UMEÅ UNIVERSITY

Molecular epidemiology approach:

Nested Case-Control Studies in Glioma and Lymphoid Malignancies

Florentin Späth

Department of Radiation Sciences, Oncology
Umeå 2019

Dissertation for PhD
ISBN: 978-91-7855-025-8
ISSN: 0346-6612
Cover figure drawn by Carl Wibom using the R software (www.r-project.org).
Electronic version available at: http://umu.diva-portal.org/
Printed by: CityPrint i Norr AB
Umeå, Sweden 2019

He will manage the cure best who has foreseen what is to happen from the present state of matters. (Hippocrates)

STUDY OVERVIEW

Paper	Disease	Reference	Design	Number Cases/Controls	Analyte	Main findings
I	Glioma	Späth et al., Tumour Biol. 2016	Nested case-control Prospective single samples (serum)	593/590	sEGFR sERBB2 12 SNPs	High blood levels of sEGFR and sERBB2 associated with risk of glioblastoma. High levels of sEGFR and sERBB2 were not associated with glioma risk SNPs.
II	B cell lymphoma (BCL)	Hosnijeh et al., Int J Cancer. 2016	Nested case-control and meta-analyses Prospective single samples (serum & plasma)	179/179 and 1452/1509 1389/1446	sCD27 sCD30	High blood levels of sCD27 and sCD30 associated with risk of BCL. High blood levels of sCD30 associated with all main BCL subtypes.
III	B cell lymphoma (BCL)	Späth et al., Cancer Res. 2017	Nested case-control Prospective serial samples (plasma)	170/170	CXCL13 sTNF-R1 sCD23 sCD27 sCD30	High blood levels of CXCL13, sCD23, sCD27, and sCD30 associated with BCL risk up to 25 years before diagnosis. Both baseline and marker slopes were associated with BCL risk.
IV	Multiple myeloma (MM)	Späth et al., Submitted 2019	Nested case-control Prospective serial samples (plasma)	65/65	MCP-3 MIP-1α MIP-1β VEGF FGF-2 Fractalkine TGF-α IL-13 TNF-α IL-10	Low blood levels of MCP-3, VEGF, FGF-2, fractalkine, and TGF-α associated with myeloma risk. These marker levels decreased among future myeloma patients over time. TGF-α can potentially improve early detection of MM.

TABLE OF CONTENTS

ABSTRACT

BACKGROUND: Nested case-control studies aim to link molecular markers with a certain outcome. Repeated prediagnostic samples may improve the evaluation of marker-disease associations. However, data regarding the benefit of repeated samples in such studies are sparse. We aimed to assess the relationship between blood levels of various proteins and risk of glioma, B cell lymphoma, and multiple myeloma to gain further understanding of disease etiology and to evaluate the clinical relevance of the studied markers. To this end, marker-disease associations were evaluated considering the natural history of the studied disease and the time between blood sample collection and diagnosis using both single (I-II) and repeated prediagnostic blood samples (III-IV).

PATIENTS AND METHODS: We conducted four nested case-control studies and one meta-analysis using samples from three prospective cohorts: the Janus Serum Bank, the Northern Sweden Health and Disease study, and the European Prospective Investigation into Cancer and Nutrition study. The following studied endpoints and relationships were included: I) glioma risk and the association with the receptor tyrosine kinases (soluble) sEGFR and sERBB2; II) B cell lymphoma risk and the association with the immune markers sCD27 and sCD30; III) B cell lymphoma risk and the association with immune markers (CXCL13, sTNF-R1, sCD23, sCD27, and sCD30) and their trends over time; and IV) multiple myeloma risk and the association with ten immune markers and growth factors (MCP-3, MIP-1α, MIP-1β, VEGF, FGF-2, fractalkine, TGF-α, IL-13, TNF-α, and IL-10) and their trends over time.

RESULTS: Risk of developing **I) glioma** was weakly associated with high blood levels of sERBB2. In addition, high levels of both sEGFR and sERBB2 assessed 15 years before diagnosis were associated with glioblastoma risk.

Risk of **II) B cell lymphoma** was associated with high levels of sCD30, whereas high levels of sCD27 were particularly associated with risk of chronic lymphocytic leukemia. Meta-analyses showed consistent results for sCD30 across cohorts and lymphoma subtypes, whereas results for sCD27 were less consistent across cohorts and subtypes.

In addition, **III) B cell lymphoma** risk was associated with levels of CXCL13, sCD23, sCD27, and sCD30 assessed in samples collected 17 years before diagnosis. Marker levels increased in cases closer to diagnosis, particularly for indolent lymphoma with a marked association for chronic lymphocytic leukemia and sCD23. Increasing marker levels closer to diagnosis were also observed for CXCL13 in future diffuse large B cell lymphoma patients.

Risk of **IV) multiple myeloma** was associated with low levels of MCP-3, VEGF, FGF-2, fractalkine, and TGF-α. Levels of these markers decreased in myeloma cases over time, especially for TGF-α. TGF-α assessed at time of the prediagnostic repeated sample seemed to help predict progression to multiple myeloma.

CONCLUSIONS: Both the natural history of the studied disease and the time between sample collection and diagnosis are crucial for the evaluation of marker-disease associations. Using repeated blood samples improves the understanding of marker-disease associations and might help to identify useful biomarker candidates.

SAMMANFATTNING PÅ SVENSKA

BAKGRUND: De sjukdomar som undersökts inom ramen för denna avhandling är gliom (hjärntumörer), samt lymfom (lymfkörtelcancer) och myelom (blodcancer som är närbesläktad med lymfom). Orsakerna till varför man får dessa sjukdomar är ofullständigt kartlagda. En del gemensamma samband har man dock kunnat se i epidemiologiska studier, till exempel att familjehistoria ger ökad sjukdomsrisk, att höga doser av joniserande strålning är kopplat med ökad sjukdomsrisk, samt att astma och allergi däremot ger minskad risk för både gliom och lymfom.

SYFTE: Syftet med denna avhandling var att undersöka sambandet mellan olika proteinnivåer i blod och risk för att senare i livet utveckla gliom, lymfom eller myelom för att bättre kunna förstå uppkomsten av dessa tumörsjukdomar. I två studier använde vi oss av upprepade blodprover tagna före diagnos, för att kunna studera förändringar i proteinnivåer över tid och samband med risk för sjukdom.

METODER OCH RESULTAT: Att mäta proteinnivåer i blod flera år innan diagnos möjliggjordes tack vare tillgång till frysta blodprover från prospektiva biobanker: NSHDS i Västerbotten, Janus i Norge och EPIC i Europa. Genom så kallade fall-kontroll-studier, d.v.s. jämförelser mellan prover från individer som senare drabbats av sjukdom (fall) och de som inte gjort det (kontroller), kan man hitta samband som eventuellt relaterar till sjukdomsuppkomst.

Risken att utveckla glioblastom, vilket är den vanligaste och mest aggressiva formen av gliom, fann vi kopplad till lätt ökade blodnivåer av sEGFR och sERBB2 15 år före diagnos. Dessa två lösliga proteiner tillhör gruppen tyrosinkinaser som är receptorer inblandade i tillväxtökande cellsignalering, vilket skulle kunna antyda att dessa proteiner är tidigt inblandade i uppkomsten av glioblastom.

Risken att utveckla lymfom fann vi kopplad till ökade nivåer av proteinerna CXCL13, sCD23, sCD27 och sCD30 i prover insamlade 15 till 25 år före diagnos. Dessa proteiner är inblandade i aktiveringen av immunsystemet och olika inflammatoriska tillstånd, och kallas därför immunmarkörer. Vi hittade även att koncentrationerna av dessa markörer ökade över tid bland framtida lymfomfall, medan de inte förändrades bland kontrollerna. Ökningen av protein-koncentrationer var särskilt stark för lymfomfall med kroniskt förlopp, vilket antyder att dessa markörer avspeglar sjukdomsaktivitet.

