological approach

Florian Wimmer¹, Wilhelm Oberaigner², Reinhard Kramar³ and Gert Mayer¹

¹Division of Nephrology, ²Division of Clinical Epidemiology, Innsbruck University Hospital and ³Austrian Dialysis and Transplantation Registry, Krankenhaus der Barmherzigen Schwestern vom Heiligen Kreuz Wels, Austria

Abstract

Background. Regional variability in the incidence of end-stage renal disease (ESRD) in Austria is reported. Our aim was to investigate the reason for low rates in the state of Tyrol.

Methods. ESRD incidence data were obtained from the Austrian Dialysis and Transplantation Registry. Additional sources were two health interview surveys, the Hospital Discharge Registry, the Mortality Registry and the Drug Wholesale Registry.

Results. Between 1995 and 1999, 4811 new cases of ESRD were recorded; the state of Tyrol (T) had a mean annual, age-adjusted incidence of 97.9/1000000 population [95% confidence interval (CI) 86.9–109.1], a number significantly lower than that for the rest of Austria [(RA), 120.9 (95% CI 116.9–124.5); P < 0.001]. This was due mainly to a difference in the incidence of ESRD patients with type 2 diabetes mellitus [(DM-2) T = 12.2 (95% CI 8.2–16.2) *vs* RA = 28.9 (95% CI 27.2–30.6); P < 0.001]. When these patients were excluded, the difference in the overall ESRD incidence disappeared. When data from various registries were analysed for the prevalence of DM, a highly significant correlation was found between ESRD incidence and DM.

Conclusion. We conclude that the variability in the ESRD incidence in Austria is explained mainly by regional differences in DM-2. Data from similar studies might be useful for predictions concerning resource allocation for ESRD programmes in the future.

Keywords: Austria; diabetes; end-stage renal disease; epidemiology; incidence

Introduction

The first successful haemodialysis treatment of a patient with acute renal failure was performed by W. J. Kolff more than 50 years ago. Since then, medical advances have made this procedure a practical mode of therapy for chronic end-stage renal disease (ESRD). During the last two decades, a dramatic increase in the use of renal replacement therapy (RRT) has been observed worldwide [1,2], which has made the treatment of ESRD a significant public health burden in several developed countries [3]. The reason for this phenomenon is not entirely clear, but ageing of the population, improved overall survival and increase in numbers of patients at higher risk for the development of ESRD (e.g. diabetes and severe cardiovascular disease) may play a great role.

Even though the number of ESRD patients is increasing worldwide, a considerable regional variability has been reported [2-6], and a better understanding of this fact might help to develop preventive measures to reduce the burden of ESRD on a large scale and also facilitate advance allocation of health care resources. In contrast to the USA, where ethnic diversity greatly influences ESRD susceptibility [7], studies in more homogeneous populations might provide additional valuable information. The Austrian Dialysis and Transplantation Registry has reported areas such as the state of Tyrol with consistently low ESRD incidence [8]. Using this and other demographic databases, we attempted to test the following hypotheses, which could explain this observation:

- (i) Real lower incidence, regional variability in the incidence of ESRD comparable with a variability in diseases that lead to ESRD.
- (ii) Missed patients, e.g. by an insufficient diagnosis of renal disease.
- (iii) Reduced patient/physician acceptance into RRT.

Correspondence and offprint requests to: Florian Wimmer, MD, Division of Nephrology, Innsbruck University Hospital, Anichstrasse 35, 6020 Innsbruck, Austria. Email: fwimmer@ yahoo.com

^{© 2003} European Renal Association-European Dialysis and Transplant Association

- (iv) Lack of treatment facilities.
- (v) Early patient death from other causes before RRT becomes necessary.

Subjects and methods

Data sources

The primary data source for our study was the Austrian Dialysis and Transplantation Registry [8]. Since 1964, this Registry, which is run by the Austrian Society of Nephrology, has been collecting data provided by the 64 dialysis and transplantation centres in Austria on all patients with ESRD treated for at least 3 months. New patients (regardless of whether initial therapy was haemodialysis, peritoneal dialysis or transplantation) between January 1, 1995 and December 31, 1999 were identified. The home address at commencement of ESRD therapy, which was identified in 97.5% of all cases, was used to locate the state of residence of each patient [Vienna (Vie), Lower Austria (LA), Upper Austria (UA), Styria (ST), Burgenland (B), Carinthia (C), Salzburg (S), Tyrol (T) and Vorarlberg (V)]. The mean ESRD incidence per year was calculated using the average of the 5-year study period for Tyrol and for all of Austria without Tyrol [rest of Austria (RA)].

The Austrian Dialysis and Transplantation Registry also receives data on the renal disease which led to ESRD. These diagnoses are grouped into eight categories which were used for analysis: vascular nephropathy, nephropathy associated with type 1 or type 2 diabetes mellitus, glomerulonephritis, kidney disease of unknown origin, interstitial nephritis/ pyelonephritis, hereditary kidney disease and others (e.g. systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, etc.).

The Austrian Federal Institute of Health provided information on the capacity of the dialysis centres in all nine Austrian states for the year 1998. Furthermore, the driving time between the patient's home address and the nearest dialysis facility was also calculated by the Austrian Federal Institute of Health [8].

The National Health Interview Survey (NHIS) covers the prevalence of chronic diseases, the extent of disability and the use of health care services. Data obtained during the 1991 [9] and 1995 [10] surveys were used.

