General introduction

AIMS

Basal cell carcinoma (BCC) is the most common cancer in white-skinned people with increasing incidence rates, burden of disease and health care costs worldwide. In addition, a significant part (≥30%) of patients with a first BCC will develop at least a second new BCC or another ultraviolet radiation related cutaneous malignancy in time (i.e., metachronous skin cancers). Therefore it is an important topic for patients, physicians and policy makers. Unfortunately it is still unclear which patients are at risk of a metachronous BCC (mBCC) and who need follow-up in the future. Both non-genetic and genetic epidemiological studies of primary/prevalent BCCs have been conducted but studies are scarce when considering patients with mBCC. In this thesis I studied the epidemiology of mBCC using robust methodological approaches. The following main questions addressed in this thesis are presented below:

1. What is already known about the epidemiology of BCC and where are the gaps?
2. What are the non-genetic and genetic predictors of a superficial first BCC?
3. How to deal with the competing risk of death when analyzing metachronous BCCs?
4. What are the non-genetic predictors, absolute risks and cumulative incidences of metachronous BCCs?
5. What are the genetic predictors of multiple/metachronous BCCs?

OUTLINE

In **chapter 2** of this thesis a scholarly (i.e., non-systematic) review of the scientific literature on the epidemiology of BCC is presented. In this review we discussed incidences, trends and differences, burden of disease, risk factors, prevention and health policies, and gaps in existing knowledge were uncovered.

In **chapter 3** we raised the issue of a lack of well-designed and large population-based cohort studies to unravel the epidemiology of mBCC patients.

In **chapter 4** the non-genetic and genetic risk factors of the superficial subtype of BCC were investigated, because previous studies pointed out that this subtype could have a different etiology compared the other BCC subtypes and may be associated with mBCC. The reproducibility of previously found predictors was tested and potential new predictors (both non-genetic and genetic) were studied.

In **chapter 5** we discussed a common problem in survival analysis regarding multiple event data (e.g., mBCC), namely competing risk of death, and showed how to overcome this problem when calculating the probability of a new event. In chapter 6 and 7 we used this knowledge and chose models that could take competing risk into account and produce valid effect measures for the included predictors.

In **chapters 6** and **7** the primary objective was to develop a prognostic model for predicting the absolute risk of mBCC. An extensive literature search showed no other prediction models existed for mBCC. We included non-genetic predictors while adjusting for the competing risk of death. In **chapter 6** the prognostic model was developed for predicting the absolute risk of a second new BCC, whereas in **chapter 7** the follow-up was extended and a third, fourth and fifth new BCC were included as well to see whether the predictors of a second BCC were predictive for the risk of further mBCC. In addition, the frequency and timing of mBCC (i.e., cumulative incidences) was determined.

In the previous two chapters the focus was on non-genetic predictors of mBCC (i.e., patient, lifestyle and tumor-specific characteristics), but genetic predisposition could play a role as well. Therefore (**chapter 8** and **9**) we performed candidate gene approaches with known BCC loci and genome-wide association studies (GWASs) to identify single nucleotide polymorphisms associated with multiple BCC (now called "multiple" instead of "metachronous" because a small part of the included patients only had multiple BCCs on their first diagnosis date and no further BCCs in time), something which had not been done before. In **chapter 8** previously found BCC loci were tested in patients with multiple BCC and a pilot GWAS was conducted to identify susceptibility single nucleotide polymorphisms for multiple BCC. In **chapter 9** we added patients with squamous cell carcinomas to our group with BCC patients, as both tumors are keratinocyte carcinomas, to increase our power and performed both a candidate gene approach and GWAS on multiple keratinocyte carcinomas in collaboration with three different USA cohort studies.

Finally, in **chapter 10**, I answer the five research questions using results derived from this thesis. In addition, limitations of our studies are discussed and implications and future perspectives are given.

DATA SOURCES

In order to answer the five research questions formulated above I have used several data sources, which will be briefly described below. More details can be found in the corresponding chapters.

For the scholarly review in **chapter 2** and to a lesser extent for the commentary in **chapter 3** we have used different comprehensive search strategies in PubMed.

Chapters 4-8 are based on histopathologically confirmed skin cancer data gathered through a linkage between the Rotterdam Study and the Dutch Pathology Registry (PALGA). The Rotterdam Study is an ongoing prospective population-based cohort study of primarily white-skinned people aged 45 years or older living in a well-defined

district of Rotterdam, the Netherlands.[1] The Rotterdam Study started in 1989 and now comprises 14,926 participants. Detailed data were acquired by interviews and by thorough examinations of the participants in a specially built research facility in their district. These steps were repeated every 3-4 years. PALGA is the Dutch nationwide network and registry of histopathology and cytopathology, which was founded in 1971 and achieved complete national coverage in 1991.[2] The skin cancer information of the Rotterdam Study participants was obtained up to 31 December 2013 and over a 1,000 BCC patients could be included in both non-genetic and genetic analyses.

Chapter 9 has also been based upon the data described above, with the addition of data of several USA prospective cohort studies. We collaborated with the research teams of the Nurses' Health Study (NHSI and II) and the Health Professionals Follow-up Study,[3] as well as the research team of the Framingham Heart Study.[4]

REFERENCES

1 Hofman A, Brusselle GG, Darwish Murad S *et al*. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol* 2015; **30**: 661-708.

2 Casparie M, Tiebosch AT, Burger G *et al*. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007; **29**: 19-24.

3 Zhang M, Song F, Liang L *et al*. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet* 2013; **22**: 2948-59.

4 Splansky GL, Corey D, Yang Q *et al*. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007; **165**: 1328-35.