Risken att utveckla myelom var däremot kopplad till låga koncentrationer av andra immunmarkörer (MCP-3, VEGF, fractalkine, FGF-2 och TGF-α). Dessa markörer visade sig att ha sjunkande blod-koncentration över tid bland fallen jämfört med kontrollerna. Avsaknad av association mellan proteinkoncentrationer och sjukdomsrisk vid det första av två provtillfällen (ca 12 år innan diagnos), samt sjunkande markörnivåer bland fallen över tid talar för att dessa markörer också kan avspegla sjukdomsaktivitet.

SLUTSATSER: Sammantaget visar våra resultat att både karaktären av själva sjukdomen och tiden mellan provtagning och diagnos spelar roll i tolkningen av sambandet mellan sjukdom och markör. Upprepade prover före diagnos kan förbättra den tolkningen, och kan därför vara en särskilt lämplig studiedesign för att hitta biomarkörer som avspeglar sjukdomsaktivitet (dvs. förändringar i proteinkoncentrationer specifika för fall).

ABBREVIATIONS

AUC	Area under the curve
CD molecules	Cluster of differentiation
CLL	Chronic lymphocytic leukemia
CLUE I	Campaign Against Cancer and Stroke
CLUE II	Campaign Against Cancer and Heart Disease
CXCL13	C-X-C motif ligand 13
DLBCL	Diffuse large B cell lymphoma
DoDSR	Department of Defense Serum Repository
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EPIC	European Prospective Investigation into Cancer and Nutrition
ERBB2	ERB-B2 receptor tyrosine kinase 2
FGF-2	Fibroblast growth factor 2
Fractalkine	also known as chemokine C-X3-C motif ligand 1
IARC	International Agency for Research on Cancer
IDH	Isocitrate dehydrogenase
Ig	Immunoglobulin
IL-10	Interleukin 10
IL-13	Interleukin 13
Janus	Janus Serum Bank
MA	The Mammography screening project (sub cohort of NSHDS)
MCP-3	Monocyte chemotactic protein 3
MIP-1α	Macrophage inflammatory protein 1α
MIP-1β	Macrophage inflammatory protein 1β
MGUS	Monoclonal gammopathy of undetermined significance
MM	Multiple myeloma
MO	The MONICA project (sub cohort of NSHDS)
95% CI	95% Confidence Interval

NSHDS	Northern Sweden Health and Disease Study
OR	Odds ratio
PLCO	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial
REMARK	Reporting recommendations for tumour marker prognostic studies
ROC	Receiver operating characteristic
s	soluble
SEER	Surveillance, Epidemiology, and End Results Program
SMM	Smoldering multiple myeloma
SNP	Single nucleotide polymorphism
sTNF-R1	Soluble tumor necrosis factor receptor 1
STROBE-ME	Strengthening the reporting of observational studies in epidemiology – Molecular Epidemiology
TGF-α	Transforming growth factor α
TNF-α	Tumor necrosis factor α
VEGF	Vascular endothelial growth factor
VIP	Västerbotten Intervention Programme (sub cohort of NSHDS)

LIST OF INCLUDED PAPERS

I. **Späth F**, Andersson U, Dahlin AM, Langseth H, Hovig E, Johannesen TB, Grankvist K, Björkblom B, Wibom C, Melin B. "Pre-diagnostic serum levels of EGFR and ErbB2 and genetic glioma risk variants: a nested case-control study" *Tumour Biol.* 2016; 37(8): 11065-72.

II. Hosnijeh FS, Portengen L, **Späth F**, Bergdahl IA, Melin B, Mattiello A, Masala G, Sacerdote C, Naccarati A, Krogh V, Tumino R, Chadeau-Hyam M, Vineis P, Vermeulen R. "Soluble B-cell activation marker of sCD27 and sCD30 and future risk of B-cell lymphomas: A nested case-control study and meta-analyses" *Int J Cancer.* 2016; 138(10): 2357-67.

III. **Späth F**, Wibom C, Krop EJ, Johansson AS, Bergdahl IA, Vermeulen R, Melin B. "Biomarker Dynamics in B-cell Lymphoma: A Longitudinal Prospective Study of Plasma Samples Up to 25 Years before Diagnosis" *Cancer Res.* 2017; 77(6): 1408-1415.

IV. **Späth F**, Wibom C, Krop EJ, Izarra A, Johansson AS, Bergdahl IA, Hultdin J, Vermeulen R, Melin B. "Immune marker changes and risk of multiple myeloma: a nested case-control study using repeated prediagnostic blood samples" *Submitted* 2019.

AIMS OF THE THESIS

The overall aim of this thesis was to evaluate marker-disease associations with respect to the natural history of the disease, biological rationale of the marker-disease association, and time between blood sample collection and diagnosis using both single (Paper I-II) and repeated prediagnostic blood samples (Paper III-IV).

Specific aims:

I. To investigate prospective blood levels of sEGFR and sERBB2 and subsequent risk of glioma and to investigate whether perturbations in protein levels are associated with known genetic glioma risk variants.

II. To investigate the consistency of marker-disease associations for circulating sCD27 and sCD30 and future risk of lymphoma using a population from five prospective cohorts.

III. To investigate earlier reported immune markers - CXCL13, sCD23, sCD27, and sCD30 - in a prospective longitudinal cohort to assess individual dynamics of the marker-disease association to better understand their natural history in lymphomagenesis.

IV. To investigate earlier reported markers - MCP-3, VEGF, FGF-2, fractalkine, and TGF-α - in a prospective longitudinal cohort to assess marker dynamics for evaluation of their potential use as biomarkers in the follow-up of multiple myeloma precursors.

1 INTRODUCTION

1.1 COMMON CHARACTERISTICS AMONG GLIOMA AND LYMPHOID MALIGNANCIES

A robust relationship between atopic disease and reduced glioma risk has been reported (Amirian et al., 2016; Linos et al., 2007). Asthma and atopic conditions have been associated with reduced lymphoma risk (Vajdic et al., 2009; Yang et al., 2017). Data for myeloma are more heterogeneous but allergy medical history has been associated with reduced myeloma risk (Landgren et al., 2006). In addition, some studies indicate inverse associations between total IgE levels and glioma risk (Schlehofer et al., 2011; Wiemels et al., 2009) and risk of lymphoid malignancies, including myeloma and B cell lymphoma (BCL) (Nieters et al., 2014). These data could indicate that immune-related factors predispose to the development of lymphoid malignancies (Kristinsson et al., 2009) and glioma. Interestingly, different circulating markers with immunomodulatory effects and growth factors have been associated with risk of glioma (Brenner et al., 2018; Schwartzbaum et al., 2017), lymphoma (Edlefsen et al., 2014; Epstein et al., 2018), and myeloma (Birmann et al., 2012; Vermeulen et al., 2018).

1.2 GENERAL ASPECTS OF GLIOMA

Glioma is the most common primary brain tumor. According to the World Health Organization, these tumors can be subdivided based on microscopy findings and routine molecular parameters. Such molecular parameters, important for glioma classification and diagnosis, include isocitrate dehydrogenase (*IDH)* mutation and 1p/19q co-deletion (Huse & Aldape, 2014; Louis et al., 2016).

Common glioma subtypes include astrocytoma, oligodendroglioma, and glioblastoma (Louis et al., 2016). Five-year survival for glioblastoma, the most aggressive glioma subtype, is estimated to be below 5% (Ostrom et al., 2013). The incidence of glioma is estimated to be about six new cases per 100,000 inhabitants per year (Ho et al., 2014; Ostrom et al., 2013). Glioma is more common in males than in females with the largest difference in glioblastoma with a 60% higher incidence for males (Ostrom et al., 2018).