The National Hospital Discharge Registry (NHDR) was screened for the years 1994–1998 to determine hospital admissions due to specific diseases using the International Classification of Diseases, Ninth Revision (ICD-9) [11]. Numbers are direct age-adjusted rates per million population (p.m.p.). In addition, data from the National Mortality Registry (MORT), which compiles and codes information on all deaths in Austria, was used for the years 1994 and 1995 [12].

From IMS-HEALTH AUSTRIA, one of the world's biggest market research institutions, data about the sale of specific categories of drugs (e.g. oral hypoglycaemic agents) in public pharmacies in each state were obtained (National Drug Wholesale Registry, provided by U. Scheithauer, IMS HEALTH AUSTRIA. Vienna, 2000). The numbers used herein represent units p.m.p. for the year 1999 (one unit corresponds to \sim 100–120 pills).

Data about the percentage of the population with a body mass index (BMI) > 30 were received from Kunze *et al.* [13].

Statistical analysis

Data are expressed as mean respective rates with a 95% confidence interval (CI).

For better comparability with other countries directly, agestandardized rates were computed according to the usual definition [14]. Age categories were defined as 0–14, 15–29, 30–44, 45–64, 65–74 and 75+ years. Weights were derived from the Austrian population in the year 1997. All computations were carried out with the SPSS programming language, which is suitable for computation of agestandardized rates (SPSS 10.0, SPSS Inc., Chicago, IL).

Significance tests for difference of age-standardized rates were based on Mantel-Haenszel score statistics [14]. We report the significance of rate differences by adapting the Bonferroni correction for the α -level which is known to be very conservative; therefore, we also present *P*-values.

Correlation between the incidence of ESRD due to type 2 diabetes mellitus (DM-2) and wholesale oral hypoglycaemic sales with respect to the percentage of the population with a BMI > 30 kg/m² was analysed by using a weighted linear regression model; weights were defined by population sizes. We report R^2 -factor and *P*-value.

Results

Between January 1, 1995 and the end of December 1999, 4811 new cases of treated ESRD were recorded in Austria. Table 1 summarizes data on the number of new ESRD cases, the numbers for each age group

Table 1. Summary registration of ESRD incidence 1995-1999 in Tyrol and rest of Austria (RA)

State	No. of cases	Age cate	egory	Age-standardized incider				
		0-14	15–29	30–44	45–64	65–74	+75	rate (95% CI) p.m.p.
Tyrol	298	3 1%	23 7.7%	41 13.8%	101 33.9%	79 26.5%	51 17.1%	97.9 (86.8–109.1)
RA	4513	33 0.7%	176 3.9%	488 10.8%	1676 37.1%	1342 29.7%	798 17.7%	120.9 (117.4–24.5)*

*Significant vs Tyrol (P < 0.001, P < 0.01 after Bonferroni correction).

Table 2.	End-s	age renal	disease	incidence	1995-	-1999	according to	o primary	renal	disease	in	the	rest	of .	Austria	(RA)) and	Tyrol	(p.m.p	·.)
----------	-------	-----------	---------	-----------	-------	-------	--------------	-----------	-------	---------	----	-----	------	------	---------	------	-------	-------	--------	-----

	RA	Tyrol	<i>P</i> -value	
Total	120.9 (116.9–124.9)	97.9 (86.9–11.0)	0.001**	
DM-2	28.9 (27.2–30.6)	12.2 (8.2–16.2)	< 0.001**	
Vascular nephropathy	18.1 (16.7–19.5)	10.3 (6.6–14.0)	0.002^{*}	
Glomerulonephritis	16.8 (15.5–18.1)	15.3 (11.0–19.6)	0.589	
Unknown origin	15.8 (14.5–17.1)	15.4 (11.0–19.8)	0.857	
Interstitial nephritis	13.9 (12.7–15.1)	11.9 (8.1–15.7)	0.424	
Hereditary kidney disease	8.6 (7.79–9.4)	6.5 (3.7–9.3)	0.197	
DM-1	7.3 (6.4–8.2)	10.8 (7.2–14.4)	0.034	
Others	11.2 (10.2–12.2)	13.7 (9.8–17.6)	0.204	
Without DM-2	92.1 (89.0–95.2)	85.7 (75.7–95.7)	0.289	

*P < 0.05, **P < 0.01 after Bonferroni correction.

and the mean annual age-adjusted incidence for Tyrol (T) and the rest of Austria (RA), which consists of all states with the exception of Tyrol.

As can be clearly seen, a large regional difference was observed, with an incidence rate of 97.9 (95% CI 86.8–109.1) in T in comparison with 120.9 (95% CI 117.4–124.5) for RA (P < 0.001, P < 0.01 after Bonferroni correction).

Additionally, we analysed the ESRD incidence rate for T and RA according to the main groups of renal diseases causing terminal renal insufficiency as defined by the Austrian Society of Nephrology (Table 2). There were significant differences with regard to the incidence of patients with DM-2 [T = 12.2 (95% CI 8.2-16.2) vs RA = 28.9 (95% CI27.2–30.6), P < 0.001, P < 0.01 after Bonferroni correction] and vascular nephropathy [T = 10.3 (95%)]CI 6.6-14.0) vs RA = 18.1 (95% CI 16.7-19.5), P = 0.002, P < 0.05]. In contrast, other renal diseases had a similar frequency throughout Austria. When patients with DM-2 were excluded from the analysis, the incidence of ESRD patients was identical in T and RA [T = 85.7 (95% CI 75.7–95.7) vs RA = 92.1 (95% CI 88.9–95.1), P = 0.289].

