Summary

Oesophageal cancer (EC) is among the top twelve cancers diagnosed in 2022, worldwide. EC has two primary histological forms – oesophageal adenocarcinoma (EAC) and oesophageal squamous cell carcinoma (ESCC). ESCC is the most prevalent histological form, accounting for over 90% of all EC incidence rates worldwide. ESCC exhibits a sharp geographical dispersion, with the highest incidence rates in regions spanning East and Southern Africa (the African ESCC corridor) and Southeast Asia (the Asian ESCC belt). Alcohol and tobacco are the well-known ESCC risk factors, while hot beverages, biomass smoke exposure, poor diet, oral hygiene, and unpiped water use are putative risk factors.

While DNA damage is a necessary step in cancer development, the link between external insults and internal injury (i.e., DNA damage) is not well illustrated in ESCC pathogenesis. DNA adducts, the covalent modifications of DNA, are important biomarkers of exposure and an essential mode of action in chemical carcinogenesis. Considering the vital contribution of external vulnerabilities (i.e., external exposome) and their equivalent biological consequence/responses (i.e., internal exposome, in context of DNA adducts), this doctoral project primarily aimed to establish the association of environmental, lifestyle, or dietary exposures (i.e., the specific components of external exposome, discussed in **Chapter 3**) with DNA adduct predisposition, and further establish links between DNA adducts and increased ESCC risk among individuals residing in high ESCC incidence regions in the African ESCC corridor (**Chapter 7**).

Chapter 1 is the Introductory chapter, providing a comprehensive overview of cancer development, with a particular emphasis on ESCC aetiology. An examination of current epidemiological study designs and their limitations in addressing the persistent knowledge gaps in ESCC aetiology is discussed. To advance knowledge of EC aetiology, a proposal to integrate traditional epidemiological and omic approaches, utilising the exposome framework, is presented. With this background, the chapter then focuses on discussing the essential elements of DNA adduct(omics) analysis in the second half of the introductory chapter.

Chapter 2 discusses the research gaps in ESCC aetiology and presents the objectives of this research. Besides the primary objective (**Chapter 7**) and the review of ESCC external exposome (**Chapter 3**) as already highlighted above, this study aimed to 1) establish knowledge, attitudes, and practices associated with increased ESCC risk in sub-Saharan Africa, (SSA) (**Chapter 4**); 2) conduct a pilot study using an (un)targeted DNA adductomics

approach as a proof of concept (**Chapter 5**); and 3) develop a targeted DNA adduct method for the simultaneous quantification of oxidative and alkylating adducts in human leukocyte DNA samples (**Chapter 6**).

Chapter 3 presents a critical review of lifestyle, socioeconomic, environmental, and dietary factors associated with ESCC risk. Evidence indicates that key ESCC risk contributions arise from specific external exposures (e.g., alcohol, tobacco, hot beverages, biomass fuel, mycotoxin, poor oral health and hygiene). Although the majority of ESCC risk factors are modifiable, the increasing number of ESCC cases observed in the SSA region poses a notable concern, highlighting a possible knowledge deficit regarding ESCC aetiology among high-risk populations. The underlying molecular mechanisms linked to external carcinogens remain limited or inconclusive in high-risk regions of Africa, offering an opportunity to explore further the role of internal exposomes in the pathogenesis of ESCC.

Chapter 4 assessed knowledge, attitudes, and practices (KAP) related to the increased risk of ESCC among individuals living in ESCC endemic regions of SSA. While aetiological evidence of ESCC points to the critical contribution of modifiable ESCC risk factors (as discussed in **Chapter 3**), knowledge concerning individuals' awareness, perceptions, and behaviours regarding ESCC aetiology is limited in the highest-burden regions of Africa. The current study aimed to assess the KAP associated with increased ESCC risk among individuals residing in Malawi, an area with the highest age-standardised ESCC incidence worldwide. The overall ESCC knowledge score was 51% (158/309), the attitude score was 57% (177/309), and the practice score was 52% (161/309). Key predictors of participants' knowledge of ESCC risk factors included being female (adjusted odds ratio [aOR] = 2.1) and having middle-income monthly earnings (aOR = 2.8). The moderate awareness among individuals of the contributing factors to increased ESCC risk and the interplay among various risk factors (as discussed in Chapter 3) may explain the region's heightened disease burden.