Previously, the only established glioma risk factors were high-dose radiation, family history, and some rare familial cancer syndromes (Bondy et al., 2008; Scheurer et al., 2010). To date, 27 genetic variants located among various chromosomal regions have been associated with increased risk of developing glioma (Melin et al., 2017; Rice et al., 2016). These findings provide evidence for an inherited genetic susceptibility to glioma (Kinnersley et al., 2015). Involved genes include *MDM4, AKT3, TERT, EGFR, LRIG1, CCDC26, CDKN2A*, *CDKN2B, PHLDB1, POL3B, TP53,* and *RTEL1.*

Discoveries involving acquired genomic variation have further enhanced the understanding of gliomagenesis. Some findings that might be involved in the key pathways driving glioma development include i) losses or alterations in tumor

suppressor genes *RB1* and *TP53* and ii) other oncogenic changes involved in gliomagenesis affect acquired mutations of receptor tyrosine kinases such as *EGFR*.

1.2.1 Treatment of glioma

Treatment goals for low-grade glioma, which is slower growing as compared to high-grade glioma, are to prolong overall survival and to minimize neurocognitive decline. These goals are achieved by early maximum safe surgical resection that eventually is followed by radiation and chemotherapy (Oberheim Bush & Chang, 2016). Similarly, surgical resection is also the initial therapeutic approach among high-grade glioma. Usually, surgery is followed by concomitant adjuvant chemotherapy and radiotherapy or adjuvant radiotherapy alone (Stupp et al., 2014).

1.3 GENERAL ASPECTS OF B CELL LYMPHOMA

Lymphoma represents a heterogeneous group of hematological malignancies arising from the lymphoid system. Traditionally, they have been divided into Hodgkin and non-Hodgkin lymphoma, accounting for about 10% and 90% of lymphoma, respectively (Swerdlow et al., 2016). For non-Hodgkin lymphoma, the vast majority of tumors (corresponding to more than 90%) are of B cell origin (Smith et al., 2015). These B cell lymphomas represent clonal proliferations of B lymphocytes that are sub-classified based on their maturity (Bennett et al., 1974). Main B cell lymphoma subtypes include aggressive subtypes such as diffuse large B cell lymphoma (DLBCL) and indolent subtypes such as follicular lymphoma and chronic lymphocytic leukemia (CLL). Five-year survival is heterogeneous among subtypes - about 80% for hairy cell leukemia and about 40% for mantle cell lymphoma based on data from the Swedish Lymphoma Registry during the period 2000 to 2005 (Dahle & Jerkeman, 2011; Juliusson, 2011). However, lymphoma survival improved markedly over time (Junlén et al., 2014; Mukhtar et al., 2018). Lymphoma incidence overall is estimated to be 19.4 per 100,000 individuals per year (SEER, 2015). Trends in lymphoma incidence have not been consistent. Lymphoma incidence increased in Europe and North America in the 1990s and stabilized thereafter (Armitage et al., 2017). Recent data indicate that trends in incidence might be heterogeneous for subtypes, with increasing incidence for indolent lymphoma (van de Schans et al., 2011).

Lymphoma etiology is incompletely understood. However, immunosuppression after transplant (van Leeuwen et al., 2009) and HIV infection, despite the introduction of antiretroviral therapy, are among the largest identified risk factors (Gibson et al., 2014). Other risk factors for developing lymphoma include preceding bacterial and viral infections (Anderson et al., 2014; Hjalgrim & Engels, 2008) and autoimmune diseases (Ekström Smedby et al., 2008; Zintzaras et al., 2005). In addition, inherited genetic susceptibility to lymphoma has been established based on observations of familial predisposition, and more than 65 inherited genetic lymphoma risk variants have been reported (Cerhan & Slager, 2015).

Physiological genetic processes that are central for B cell receptor diversity, referred to as V(D)J recombination, somatic hypermutation, and class switch recombination

predispose B lymphocytes to genetic instability, contributing to lymphoma development (Nogai et al., 2011). Lymphomagenesis is driven by genetic alterations, occurring during different steps of B cell differentiation, that allow the cells to escape from normal physiological restrictions related to growth, differentiation, proliferation, and cell death, resulting in acquired subtype-specific genetic lesions such as t(14;18) described for follicular lymphoma (Blombery et al., 2015).

1.3.1 Treatment of B cell lymphoma

Patients with low-bulk asymptomatic indolent lymphoma may initially be followed when no disease progression is evident (Gribben, 2007). If treatment is required, treatment options are similar to those used for treatment of aggressive lymphoma, including the use of chemo immunotherapy combining monoclonal antibodies with often anthracycline-based chemotherapeutic regimens (Zinzani et al., 2015). Other lymphoma treatment options may include novel biological agents (Chiappella et al., 2017) and radiotherapy (Yahalom, 2014).

1.4 GENERAL ASPECTS OF MULTIPLE MYELOMA

Multiple myeloma (MM), one of the most common hematologic malignancies, is characterized by clonal expansion of plasma cells (Swerdlow et al., 2016). All MM cases are preceded by monoclonal gammopathy of undetermined significance (MGUS) (Landgren et al., 2009; Weiss et al., 2009), a premalignant precursor, and smoldering multiple myeloma (SMM), which is characterized as an asymptomatic disease stage (Bianchi & Munshi, 2015). The risk of progression from precursor state to full-blown disease is highly heterogeneous (Zingone & Kuehl, 2011) but has been estimated to be approximately 1% and 10% per year for MGUS and SMM patients, respectively (Kyle et al., 2018; Kyle et al., 2007). Five-year overall survival for myeloma patients was recently shown to be 42%, although survival substantially depends on age at diagnosis (Blimark et al., 2018). Despite the observed survival improvement of myeloma patients diagnosed in more recent years (Costa et al., 2017), MM still is considered to be incurable (Ravi et al., 2018). Myeloma incidence has been described to be about 6.7 per 100,000 individuals per year (SEER, 2013). The age-adjusted incidence rate of MM seems to be stable while the total number of MM patients is increasing due to both aging populations and access to improved treatment options (Turesson et al., 2018).

Although the etiology of MM remains elusive, diversity of risk factors points to component causes of the disease (Sergentanis et al., 2015). Myeloma risk factors described include exposure to chemicals, pesticides, or unprecedented environmental risks. Other risk factors reported include overweight and obesity, African ancestry, family history, and inherited genetic polymorphisms (De Roos et al., 2018; Landgren et al., 2018; Marinac et al., 2018; Sergentanis et al., 2015; Smith et al., 2018).

MM is characterized by uncontrolled, destructive growth of mutated plasma cells within the bone marrow (Fairfield et al., 2016). Identical to lymphomagenesis, normal B cell development aiming to increase the quality of the immune response, predispose to malignant transformation and the acquisition of chromosomal translocations

involving immunoglobulin loci in MM (Boyle et al., 2014). In addition, MM is characterized by genetic and epigenetic changes involving a number of oncogenes and tumor-suppressor genes (Anderson & Carrasco, 2011). Many of these genetic events are common for both MM and its precursors such as the dysregulation of cyclin D (Kuehl & Bergsagel, 2012). Furthermore, several studies have shown that interplay between MM cells and the bone marrow microenvironment influences tumor growth and survival (Anderson & Carrasco, 2011).

1.4.1 Treatment of multiple myeloma

According to the most recent criteria of the international myeloma working group, treatment is initiated in patients with active myeloma, requiring evidence of end-organ damage attributable to the underlying plasma cell disease, or in smoldering myeloma patients who fulfill high risk criteria of progression (Rajkumar et al., 2014). The most common treatment options include newer medicines such as immunomodulatory drugs, proteasome inhibitors, and monoclonal antibodies. Other common therapeutic options include conventional chemotherapeutic regimens and for younger patients autologous stem cell transplantation (Gandolfi et al., 2018).

1.5 GENERAL REMARKS REGARDING BIOMARKERS

According to the National Institutes of Health (NIH), a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a given therapy (NIH, 2001). Based on this definition, biomarkers are classified into different categories: prognostic biomarkers estimate the likely outcome; predictive biomarkers indicate the expected response to a specific therapy; and pharmacodynamic biomarkers provide pharmacological data of a certain drug (Gainor et al., 2014; Goossens et al., 2015).