From these data, we attempted to verify whether the low incidence of ESRD patients with DM-2 in Tyrol can be explained by a lower frequency of this disease in the general population. As can be seen from Table 3, each data source screened revealed a lower rate of diabetes mellitus in Tyrol as compared with the rest of Austria. Furthermore, we carried out

 Table 3. Frequency of diabetes mellitus in Tyrol and the rest of Austria (p.m.p.)

County	NHIS 91	NHIS 95	NHDR	MORT	DRUG
Tyrol	11 830	13 000	3941	165	195 291
RA	16 790	24 220	4533	218	275 175

NHIS 91, National Health Interview Survey 1991; NHIS 95, National Health Interview Survey 1995; NHDR, National Hospital Discharge Registry 1994–1998; MORT, National Mortality Registry 1994–1995; DRUG, National Drug Wholesale Registry 1999, oral hypoglycaemic agents.



Fig. 1. Incidence of end-stage renal disease due to diabetes mellitus type 2 and wholesale oral hypoglycaemic sales in 1999 (p.m.p.) for Austria's nine states.

a weighted linear regression model between the ESRD incidence rate and the wholesale oral hypoglycaemic sales for each of the nine Austrian counties; for this an $R^2 = 0.59$, P = 0.026 was found (Figure 1).

When data on BMI for Austria were obtained from Kiefer *et al.* [13], a weighted linear regression between the prevalence of a BMI > 30 kg/m^2 and the incidence rate of ESRD patients with DM-2 was also significant ($R^2 = 0.47$, P = 0.041) (Figure 2).

Discussion

Epidemiological studies have long been used to aid public health investigations by identifying geographic areas of elevated incidence of disease, from John Snows' cholera survey to recent investigations of cancer, stroke or myocardial infarction clusters. So far, only a few studies have dealt with the issue of



Fig. 2. Incidence of end-stage renal disease due to diabetes mellitus type 2 and percentage of the population with a BMI $> 30 \text{ kg/m}^2$ for Austria's nine states.

regional differences in the incidence of ESRD. Relman *et al.* [15], studying the US population, considered patient selection by doctors to be largely responsible for marked variations. Another study by Rosansky *et al.* [4] stratified the various states of the USA for differences in race, gender and age composition, and performed several regression analyses. This study mainly showed that states with relatively low rates of ESRD treatment tended to have fewer diabetic nephropathy cases, which was also suggested by two other studies [5,6].

The Austrian Dialysis and Transplant Registry consistently reports remarkable regional variability in new patients for RRT in Austria, a small country with a quite homogeneous population. The lowest incident patient rates are reported for Tyrol. It could be possible that the regional variability in ESRD is paralleled by a variability in the prevalence of diseases that lead to ESRD such as diabetes mellitus (real lower incidence). However, several other factors might also be involved such as missed patients due to failure to diagnose renal disease, patient or physician reluctance to accept RRT as a treatment option, a lack of dialysis facilities that would limit access, or early patient death before ESRD has developed.

Real lower incidence

As an initial step, we categorized the incident ESRD patient population in Tyrol and the rest of Austria according to eight main diagnostic groups as defined in the questionnaire drawn up by the Austrian Society of Nephrology, and this clearly showed that the observed difference was due mainly to a lower incidence of patients entering RRT with the diagnosis of DM-2. We then proceeded to evaluate whether the finding of a low prevalence of DM-2 is confined only to the dialysis population. As no diabetes registry is available in Austria, we had to rely on several other data sources, which included two National Health Interview Surveys (NHIS 1991 and 1995), the National Hospital Discharge Registry (NHDR), the National Mortality Registry (MORT) and the National Drug Wholesale Registry (DRUG). These registries provide estimates of the frequency of diabetes at different levels of the health care system. The National Health Interview Survey reports the participant's self-perception of disease. The National Drug Wholesale Registry covers patients receiving medication, the National Hospital Discharge Registry records patients discharged from hospitals with a specific ICD-9 code, and the National Mortality Registry identifies subjects suspected of having died from complications related to a disease. It was very convincing for us to find that all these different sources showed a uniform situation, with Tyrol having low rates of diabetes mellitus, e.g. the prevalence rate stated in the two National Health Interview Surveys gives a number of 11831/p.m.p. for 1991 and 13 000 for 1995 for the Tyrolean population reporting diabetes, in comparison with the Austrian average of 16787 (1991) and 24216 (1995) (Table 3).

However, even though all these data show a picture of low diabetes rates in Tyrol, several limitations have to be mentioned. One might argue that the National Drug Wholesale Registry cannot be used to calculate the actual number of patients treated. As far as the results of our study are concerned, however, one would have to assume that patients in various parts of Austria are prescribed different numbers of pills for treatment of diabetes. This seems quite unlikely. Under-reporting of diabetes in all presented data could influence the result of our study, but we assume that this presents more a nationwide problem and thus should not affect state-specific differences. We are also aware of the potential pitfalls of diagnosing DM-2 nephropathy. In the majority of patients with DM-2, it is undoubtedly classical Kimmenstiel-Wilsons's glomerulosclerosis which leads to ESRD, but primary renal diseases such as glomerulonephritis, ischaemic nephropathy, etc. may occur more frequently than expected by chance. However, on the other hand, the prevalence of DM-2 probably could be underestimated.