In **Chapter 5**, the (un)targeted DNA adduct method (i.e., ultra-high-performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS), intended for use in **Chapter 7**, was tested using human leukocyte DNA samples obtained from kidney transplant recipients (KTRs). Kidney transplantation is a life-saving treatment for patients with kidney failure. While immunosuppressive therapy is crucial for KTRs to prevent graft rejection, it may also contribute to long-term complications, including graft failure and other comorbidities. Additionally, lifestyle and environmental factors may further impact patient outcomes. Aside from the overarching goal (i.e. using it as a proof of concept for **Chapter 7**), this chapter aimed to investigate the link between certain intrinsic (e.g., age) and exogenous risk factors (e.g.,

immunosuppressive drug intake) among KTRs by studying DNA adduct formation. Specific intrinsic and exogenous factors (e.g., smoking) were associated with elevated levels of 8-oxoguanine (8-oxoG), while immunosuppressive drugs were linked to higher levels of DNA alkylation, specifically *N*3-methyladenine (*N*3-meA). Lifestyle habits, particularly alcohol consumption, were associated with elevated levels of α-methyl-γ-hydroxy-1,N^2-propano-2′-deoxyguanosine (CrG) and N^6-methyladenine (N^6-meA). The detection of these DNA adducts in KTRs warrants further research.

In **Chapter 6**, a targeted method using ultra-high performance liquid chromatography–tandem mass spectrometry (UHPLC-MS/MS) was developed to profile oxidative and alkylating DNA adducts in human leukocyte DNA samples. The UHPLC-MS/MS method was assumed to have comparable sensitivity to UHPLC-HRMS for targeted analysis of oxidative and alkylating DNA adducts; therefore, it would be utilised to confirm the structural identities of DNA adducts obtained in **Chapters 5** and **7**. When applied to 30 human leukocyte DNA samples from KTRs, this new UHPLC-MS/MS method enabled the detection of *N*3-meA, *N*7-meG, O^6-meG, M_1G, and CrG. N^2-etG and AcrG could not be detected in the samples. It was concluded that the new UHPLC-MS/MS method may serve as a complementary method for quantifying *N*3-meA and AcrG to the UHPLC-HRMS method (discussed in **Chapter 5**), which overall has better sensitivity. Due to the limited sample volume (for the ESCC cases and controls used in **Chapter 7**), no further UHPLC-MS/MS analysis was performed on the samples using this newly developed method.

Chapter 7 primarily aimed to screen DNA adducts in human leukocyte DNA samples obtained from ESCC cases and controls using the UHPLC-HRMS method (developed in **Chapter 5**). The samples were obtained from Tanzania and Kenya, as part of a large multi-country study on oesophageal squamous cell carcinoma African prevention research (ESCCAPE). Aside from the main objective (untargeted screening of DNA adducts), targeted analysis of oxidative and alkylating DNA adducts was also performed on the samples. However, due to administrative constraints, untargeted data could not be processed. Findings from the targeted analysis did not reveal a significant association between oxidative or alkylating DNA adducts and increased ESCC risk, except for N^6-meA (where an inverse association was observed). The biological reverse of N^6-meA remains to be validated in future studies.

Chapter 8 summarises the key findings (from **Chapters 1**-**7**) and draws general conclusions.

Chapter 9 describes the future research needed to investigate the role of the DNA adductome in oesophageal carcinogenesis. This field of research demands significant methodological,

technological, and conceptual advancements. Methodologically, to increase the statistical power of findings, future study designs should use large-scale, prospective cohort designs. These study designs must also address potential biases arising from uneven participant inclusion (e.g., age, sex, or country imbalances). Technological progress would benefit from the adoption of more sensitive and selective mass spectrometry techniques, such as multi-stage high-resolution mass spectrometry (HRMS/MS^n). Researchers should also validate DNA adduct profiles from surrogate tissues (such as blood, saliva, or urine) against those from the target tissue (i.e., blood vs oesophageal biopsy) to make informed conclusions in cases where access to or availability of target tissues is not always feasible (i.e., due to ethical or logistical reasons). Conceptually, DNA adductomics research needs to move beyond single-time-point measurements. It should assess the dynamic exposome at different key life stages to capture the multiplicity of the exposome. Ultimately, advances in DNA adductomics research and biomarker identification should lead to actionable public health measures that enable cost-effective screening and improved ESCC risk stratification, especially for high-risk populations in low-income settings.