Figure 1. Biomarker development adapted from Goossens et al. (2015).

Biomarker development is an extensive and costly multiphase procedure beginning with discovery of a biomarker candidate followed by several qualification steps, including the evaluation of the analytical and clinical validity, and finally the prospective assessment of the clinical utility of a particular biomarker (Kraus, 2018).

Biomarker discovery can be performed using high-throughput technologies such as omics-based techniques - e.g., genomics, transcriptomics, proteomics, metabolomics, and microbiomics (Quezada et al., 2017). Material screened for biomarker candidates can be of various origin, although samples collected closest to the disease process itself have the highest probability of providing a direct biomarker (i.e., related to the outcome) (Kraus, 2018).

Analytic validity evaluates the technical aspects of biomarker assays themselves. Clinical validity evaluates the reliability of a biomarker to differentiate the overall population into diseased and normal subjects. Finally, biomarker candidates must be evaluated in prospective trials regarding their clinical utility, including an assessment of their effectiveness and benefit-to-harm ratio (Henry & Hayes, 2012).

1.5.1 Do we need biomarkers?

Biomarkers may increase our molecular understanding of oncogenic mechanisms contributing to tumorigenesis (Vasan, 2006). Precision medicine relies on the availability of biomarkers to better classify patients by their probable disease risk, prognosis, and/or response to treatment (Vargas & Harris, 2016). There are several examples of important molecular markers successfully incorporated in clinical oncology: i) *BCR-ABL* mutation which has been identified as therapeutic target for tyrosine kinase inhibition in chronic myeloid leukemia (Druker et al., 2001; Faderl et al., 1999); ii) the mutation of *EGFR*, which predicts response to targeted kinase inhibition in non-small cell lung cancer (Pao et al., 2004); and iii) the light chain of the immunoglobulin in the urine, which is one of the first reported biomarkers (Jones, 1847). Serum free light chains are significant for diagnosis and prognosis of plasma cell disorders (Ozsan & Dispenzieri, 2011).

To conclude, individualized treatment of cancer patients is associated with significant improvement of outcome in thousands of cancer patients (Fontes Jardim et al., 2015; Schwaederle et al., 2016; Schwaederle et al., 2015). Because this improvement depends on biomarker-guided decisions (Mordente et al., 2015), reliable biomarkers are an essential part of clinical oncology.

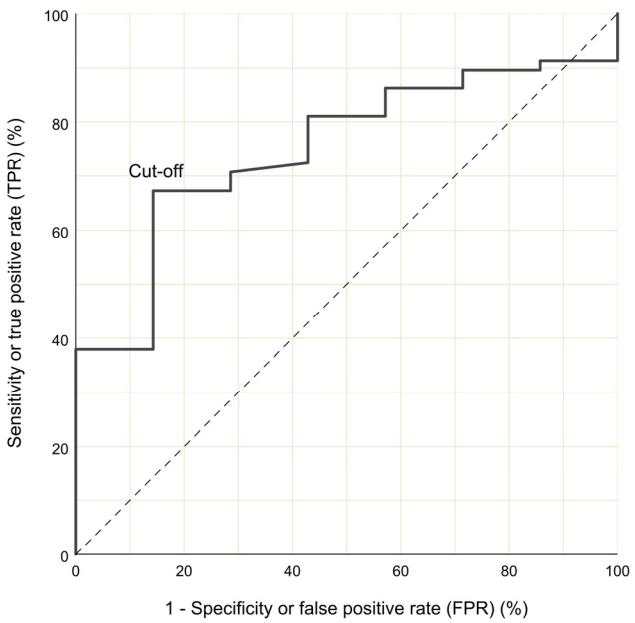
1.5.2 Sensitivity and specificity

Sensitivity and specificity are used as biomarker performance measures (Simon, 2015). Sensitivity, or the true positive rate, is defined by the proportion of individuals correctly assessed as diseased out of all individuals assessed as diseased. Specificity, or the true negative rate, is the proportion of individuals correctly assessed as healthy out of all individuals assessed as healthy. Thus, the false positive rate of a biomarker is calculated by $1 - specificity$. Ideal biomarkers should be both highly sensitive and specific (Yotsukura & Mamitsuka, 2015). Of note, the usefulness of a biomarker might increase as values for the false positive rate decrease and the prevalence of the outcome of interest increases (Pepe et al., 2016).

1.5.3 Receiver operating characteristic

The receiver operating characteristic (ROC) curve is a plot that illustrates the ability of a marker to predict a binary outcome among individuals (e.g. diseased vs. non-diseased) (Pepe et al., 2009). Figure 2 illustrates the ability of TGF-α to predict progression to multiple myeloma (MM) (58 individuals who progressed vs. 7 individuals who did not progress) (Paper IV). Ideal marker performance would be indicated by a marker rising steeply along the y-axis towards the upper left corner, yielding 100% sensitivity and 100% specificity. Uninformative marker performance is indicated by the dashed diagonal reference line.

Figure 2. ROC curve for TGF-α at the time of the prediagnostic repeated sampling (Paper IV).



The best discriminatory cut-off point for TGF-α measured at time of the prediagnostic repeated sampling is represented by the point on the solid line that is closest to the upper left corner (indicated as "cut-off", Figure 2). This cut-off value corresponds to a sensitivity of 67% and a specificity of 86%, implying a false negative and false positive rate of 33% and 14%, respectively.

Marker accuracy is measured by the area under the ROC curve (AUC); an AUC of one represents a perfect marker or test and an AUC of 0.5 represents a marker without any discriminatory potential. At prediagnostic repeated sampling, TGF-α had an AUC of 0.75 (95% CI 0.60-0.90) (Figure 2 and Paper IV).

1.6 NESTED CASE-CONTROL STUDIES

The nested case-control design is an efficient and well-established epidemiologic study design (Langholz & Richardson, 2009). The study design allows for the comparison of diseased and normal to determine the risk of disease (García-Closas, 2011; Rifai et al., 2006). Prospective cohorts, collecting biological specimens and life-style data from healthy participants, typically provide the basis of nested case-control studies. Participants of prospective cohorts need to be continuously followed to identify individuals developing the outcome of interest. Normally, a specified number of matched controls is selected among participants of the same cohort who did not develop the disease until the time of diagnosis of the incident case (Ernster, 1994).

Matching of controls to cases allows for the evaluation of the discriminatory performance of a marker while minimizing the influence of potential confounding variables (Pepe et al., 2001). However, effects of matching variables are not estimable, which may result in a biased calculation of the individual risk and overestimation of the discriminative power (Ganna et al., 2012). Therefore, matching strategies should be considered carefully in relation to the research question (Janes & Pepe, 2008). Usually applied matching factors in such studies comprise sex, age, and date of specimen collection (i.e., sample storage time).

Table 1. Common advantages and limitations of nested case-control studies (García-Closas, 2011; Sedgwick, 2014).