How can we explain low diabetes rates in Tyrol? The prevalence of DM-2 is partly determined by genetic factors [16] but, unlike in the USA, where racial and cultural differences affect the incidence of ESRD [7], the Austrian population seems to be relatively homogeneous and no study until now has identified genetic isolation in Tyrol, an area where the majority of the population lives in mountainous areas.

On the other hand, obesity has been shown to be one of the most important factors triggering DM-2 [17]. Kiefer *et al.* [8] demonstrated a regional Downloaded from ndt.oxfordjournals.org at Medizen Biolog Fachbibliothek on November 14, 2010

1566

variability in the distribution of subjects with a BMI $> 30 \text{ kg/m}^2$ in Austria, which paralleled our data on DM-2-associated ESRD (Figure 2). The National Health Interview Survey of 1991 also showed Tyrol to be the state with the highest percentage of persons (50.8%) in Austria who take regular physical exercise [9].

Nonetheless, other factors might also contribute to a regional variability in RRT incidence, and we also tried to assess their importance.

Insufficient diagnosis

Insufficient diagnosis of renal disease in general might also influence RRT incidence, as suggested by several authors [6]. However, a search of the National Hospital Discharge Registry for diseases such as glomerulonephritis, nephritis and/or infectious kidney diseases showed the number of hospital stays in Tyrol (1560/year p.m.p.) to be identical to that for the rest of Austria (1562/year p.m.p.) (P > 0.05).

Reduced acceptance into RRT

Sekkarie *et al.* [18] suggested that the extent of comorbidity could influence a physician's decision to refer a patient to RRT. The Austrian Dialysis and Transplant Registry collects data on patients who have undergone at least 3 months of RRT. If 'sick' patients were indeed excluded from RRT in Tyrol, one would expect the mortality rate for these 3 months to be much lower than in the rest of Austria. However, in Tyrol, 5.61% of ESRD patients die within 90 days of initiation of RRT, a figure similar to that observed in the rest of Austria (7.52, P > 0.05). Another hypothesis states that the distance to the nearest dialysis facility could affect patients' reluctance [19], but a study showed that almost all patients in Austria could reach a unit by car within 40 min [8].

Lack of facilities

Data obtained from the Austrian Federal Institute of Health demonstrate that dialysis capacity in Tyrol is much higher than in the rest of Austria [8], which excludes a lack of access to dialysis treatment due to limited health care resources. Additionally, since all costs for ESRD care in Austria are covered by public health insurance, we feel that socio-economic reasons cannot explain regional variabilities.

Early death

Finally, it has been suggested that patients die from other causes before reaching ESRD [20]. Data on

cardiovascular mortality in general show that this rate in Tyrol is $\sim 15\%$ lower than the Austrian average [12].

The aim of the present study was to identify factors involved in the development of renal failure for possible use in a preventive programme. We show here that the low incidence of RRT in Tyrol can be explained mainly by low rates of DM-2. As the population in Tyrol has the lowest percentage of persons with a BMI $> 30 \text{ kg/m}^2$ and the highest percentage taking regular physical exercise in Austria, our findings suggest that general health care preventive measures are able to reduce target organ diseases significantly.

Acknowledgements. We thank all 64 dialysis and transplantation centres in Austria for providing their data to the Austrian Dialysis and Transplantation Registry, as well as Statistic Austria, IMS HEALTH and the Austrian Federal Institute of Health for providing their valuable data.

Conflict of interest statement. None declared.

References

- 1. Van Dijk PC, Jager KJ, de Charro F *et al*. Renal replacement therapy in Europe: the results of a collaborative effort by the ERA-EDTA registry and six national or regional registries. *Nephrol Dial Transplant* 2001; 16: 1120–1129
- Usami T, Koyama K, Takeuchi O et al. Regional variations in the incidence of end-stage renal failure in Japan. J Am Med Assoc 2000; 284: 2622–2624
- Ruggenenti P SA, Remuzzi G. Progression, remission, regression of chronic renal diseases. *Lancet* 2001; 357: 1601–1608
- Rosansky SJ, Huntsberger TL, Jackson K et al. Comparative incidence rates of end-stage renal disease treatment by state. Am J Nephrol 1990; 10: 198–204
- Foxman B, Moulton LH, Wolfe RA *et al.* Geographic variation in the incidence of treated end-stage renal disease. *J Am Soc Nephrol* 1991; 2: 1144–1152
- Moulton LH, Port FK, Wolfe RA *et al.* Patterns of low incidence of treated end-stage renal disease among the elderly. *Am J Kidney Dis* 1992; 20: 55–62
- Klag MJ, Whelton PK, Randall BL et al. End-stage renal disease in African-American and white men. 16-year MRFIT findings. J Am Med Assoc 1997; 277: 1293–1298
- 8. Kramar R. Austrian Dialysis and Transplantation Registry, OEDTR 1999 Annual Report: Atlas of End-stage Renal Disease in Austria; 2000
- 9. Statistic Austria. Mikrozensus 1991. Gesundheitszustand der oesterreichischen Bevölkerung [in German]. Vienna; 1996
- Statistic Austria. Mikrozensus 1995. Personen mit koerperlichen Beeintraechtigung [in German]. Vienna; 1996: 226–227
- Statistic Austria. Gesundheitsstatistisches Jahrbuch 1997/1998/ 1999 [in German]. Vienna; 1999
- 12. Statistic Austria. Österreichischer Todesursachenatlas 1988/94 [in German]. Vienna; 1998
- Kiefer I, Kunze U, Mitsche N et al. Übergewicht in Österreich: Epidemiologische und sozialmedizinische Aspekte [in German]. Acta Med Austriaca 1998; 25: 126–128
- Rothman KJ, Greenland S. Modern Epidemiology, 2nd edn. Lippincott-Raven, Philadelphia; 1998
- 15. Relman AS, Rennie D. Treatment of end-stage renal disease: free but not equal. *N Engl J Med* 1980; 303: 996–998

- Krolewski AS. Genetics of diabetes nephropathy: evidence for major and minor effects. *Kidney Int* 1999; 55: 1582–1596
- Hillier TA, Pedula KL. Characteristics of an adult population with newly diagnosed type 2 diabetes: the relation of obesity and age of onset. *Diabetes Care* 2001; 24: 1522–1527
- 18. Sekkarie M, Cosma M, Mendelssohn D. Nonreferral and nonacceptance to dialysis by primary care physicians and

nephrologists in Canada and the United States. Am J Kidney Dis 2001; 38: 36-41

- Boyle PJ, Kudlac H, Williams AJ. Geographical variation in the referral of patients with chronic end stage renal failure for renal replacement therapy. Q J Med 1996; 89: 151–157
- Lippert J, Ritz E, Schwarzbeck A et al. The rising tide of endstage renal failure from diabetic nephropathy type II—an epidemiological analysis. Nephrol Dial Transplant 1995; 10: 462–467

Received for publication: 21.5.02 Accepted in revised form: 12.2.03

B2-3 Appendix Full Papers

Anni Heinaioi (2003) 04. 217–222 DOI 10.1007/s00277-004-0936-x

ORIGINAL ARTICLE

Michael A. Fridrik · Hubert Hausmaninger · Werner Linkesch · Richard Greil · Otto Krieger · Christian Baldinger · Joerg Klocker · Joerg Pont · Willhelm Oberaigner

Long-term results of dose density therapy in patients with aggressive lymphoma

Received: 10 February 2004 / Accepted: 31 July 2004 / Published online: 6 November 2004 © Springer-Verlag 2004

Abstract To evaluate the long-term outcome of dose density chemotherapy in the treatment of aggressive lymphoma, we analyzed 142 patients with untreated aggressive lymphoma. Chemotherapy was an eight-drug regimen given in weekly intervals in two prospective trials. The median observation period was 8 years; the longest follow-up was 13 years. Overall survival at 8 years was 0.583. The 8-year survival of patients \leq 60 years was significantly better than that of older patients, namely

Conflict of interest: there are no conflicts of interest.

M. A. Fridrik (⊠) Department of Oncology, AKH Linz, Linz, 4020, Austria e-mail: michael.fridrik@akh.linz.at

H. Hausmaninger Department of Medicine 3, LKH Salzburg, Salzburg, 5020, Austria

W. Linkesch Department of Hematology, University of Graz, Graz, 8020, Austria

R. Greil Department of Medicine, University of Innsbruck, Innsbruck, 6020, Austria

O. Krieger KH Elisabethinen, Linz, 4010, Austria

C. Baldinger Department of Hematology-Oncology, KH Barmh. Schwestern, Wels, 4600, Austria

J. Klocker Department of Medicine, LKH Klagenfurt, Klagenfurt, 8020, Austria

J. Pont Department of Medicine, KFJ, Vienna, 1100, Austria

W. Oberaigner TILAG, Institute for Clinical Tumor Epidemiology, 6020 Innsbruck, Austria 0.713 vs 0.304 (p=0.000000697). This excellent survival of patients aged ≤ 60 years was identical for high-risk and high-intermediate-risk patients compared with low-risk and low-intermediate-risk patients in the age-adjusted international prognostic index (IPI). The excellent long-term results of the CEOP/IMVP-Dexa regimen (cyclophosphamide, epirubicin, vincristine, and prednisone/ ifosfamide with systemic mesna, methotrexate, etoposide, and dexamethasone) for patients aged ≤ 60 years suggest that this regimen might be superior to the standard CHOP regimen (cyclophosphamide, doxorubicin, vincristine, and prednisolone) and needs to be tested in comparison to high-dose regimens and novel approaches including antibody treatment.

Keywords Lymphoma · Large-cell lymphoma · Drug therapy · Antineoplastic combined chemotherapy protocols

Introduction

Aggressive non-Hodgkin's lymphoma (NHL) is a fatal disease. Survival depends on the histological subgroup [20] and the international prognostic index (IPI) [19]. The CHOP regimen (cyclophosphamide, doxorubicin, vincristine, and prednisolone) is the current standard in first-line treatment of aggressive lymphoma. Despite a high remission rate, more than 50% relapse and will eventually die.

So far it has not been possible to show that more intensive standard dose regimens achieve better survival than CHOP [12]. CEOP/IMVP-Dexa (cyclophosphamide, epirubicin, vincristine, and prednisone/ifosfamide with systemic mesna, methotrexate, etoposide, and dexamethasone) is a multidrug dose density regimen with weekly chemotherapy. It is a hybrid of a CHOP-like regimen fused with IMVP-16. In order to maintain a high-dose density, the doses were not reduced as long as neutrophil counts were higher than 1.0 G/l. With this regimen we were able to achieve a high remission rate, a long time to treatment failure (TTF), and a long time to relapse (TTR) [13]. These

are excellent surrogate markers for the outcome of treatment in aggressive lymphoma. However, overall survival is the most important endpoint. Another concern is long-term toxicity in terms of secondary cancer, which can occur up to 25 years after therapy [17].