Advantages	Limitations
(Serial) specimens collected before outcome	Cohort size limits power (especially serial samples)
Transient and disease biomarkers can be studied	Potential reverse causation bias
Multiple disease endpoints can be investigated	Not all pertinent risk factors available
Selection bias minimized	Reliability on cancer registries (misclassification)

1.7 OVERVIEW OF INVESTIGATED ANALYTES

Paper	Analyte	Short description	Reference
I	sEGFR	Soluble form of EGFR, a RTK; RTKs involved in cell proliferation, differentiation, and survival; sEGFR could indicate disease activity in non-small cell lung cancer.	(Wieduwilt & Moasser, 2008) (Jorissen et al., 2003) (Gregorc et al., 2004)
	sERBB2	Soluble form of ERBB2, a RTK; RTKs are over-expressed in various tumors; high levels of sERBB2 could be a negative prognostic marker in HER2-positive breast cancer.	(Wieduwilt & Moasser, 2008) (Moreno-Aspitia et al., 2013)
II	sCD27	Soluble form of CD27, as part of TNF receptor family; CD27 is involved in immune activation; high levels of sCD27 have been associated with, e.g., HIV, SLE, and lymphoma.	(Maurer et al., 1990) (Nolte et al., 2009) (Ciccarelli et al., 2009)
	sCD30	Soluble form of CD30, as part of TNF receptor family; CD30 mediates its effects via NF-κB; sCD30 has been associated with autoimmune diseases and risk of (AIDS) lymphoma.	(Horie & Watanabe, 1998) (Caligaris-Cappio et al., 1995) (Breen et al., 2006)
III	CXCL13	Chemokine important for B cell homing; CXCL13 has been identified as marker of germinal center activity; levels of CXCL13 associated with several autoimmune diseases.	(Ansel et al., 2000) (Havenar-Daughton et al., 2016) (Nocturne et al., 2015)
	sTNF-R1	Soluble form of TNF-R1; binds specifically to TNF-α; high levels of TNF-Rs inhibit TNF-α activity; circulating sTNF-R1 was associated with total mortality.	(Aderka et al., 1992) (Carlsson et al., 2014)
	sCD23	CD23 is expressed on CLL; sCD23 stimulates differentiation of B-cells and Ig isotype switching; involved in bcl-2 expression (anti-apoptotic); associated with (AIDS) NHL risk.	(Cooper et al., 2012) (Brizard et al., 1994) (Martínez-Maza et al., 1998)
	sCD27	See above (Paper II).	
	sCD30	See above (Paper II).	

RTK, receptor tyrosine kinase; SLE, systemic lupus erythematosus; NF-κB, nuclear factor kappa light-chain-enhancer of activated B cells; Ig, immunoglobulin; Analyte overview to be continued on next page.

Paper	Analyte	Characteristic	Reference
IV	MCP-3	Pluripotent chemokine; ligand to various chemokine receptors; involved in normal and abnormal immune responses; both pro- and anti-tumor effects have been described.	(Sallusto et al., 1998) (Dempe et al., 2012) (Lee et al., 2016)
	MIP-1α	Macrophage inflammatory protein 1-alpha or CCL3 is a chemokine that enhances leukocyte infiltration to sites of inflammation; serum levels correlates with myeloma bone disease.	(Schaller et al., 2017) (Terpos et al., 2003)
	MIP-1β	Or CCL4 (see above); highly related to MIP-1α; MIP chemokines have inflammatory activities; inducible in hematopoietic cells; elevated in bone marrow of MGUS and myeloma patients.	(Menten et al., 2002) (Goodyear et al., 2017)
	VEGF	Vascular endothelial growth factor is crucial to angiogenesis; elevated serum levels are negative prognostic in many cancers including lymphoma.	(Matsumoto & Ema, 2014) (Salven et al., 2000)
	FGF-2	Fibroblast growth factor 2 signals via RTKs; serum FGF-2 is higher in MM as in controls; FGF-2 might be correlated to disease activity in MM but did not change during therapy.	(Adamis, 2010) (Greco et al., 2009) (Kyrtsonis et al., 2004)
	Fractalkine	Known as chemokine, C-X3-C motif ligand 1 also has a soluble form; CX3CL1 and its receptor (CX3CR1) acts in inflammation and development of B cell malignancies.	(Bazan et al., 1997) (Ferretti et al., 2014)
	TGF-α	Ligand to EGFR; produced by various cells; a small study suggested that TGF-α might be higher in bone marrow in MM but lower in blood as in healthy controls (n.s.).	(Lee et al., 1995) (Singh & Coffey, 2014) (Kara et al., 2006)
	IL-13	Both pro- and anti-inflammatory cytokine; associated with inflammatory disease such as asthma and autoimmune disease; IL-13 can inhibit tumor immunosurveillance.	(Huang et al., 2015) (Wynn, 2003)
	TNF-α	Tumor necrosis factor; inflammatory cytokine mediating apoptosis; high TNF-α was found in myeloma bone marrow; advanced myeloma stage was associated with serum TNF-α.	(Idriss & Naismith, 2000) (Berenson et al., 2000) (Jurišić & Čolović, 2002)
	IL-10	Anti-inflammatory cytokine can both, impede and improve pathogen clearance; high serum levels of IL-10 were associated with poor prognosis in myeloma patients.	(Couper et al., 2008) (Wang et al., 2016)

RTK, receptor tyrosine kinase; n.s., non-significant.

1.8 CHARACTERISTICS OF SELECTED ANALYTES

1.8.1 sEGFR, sERBB2, and TGF-α

The epidermal growth factor receptor (EGFR) belongs to a family of four receptor tyrosine kinases including EGFR (ERBB1, HER1), ERBB2 (HER2), ERBB3 (HER3), and ERBB4 (HER4) (Wieduwilt & Moasser, 2008). These transmembrane receptors are similar in structure and function and consist of an extracellular receptor domain, a transmembrane region, and an intracellular domain with tyrosine kinase function (Herbst, 2004). Receptor dimerization following ligand-binding leads to phosphorylation of the tyrosine kinase and activation of downstream signaling involved in cellular proliferation, differentiation, and survival (Jorissen et al., 2003). EGFR is expressed on the surface of multiple cell types (Hutchinson et al., 2015) and dysregulation or overexpression of EGFR family members is involved in several types of cancer, including glioblastoma (Yarden & Pines, 2012). Thus, EGFR is a reasonable therapeutic target in cancer (Dutta & Maity, 2007).

Soluble (s)EGFR and sERBB2 are proteolytically cleaved from the extra-cellular domain of the transmembrane receptor and released into the peripheral circulation (Choi et al., 1997; Codony-Servat et al., 1999). Interestingly, a decrease in sEGFR during treatment with tyrosine kinase inhibition has been associated with disease control in patients with non-small cell lung cancer (Gregorc et al., 2004). Similarly, increasing blood levels of sERBB2 were associated with shorter survival for HER2-positive breast cancer (Moreno-Aspitia et al., 2013), implying that both sEGFR and sERBB2 might reflect tumor activity.

Transforming growth factor alpha (TGF-α) is a transmembrane protein and growth factor belonging to the epidermal growth factor family. By binding its receptor (EGFR), TGF-α activates EGFR mediated signaling and cell proliferation. TGF-α is produced by various malignant and normal cells such as skin and hematopoietic cells (Lee et al., 1995). Overexpression of TGF-α has been associated with human cancer (Rusch et al., 1997). Proteolytically cleavage of the extracellular domain of TGF-α by metalloproteases leads to the release of the soluble form of the ligand that is able to bind and activate EGFR (Singh & Coffey, 2014).

1.8.2 CXCL13

The chemokine C-X-C motif ligand 13 (CXCL13) is critical for homing and motility of B cells in lymphoid tissue. Mice deficient in either CXCL13 or its receptor - CXCR5 - lack lymphoid structures such as lymph nodes or spleen (Ansel et al., 2000). CXCR5 is expressed by B cells, whereas its ligand CXCL13 is expressed by T follicular helper cells, dendritic cells, and stromal cells in secondary lymphoid organs (Ohmatsu et al., 2007). CXCL13 is important for the formation of ectopic lymphoid tissue and plays a crucial role for B cell migration to sites of inflammation (Rupprecht et al., 2009). The development of such ectopic lymphoid-like microenvironments has been referred to as tertiary lymphoid structures (Weinstein & Storkus, 2015), which can originate in any

non-lymphoid tissue exposed to chronic inflammation (Aloisi & Pujol-Borrell, 2006). The formation of tertiary lymphoid tissues is involved in various pathological conditions, including cancer, autoimmunity, and infection (Colbeck et al., 2017). Interestingly, CXCL13 has been identified as plasma biomarker of germinal center activity. Germinal center reaction occurring in lymph nodes is critical for generation of high affinity antibodies for improved immune response (Havenar-Daughton et al., 2016). CXCL13 blood levels are associated with several autoimmune diseases such as rheumatoid arthritis, Sjögren syndrome, and systemic lupus erythematosus (Nocturne et al., 2015). Further, pre-diagnostic levels of CXCL13 have been shown to be associated with future lymphoma risk (De Roos et al., 2012).