Our trials were based on the hypothesis that dose density is an important factor in curing aggressive lymphoma. We started our trials in 1988 and followed this concept over three consecutive trials. After a median observation period of 8 years, the results of the first two trials seem to be stable with no further decrease in the outcome.

Materials and methods

One hundred forty-two patients from two consecutive trials were pooled for this analysis. In the first study, a phase 2 trial from October 1988 to March 1991, we assessed the feasibility, toxicity, and efficacy of this new dose-dense regimen in a multicenter setting [13]. From this study two patients were excluded because of human immunodeficiency virus (HIV)-positive serology. The patients of the second trial (July 1991 to November 1995) were randomized to receive CEOP/IMVP-Dexa with and without filgrastim [14]. The patient population was comparable with inclusion and exclusion criteria identical to the first trial. The endpoint was febrile neutropenia. There was no difference between the two randomized groups in respect of remission rates, TTR, TTF, and survival. Inclusion and exclusion criteria, staging modalities, toxicity, and details of the chemotherapy regimen have been published elsewhere [13, 14]. Both studies were done according to the Helsinki Declaration; the local Ethics Committees of the participating centers approved them. All patients signed a confirmed consent. All patients had a previously untreated aggressive lymphoma. The chemotherapy regimen is outlined in Table 1; no intrathecal prophylaxis was given. Irradiation to residual lymphoma or to regions with bulky disease at diagnosis was allowed and was given on discretion of the treating physician. The longest observation is 13 years; the median observation time for living patients is 8 years.

The median age of the patients was 52 years (range: 19– 72 years). Forty-eight patients (33.8%) were >60 years. The number of patients in the low-risk, low-intermediaterisk, high-intermediate-risk, and high-risk groups according to the IPI were 71 (50%), 34 (24%), 25 (18%), and 12 (8%), respectively. Of the 142 patients, 22 (15.5%), 51 (35.9%), 24 (16.9%), and 45 (31.7%) had stage 1, 2, 3, and 4 disease, respectively. Histologies were centrally reviewed. Diffuse large-cell lymphoma was found in 116 patients (81.46%), lymphoblastic B in 5, Burkitt-like in 1, and anaplastic large cell 0-cell primary systemic type in 3 patients. Seventeen (12%) patients had a lymphoma of the T-cell phenotype, two lymphoblastic T, nine peripheral Tcell, and six anaplastic large cell T-cell primary systemic type.

Biostatistics

All eligible patients were included in the analysis. Survival estimates were calculated by the Kaplan-Meier method. Comparison of survival between the different risk groups was done by the log-rank test, between remission rates or other events by the κ^2 -test.

The TTR and the TTF were defined as proposed by Dixon et al. [9]. Briefly, survival included all eligible patients and counted all deaths as events. TTF was the time from registration until relapse, progression, toxic death, withdrawal, or the date the patient was last known to be alive, excluding deaths from unrelated causes. TTR was the time from registration until relapse or the date the patient was last known to be alive, including only complete responses (CRs) and counting only relapses as events.

For comparison of the occurrence of second primaries in the study with the expected rate of malignancies in a normal population, an age-adjusted and sex-adjusted sample was used.

Results

One hundred nine patients (76.8%) achieved a complete (CR) and 22 (15.5%) a partial response (PR). Stable disease (SD), and progression (PD) during therapy, was observed in three patients each. Five patients were not evaluable for response, four because of early death and one patient because of loss to follow-up. The mean dose intensity was 74% of the planned dose.

Fifty-four patients (38%) have relapsed to date. Twentyfive (46.3%) relapsed within an originally involved site,

Table 1 Chemotherapy men CEOP/IMVP-Dexa	regi-	Cyclophosphamide	750 mg/m^2	i.v. 30 min	Day 1
		Epirubicin	70 mg/m^2	i.v. 2 h	Day 1
		Vincristine	1.4 mg/m^2	i.v. push	Days 1+8
		Prednisolone	100 mg	p.o.	Days 1–5
		Ifosfamide	2000 mg/m^2	i.v. 30 min.	Days 15–17
		Uromitexan	400 mg/m^2	i.v. push at 0, 4, and 8 h after ifosfamide	Days 15–17
		VP-16	100 mg/m^2	i.v. 2 h	Days 15–17
		Dexamethasone	40 mg	p.o.	Days 15–19
		Methotrexate	800 mg/m^2	i.v. 4 h	Day 22
		Ca folinate	15 mg/m^2	p.o. every 6 h	Days 23-25

Table 2 Toxic deaths

Age (years)	Histology	IPI	No. of risk factors	No. of days after entry	Cause of death
69	DLBCL	Low	1	11	Pneumonia
59	DLBCL	High	4	12	Neutropenic fever
66	DLBCL	Low-intermediate	2	14	Pulmonary embolism
72	DLBCL	High	4	22	Cardiopulmonary failure
63	DLBCL	Low	1	29	Neutropenic fever
64	Peripheral T-cell	High	3	30	Urosepsis
63	DLBCL	Low	1	51	Colitis, pneumonia
62	DLBCL	High-intermediate	3	63	Pneumonia
52	DLBCL	Low	0	121	Pneumonia
68	DLBCL	Low-intermediate	2	133	Pneumonia

12 (22.2%) within an originally involved site and within an originally uninvolved site, and 11 (20.4%) within an originally uninvolved site. For six patients the site of relapse was unknown. Only two patients had a relapse in the central nervous system (CNS). Of the relapsing patients, 27 (50%), 13 (24%), 6 (11%), 3 (5.6%), and 4 (4.7%) relapsed in the 1st, 2nd, 3rd, 4th, and 5th to 8th year, respectively. The only relapse after 8 years occurred as late as 12 years after study entry. Biopsy results of relapses were not systematically collected. All relapses from which we know that they had a biopsy were of the same histological subtype as the first manifestation.

Fifty-five patients have died: 39 (70.9%) of lymphoma, 10 (18.2%) of toxicity, and 6 (10.9%) due to unrelated causes. The treatment-related death rate was 7%. Eight of ten toxic deaths occurred in patients >60 years (p=0.01667), most of them in the first cycle. IPI score at diagnosis had no influence on toxic deaths (Table 2). Acute toxicity has been reported elsewhere [13, 14].

In 980 patient years, six second primaries occurred. One carcinoma of the bile duct occurred 20 months after entry to the study and the patient died 3 months later. One esophageal carcinoma occurred 42 months after entry and the patient died 52 months later of the tumor. One patient developed a melanoma 6 years after entry into the study, but was lost to follow-up thereafter. One patient was nephrectomized because of a right-sided renal cell carcinoma 90 months after entry to the study. He was free of disease 50 months after the nephrectomy. In one single patient two cancers developed. First, a papillary thyroid cancer 92 months after study entry developed. A thyroidectomy and therapy with radiolabeled iodine were done. The patient was cured from this disease. Consecutively, chronic myeloid leukemia developed 140 months after entry to the study. He was treated with imatinib and hydroxyurea thereafter, never achieved a hematological remission, and died 37 months after the diagnosis of chronic myeloid leukemia. No other long-term toxicities were observed.







TTR at 8 years for patients with CR was 0.619 [95% confidence interval (CI): 0.53–0.72]. The TTF at 8 years was 0.536 (95% CI: 0.457–0.63). Overall survival at 3, 5, and 8 years was 0.695 (95% CI: 0.623–0.777), 0.648 (95% CI: 0.572–0.7340), and 0.583 (95% CI: 0.503 – 0.665), respectively (Fig. 1).

CR rates for patients ≤ 60 years were not significantly higher than those for patients >60 years (81.9 vs 66.7%, p=0.25). Overall survival at 8 years for patients ≤ 60 years was 0.713 (95% CI: 0.662–0.816). In patients >60 years the overall survival was only 0.304 (95% CI: 0.192–

difference 0.453). This was highly significant (p=0.000000697) (Fig. 2). This difference may be due to several reasons. Only 4 of 16 patients dying of causes other than lymphoma were >60 years. Remissions in elderly patients are not as stable as in younger patients [19]. Salvage regimens are more effective in younger patients; especially high-dose therapy is preferably used in patients ≤60 years of age. Patients in the high-risk IPI group had a significantly worse survival than patients in the other groups. However, there was no difference in overall survival in patients ≤ 60 years with 0 or 1 point vs 2









or 3 points according to the age-adjusted IPI (Fig. 3). In terms of TTF for younger patients the IPI did not make a difference (Fig. 4). For older patients there seems to be a difference, but the patient numbers in the different groups was too low to reach statistical difference (Fig. 5). There was no statistically significant overall survival difference between the different histological subgroups (data not shown).

Discussion

A long follow-up is necessary to assess the final value of a treatment regimen in aggressive lymphoma. The results are mature and the final value of this dose density therapy can be assessed. In our analyses 15% of relapses occurred



Fig. 5 Time to treatment failure for patients ≤ 60 years according to IPI (low=0-1, low-intermediate=2, high-intermediate=3, and high=4-5 risk factors)

after 3 years of follow-up. This compares well with the outcome of other dose-intense regimens [10, 16].

Patients >60 years had an unacceptably high toxic death rate and did not seem to benefit from this regimen (Fig. 2). Treatment options other than dose density should be used for older patients. Targeted therapy may be one of these strategies [6].

CNS recurrence is common in advanced aggressive NHL. The treatment recommendations for CNS disease are radiotherapy, intrathecal therapy, or intravenous highdose methotrexate [11]. Usually a methotrexate dose higher than 800 mg/m² is necessary to reach adequate levels in the CNS. However, only two CNS relapses occurred in our patients although no intrathecal prophylaxis was given. This is remarkable because 79 patients had stage 3 or 4 and were, therefore, at risk for a CNS relapse. High-dose dexamethasone may have played a role in the low CNS relapse rate. Recently, a lower CNS relapse rate could be found for dexamethasone over prednisone in a randomized trial in children with lymphoblastic leukemia [4].

Second malignancies are a serious concern for all antineoplastic therapies [17]. According to estimates, a cohort of normal persons comparable to the patients treated in our studies would experience 8.8 second malignancies. However, in the present study we encountered six second malignancies. Although the rate of second malignancies is lower than expected for a normal population, even lower incidence rates of 2.75% are described for other dose intensive regimens such as the ACVBP of the GELA Group [1]. It was surprising that we had no patient with myelodysplastic syndrome or acute myeloid leukemia. This may be just by chance because of the low number of patients in comparison to other reports [1].