1.8.3 sCD23, sCD27, and sCD30

CD23 is a transmembrane protein that is upregulated on activated B cells and serves as low-affinity receptor for IgE. Binding of IgE to CD23 leads to negative regulation of IgE synthesis (Delespesse et al., 1992). Except the primarily expression of CD23 on B cells it is expressed on other cell types including certain lymphoma cells.

CD23 is cleaved from the B-cell membrane to release soluble CD23 (sCD23) which is able to promote survival and differentiation of B cells (Gordon, 1992). In addition, sCD23 induces IgE class switching and synthesis (Cooper et al., 2012; Perez-Witzke et al., 2016). A meta-analysis (including almost 1000 future B cell lymphoma patients) supports previous results showing an association between elevated circulating sCD23 and lymphoma risk in the general population (Huang et al., 2018).

CD27 is a transmembrane protein and a member of the tumor necrosis factor receptor superfamily expressed on lymphocytes (Maurer et al., 1990). CD27 acts as a co-stimulatory receptor on T and B cells in which the interaction between CD27 and its ligand CD70 plays a critical role in T cell dependent humoral immune response (Jacquot, 2000). Furthermore, CD27 signaling activates nuclear factor-kappa B, promoting cell survival and increasing effector function (Borst et al., 2005). Transient CD27-mediated co-stimulatory signaling might improve immune response. Sustained CD27-CD70 signaling, on the other hand, might result in immune dysregulation associated with chronic immune activation such as persistent virus infection and autoimmune disease (Nolte et al., 2009). Interestingly, CD27 is expressed in various neoplastic B cells, including CLL (Lens et al., 1998).

Soluble CD27 (sCD27) is proteolytically cleaved from cell surface CD27 after lymphocyte activation. Present in different body fluids, sCD27 can be used to monitor immune activation. Blood levels of sCD27 increase when the immune system is activated as the result of, for example, autoimmune disease or viral infection (Lens et al., 1998). sCD27 blood levels also increase in lymphoma patients, reflecting disease burden in B cell malignancies (Ciccarelli et al., 2009; van Oers et al., 1993).

CD30 is a transmembrane receptor and together with its ligand (CD30L) is a member of the tumor necrosis factor receptor superfamily. Expression of CD30 has been observed in activated T and B lymphocytes and in neoplasms of lymphoid origin, including Hodgkin lymphoma (Horie & Watanabe, 1998). CD30 mediates its effects through a

number of signaling pathways such as nuclear factor-kappa B, contributing to pro-survival benefits of cells with upregulated CD30 (van der Weyden et al., 2017).

Soluble CD30 (sCD30) is proteolytically cleaved from the extracellular domain of CD30 (Josimovic-Alasevic et al., 1989). Elevated blood levels of sCD30 have been observed in pathophysiological conditions, including HIV (Pizzolo et al., 1997) and autoimmune diseases such as systemic lupus erythematosus (Caligaris-Cappio et al., 1995), systemic sclerosis (Giacomelli et al., 1997), and ulcerative colitis (Giacomelli et al., 1998). Elevated levels of circulating sC30 have been associated with increased future lymphoma risk both among HIV positive individuals (Breen et al., 2006) and the general population (Purdue et al., 2009).

1.8.4 MCP-3

C-C motif ligand 7 or monocyte chemoattractant protein-3 (MCP-3) is a pluripotent chemokine belonging to the MCP subfamily (Proost et al., 1996). MCP-3 is produced by certain tumor cell lines and by macrophages (Opdenakker et al., 1993). Furthermore, it has been shown that the production of MCP-3 is enhanced in areas of inflammation (Wedemeyer et al., 1999). Chemokines play an important role in the control of leukocyte trafficking and are involved in normal and abnormal immune responses (Sallusto et al., 1998). For example, recruitment of monocytes to sites of inflammation is critical for host defense. Rodent models showed that MCP-3 is important for monocyte mobilization and recruitment to sites of inflammation (Tsou et al., 2007).

MCP-3 can interact with various chemokine receptors, including CCR1, CCR2, and CCR3. MCP-3 is chemotactic for and activates different types of leukocytes such as macrophages, T cells, NK cells, and basophils (Proost et al., 1996; Xu et al., 1995). In animal models, anti-tumor activity has been observed as an additional immunological effect of MCP-3 (Dempe et al., 2012; Fioretti et al., 1998). In addition to anti-tumor activity, pro-tumor effects of MCP-3 have been described in several cancers such as the MCP-3 mediated enhancement of metastasis in colorectal cancer (Lee et al., 2016).

1.8.5 VEGF and FGF-2

Vascular endothelial growth factors (VEGFs) and fibroblast growth factors (FGFs) are key regulators of angiogenesis in physiologic and pathologic conditions. VEGF (or VEGF-A) signaling is important for progression of angiogenesis-related diseases such as cancer (Matsumoto & Ema, 2014) including the development of myeloma (Vincent et al., 2005). Angiogenic stimulatory effects of VEGF are largely mediated by binding of VEGF to one of its receptors (VEGFR2). VEGF is produced by different cell types like fibroblasts, inflammatory cells, and tumor cells, often in response to increasing tumor hypoxia (Kieran et al., 2012).

Besides the previously mentioned pro-angiogenic function, FGF-2 and its receptor (FGFR) mediate mitogenic activity by stimulating the growth of fibroblasts, endothelial, and cancer cells. Therefore, FGF-2 is frequently dysregulated in cancer, particularly in advanced disease stages (Akl et al., 2016). Advanced myeloma stage has been associated with increasing blood concentrations of FGF-2 (Sezer et al., 2001).

2 PATIENTS AND METHODS

2.1 MOLECULAR EPIDEMIOLOGY STUDIES: THE BIOREPOSITORIES

In prospective cohort studies, exposure information such as life-style data and/or biological specimens of a defined population is used to assess exposure-outcome relations (Szklo, 1998). Establishing prospective cohorts is initially costly and time-consuming as large populations must be recruited and followed for a long period (Rothman, 2011). Nevertheless, prospective cohorts constitute an important investment in research infrastructure as they improve our understanding of biological processes that contribute to the development of complex diseases (Foster & Sharp, 2005). This thesis uses prospective cohorts that include more than 900,000 participants, providing a powerful tool for molecular epidemiology studies.

2.1.1 Janus Serum Bank

The Janus Serum Bank is a population-based biobank in Norway that includes 450,000 blood samples collected from 320,000 individuals. Blood samples preserved from Janus consist of serum containing small amounts of leukocytes due to incomplete separation (Langseth H, 2009). The large majority of Janus participants (about 90%) were recruited from individuals who participated in one or more of five health surveys carried out in different Norwegian counties between 1972 and 2003. About 10% of Janus participants were recruited from Red Cross blood donors between 1973 and 2004. Serially collected prospective blood samples are available in about one-fifth of Janus donors. The cohort is balanced regarding sex - 52% males and 48% females. The average age at sample donation was 41 years and the number of incident cancer cases in the Janus cohort was 69,221 in 2017 (Hjerkind et al., 2017; Langseth et al., 2017). Information on cancer diagnosis is provided by an annual linkage between Janus and the Cancer Registry of Norway (Langseth et al., 2017).

Limitations of the Janus cohort include sparse documentation and variation in blood sample processing, including the addition of iodoacetate to samples collected between 1972 and 1978 (Langseth et al., 2017). In addition, according to present biobank requirements, the blood sample's high storage temperature (-25°C) is suboptimal for long-term storage. This is of interest, as it has been shown previously that long-term storage influences blood protein concentrations (Enroth et al., 2016; Kugler et al., 2011). The strengths of the Janus cohort include the large number of participants, the long follow-up time, and the availability of longitudinally collected prospective blood samples (Langseth et al., 2017). In addition, the cancer registry's accuracy (94%) and completeness (99%) have recently been evaluated and reported to be among the highest, including all cancer sites, compared to other European registries (Larsen et al., 2009).