Although comparisons of different trials are prone to several biases, CEOP/IMVP-Dexa with an overall survival of 58% may be better than the long-term results of the CHOP regimen [12, 8, 2]. An even better survival was observed for patients aged ≤ 60 years, including those in the high-risk and high-intermediate-risk groups of the ageadjusted IPI [19] [(Figs. 2, 3)]. For these patients, highdose chemotherapy with stem cell support has proven to be without any benefit in the majority of the trials [7]. Similarly, dose intensification after incomplete chemotherapy could not improve the results [15]. The same is true for dose intensification in slowly responding patients [21]. However, several lines of evidence point to a role of early dose intensification [3, 5]. The German Non-Hodgkin's Lymphoma Study Group reported a significant improvement of the CR rate and the 5-year event-free survival but not for overall survival by adding etoposide to the CHOP regimen for patients ≤60 years of age with a normal lactic dehydrogenase (LDH) level [18]. These studies are supported by our data, where patients ≤ 60 years in the high-risk and intermediate-high-risk groups according to the age-adjusted IPI [19] achieved a long-term survival rate as good as that of the low-risk patients receiving dose density therapy (Fig. 3).

Several biases can influence comparisons between different trials. To prove the superiority of CEOP/IMVP-Dexa, we recently finished accrual to a randomized trial with standard CHOP. Final results are pending.

References

- 1. André M, Mounier N, Leleu X, et al. (2004) Second cancers and late toxicities after treatment of aggressive non-Hodgkin lymphoma with the ACVBP regimen: a FELA cohort study on 2837 patients. Blood 103:1222–1228
- Armitage JO, Dick FR, Corder MP, et al. (1982) Predicting outcome in patients with diffuse histiocytic lymphoma treated with cyclophosphamide, adriamycin, vincristine and prednisone (CHOP). Cancer 50:1695–1702
- 3. Blayney DW, LeBlanc ML, Grogan T, et al. (2003) Dose intense, every 2 weeks with dose-intense cyclophosphamide, doxorubicin, vincristine, and prednisone may improve survival in intermediate- and high-grade lymphoma: a phase II study of the Southwest Oncology Group (SWOG 9349). J Clin Oncol 21:2466–2473
- 4. Bostrom BC, Sensel MR, Sather HN, et al. (2003) Dexamethasone versus prednisone and daily oral versus weekly intravenous mercaptopurine for patients with standard-risk acute lymphoblastic leukemia: a report from the Children's Cancer Group. Blood 101:3809–3817
- Coiffier B (1995) Fourteen years of high dose CHOP (ACVBregimen): preliminary conclusions about the treatment of aggressive lymphoma patients. Ann Oncol 6:211–217

- Coiffier B (2003) Increasing chemotherapy intensity in aggressive lymphomas: a renewal? J Clin Oncol 21:2457–2459
- Coiffier B, Lepage E, Briere J, et al. (2002) CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med 346:235–242
- Coltman CA, Luce JK, McKelvey EM, et al. (1977) Chemotherapy of non-Hodgkin's lymphoma: ten years experience in the Southwest Oncology Group. Cancer Treat Rep 61:1067– 1078
- Dixon DO, McLaughlin P, Hagemeister FB, et al. (1987) Reporting outcomes in Hodgkin's disease and lymphoma. J Clin Oncol 5:1670–1672
- 10. Dumontet C, Bastion Y, Felman P, et al. (1992) Long-term outcome and sequelae in aggressive lymphoma patients treated with the LNH-80 regimen. Ann Oncol 3:639–644
- 11. Ferreri AJM, Abrey LE, Blay JY, et al. (2003) Summary statement on primary central nervous system lymphomas from the Eighth International Conference on Malignant Lymphoma, Lugano, Switzerland, June 12 to 15, 2002. J Clin Oncol 21:2407–2414
- Fisher RI, Gaynor ER, Dahlberg S, et al. (1993) Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. N Engl J Med 328:1002–1006
- Fridrik MA, Hausmaninger H, Linkesch W, et al. (1996) CEOP-IMVP-Dexa in the treatment of aggressive lymphomas: an Austrian multicenter trial. J Clin Oncol 14:227–232
- 14. Fridrik MA, Greil R, Hausmaninger H, et al. (1997) Randomized open label phase III trial of CEOP/IMVP-Dexa alternating chemotherapy and filgrastim versus CEOP/IMVP-Dexa alternating chemotherapy for aggressive non-Hodgkin's lymphoma (NHL). A multicenter trial by the Austrian Working Group for Medical Tumor Therapy. Ann Hematol 75:135–140
- Gisselbrecht C, Lepage E, Molina T, et al. (2002) Shortened first-line high-dose chemotherapy for patients with poorprognosis aggressive lymphoma. J Clin Oncol 20:2472–2479
- Lee AY, Connors JM, Klimo P, et al. (1997) Late relapse in patients with diffuse large-cell lymphoma treated with MACOP-B. J Clin Oncol 15:1745–1753
- Leung W, Sandlund JT, Hudson MM, et al. (2001) Second malignancy after treatment of childhood non-Hodgkin lymphoma. Cancer 92:1959–1966
- Pfreundschuh M, Truemper L, Kloess M, et al. (2004) 2-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of young patients with good prognosis (normal LDH) aggressive lymphomas: results of the NHL-B1 trial of the DSHNHL Blood First Edition Paper, prepublished online 24 February 2004. Blood 104:626–633
- Shipp MA (1993) The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 329:987–994
- The Non-Hodgkin's Lymphoma Classification Project (1997) A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. Blood 89:3909– 3918
- 21. Verdonk LF, Vanputten W, Hagenbeek A, et al. (1995) Comparison of CHOP chemotherapy with autologous bone marrow transplantation for slowly responding patients with aggressive non-Hodgkin's lymphoma. N Engl J Med 332:1045–1051