2.1.2 Northern Sweden Health and Disease Study (NSHDS)

Initiated in 1985, NSHDS is a prospective cohort designed for cardiovascular disease and cancer research. The cohort includes three sub-cohorts: the Västerbotten Intervention Programme (VIP), the Mammography Screening Project (MA), and the Northern Sweden Monica Project (MO). VIP, the largest of these sub-cohorts, continuously recruits individuals aged 40, 50, and 60 years old. At recruitment, all subjects provide informed consent and are asked to complete a questionnaire to assess medical and lifestyle data (Hallmans et al., 2003). For most participants, blood samples are collected after overnight fasting. These samples are collected in 10 mL tubes (1 EDTA and 1 heparin) and then processed according to routine protocols. In addition, samples are aliquoted into plasma, buffy coat, and red blood cells and are frozen within one hour from collection. Frozen samples are transported to the biobank at Umeå University Hospital within one week for long-term storage at -80°C. As of August 2018, the biobank includes around 230,000 blood samples donated by 135,000 individuals (NSHDS, 2018). The median age at blood sample collection within the total cohort is 50 years. The sex ratio among participants of the NSHDS is slightly in favor of women (53%) (personal communication with Robert Johansson, October 2018). Incident cancer cases occurring during follow-up can be identified with high accuracy and completeness by linkage with the Swedish Cancer Registry (Barlow et al., 2009; Pukkala et al., 2018).

NSHDS lacks cryopreserved samples. Strengths include high number of longitudinally collected samples, high blood sample quality, availability of buffy coat, and same standardized procedures applied for all samples collected during the study period. The long study period (> 33 years, VIP is still ongoing) allows for screening of substantial changes in a homogenous population.

2.1.3 European Prospective Investigation into Cancer and Nutrition (EPIC)

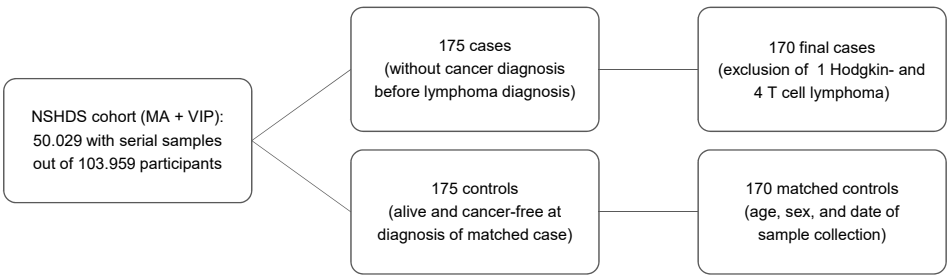
EPIC is a prospective multi-center cohort that includes 520,000 participants recruited from more than 25 European centers between 1992 and 1999 (IARC, 2018). The International Agency for Research on Cancer (IARC) initiated EPIC to investigate the relation between diet, nutritional characteristics, various lifestyle factors, and the risk of cancer (Riboli & Kaaks, 1997). Study participants were between 35 and 70 years old at enrollment. The participants provided information on dietary intake, lifestyle, and anthropometric data and a 30 mL blood sample was collected. Available blood samples include plasma, serum, buffy coat, and erythrocytes. Most of these samples are stored in liquid nitrogen at -196°C in the central EPIC repository at IARC. Identification of incident cancer cases varies by participating country and is based on either national cancer registries or on a combination of methods, including health insurance records, pathology registries, and active follow-up of study participants (Riboli et al., 2002).

One limitation of EPIC is the lack of repeated prospective samples. However, the main strengths are its multi-center character and the high number of enrolled participants, making EPIC one of the largest prospective cohorts worldwide (IARC, 2018).

2.2 STUDY DESIGN AND POWER CONSIDERATIONS

Paper I to IV were designed as nested case-control studies using both single (Paper I and II) and repeated prediagnostic blood samples (Paper III and IV). The process of case-control selection is illustrated for Paper III (Figure 3). For each cohort, specific study design considerations were discussed repeatedly with biostatisticians.

Figure 3. Case-control selection for Paper III (October 2013).



2.2.1 Considerations regarding statistical power

The power to detect differences in protein levels between cases and controls of small effect size (Cohen's d) was calculated to be high (> 95%) based on the large number of accessible case-control pairs for Paper I (N = 590). In Paper II and Paper III, sample sizes corresponded roughly to previously published studies reporting on the same marker-disease associations. Therefore, the power was considered sufficient for these studies. Based on findings observed by Vermeulen et al. (2018), we estimated a medium marker effect size for Paper IV. Thus, the power to detect differences between cases and controls was calculated to be above 90% despite the small number of available case-control pairs in this study (N = 65).

2.3 LABORATORY ANALYSES

All circulating proteins were measured by either enzyme-linked immunosorbent assays (ELISA) (Paper II and III) or multiplex-based immunoassays (Paper I, III, and IV). In ELISAs, the target protein or ligand is identified by a primary antibody that binds to the specific epitope of the target protein. Subsequently, the primary antibody (bound to the ligand) is detected by an enzyme-labeled secondary antibody. Finally, a colorimetric substrate is added, which quantifies the amount of captured ligand. As with this technique, Luminex immunoassays are based on an antigen-antibody interaction. Luminex allows the simultaneous detection and quantification of multiple target proteins using different antigen-specific capture and detection antibodies. Analyte quantification is then performed using a fluorescence reporter system. Laboratory analyses were performed by Ulrika Andersson at the oncology laboratory (Department of Radiation Sciences, Umeå University) using multiplex immunoassays (Paper I). Roel Vermeulen's laboratory (Institute for Risk Assessment Sciences, Utrecht University, The Netherlands) performed all other analyte assessments (Paper II-IV).

2.4 BIOSTATISTICAL CONSIDERATIONS

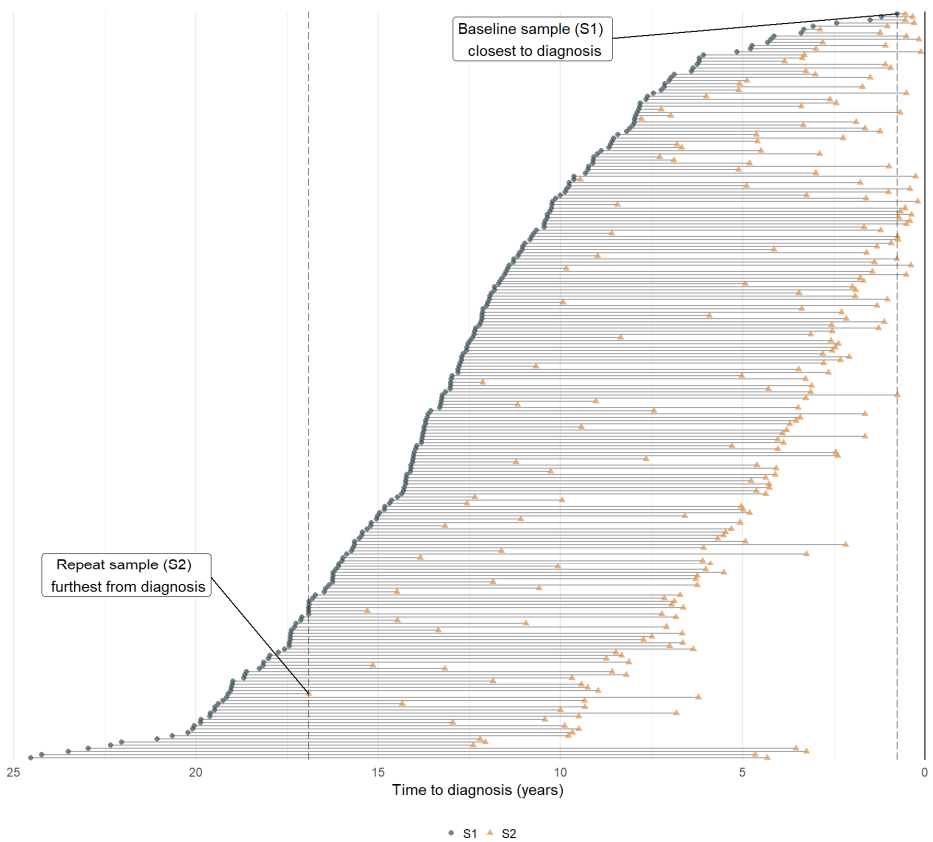
2.4.1 Linear mixed modeling

Used to assess linear associations, linear regression assumes independence among measures. When repeated measures from the same individuals are included in the dataset, the assumption of independence is violated. In this situation, it is preferable to assume different baseline levels for each individual. This may be achieved by applying linear mixed effects models, which adds a random effect for each individual (i.e. each individual is assigned a different intercept value). The model used in Paper III and Paper IV modeled protein levels according to (simplified):

$$Protein\ level \sim Case\text{-}control\ status\ +\ case\text{-}control\ status \times time\ +\ (1|individual)\ +\ (1|case\text{-}control\ pair)\ +\varepsilon$$

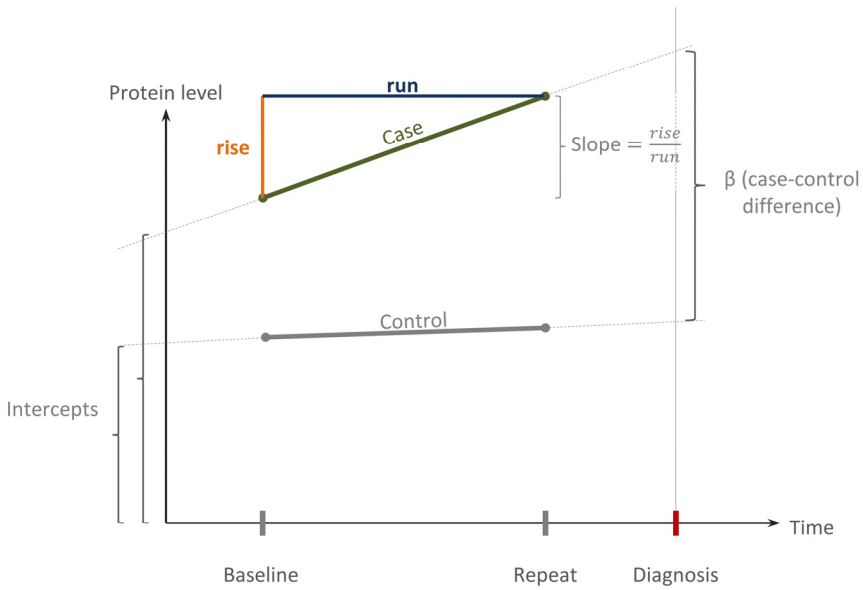
Fixed effects included were *case-control status* and the interaction term between *case-control status* and *time*. The t*ime* variable defined the time between sample collection and diagnosis (continuous and set to zero at the time of case diagnosis) (Figure 4).

Figure 4. Prediagnostic sampling among cases with future B cell malignancy (N = 235).



Random effects included in the present model were intercepts of each individual *(1|individual)* and each matched case-control pair *(1|case-control pair)*, accounting for both variation within and between individuals. The residual error (*ε*) describes the difference between measured protein levels and values predicted by the model (Winter, 2013).

Figure 5. Interpretation of the linear mixed effects model.



The β estimate for the case-control term is interpreted as described below:

a) β is assumed to be 0.25 (e.g., CXCL13 and B cell lymphoma risk, Paper III).
b) CXCL13 is on average 0.25 units higher for cases than for controls at time of diagnosis (i.e., *time* = 0), on a log10 transformed scale.
c) To re-transform the β estimate to the original scale calculate: $10^{0.25}$ = 1.78; $CXCL13_{Cases}/CXCL13_{Controls}$, thus corresponds to an increase in the CXCL13 level from controls to cases of 78% at time of the case diagnosis (Figure 5).

The β estimate for the *time* interaction term (i.e., slope) is interpreted similarly:

a) β for the interaction term in cases (*cases x time)* is assumed to be 0.011 (e.g., CXCL13 and B cell lymphoma, Paper III).
b) CXCL13 increases 0.011 units per time unit (i.e., year), on a log10 transformed scale.
c) To re-transform to the original scale calculate: $10^{0.011}$ = 1.026; this is corresponding to a 2.6% increase in CXCL13 level per year among B cell lymphoma cases (Figure 5).

2.4.2 Multiple imputation and missing data

Various strategies are employed to handle missing values for measured concentration data such as undetectable values below or above the limit of quantification. To ensure unbiased estimates, multiple imputation is a commonly recommended computational strategy to replace such missing values (Lubin et al., 2004).

Missing immune marker levels, as they were either unmeasurable low or high, were imputed (Paper III and IV). The finally selected imputation model included log transformed immune marker concentrations, case-control status, and analysis plate as previously recommended (Sterne et al., 2009). Estimations received by applying this model were rigorously evaluated by visual inspection of plotted marker abundance data. Furthermore, it was previously suggested to evaluate the influence of multiple imputation by excluding all imputed marker concentrations (Thabane et al., 2013). Performing such sensitivity analyses did not reveal a significant influence of multiple imputation on received results (Paper III and IV). However, because excluding imputed values will bias the results, we chose not to present such analyses in the papers.

3 RESULTS AND INTERPRETATION

3.1 sEGFR AND sERBB2 ASSOCIATED WITH GLIOBLASTOMA RISK (PAPER I)

Future risk of developing glioblastoma was associated with high levels of sEGFR and sERBB2 assessed in blood samples in median 15 years (range, 2 months - 35 years) before diagnosis (OR = 1.58, 95 % CI = 1.13-2.22, *P* = 0.008 and OR = 1.63, 95 % CI = 1.09-2.44, *P* = 0.017, respectively). High levels of sERBB2 were also associated with glioma risk overall (OR = 1.39, 95 % CI = 1.00-1.93, *P* = 0.049) (all ORs represent high vs. low protein levels). Linear regression analyses did not reveal obvious associations between time to diagnosis and levels of sEGFR and sERBB2.

Considering the overexpression of EGFR and ERBB2 described among different tumors including glioma (The Cancer Genome Atlas Research et al., 2008) and the potential oncogenic signaling effects of receptor tyrosine kinases (Lemmon & Schlessinger, 2010), these findings might be biologically plausible. Furthermore, significantly elevated blood levels of sEGFR have been found in several tumor types, including gastric cancer and ovarian cancer in situ, as compared to healthy individuals. This suggests that sEGFR could be used as an early diagnostic marker in ovarian cancer (Oh et al., 2000). The detection of measurable levels of sEGFR and sERBB2 among the majority of cancer-free controls in our cohort also suggests a physiological function of these proteins in normal conditions (Maramotti et al., 2016).

Of note, observed marker-disease associations were not confirmed when modeling protein concentrations on a continuous scale. Because dichotomizing continuous data may lead to loss of information and reduced power (Altman & Royston, 2006), it cannot be excluded that these results are false-positive.

3.1.1 Lack of genotype-phenotype association

It has been shown that a panel of single nucleotide polymorphisms (SNPs) identified in the *CRP* gene influences cancer susceptibility likely by inducing elevated levels of C-reactive protein (CRP) (Geng et al., 2016). Similarly, we hypothesized a causal relationship between genetic polymorphisms associated with glioma risk and elevated levels of sEGFR and sERBB2.

However, we did not find evidence for an association between high blood levels of sEGFR and sERBB2 and the genetic glioma risk variants known at this time.

3.2 sCD27 AND sCD30 ASSOCIATED WITH B CELL LYMPHOMA RISK (PAPER II)

For the nested case-control study, blood samples were collected in median 6.2 years (range, 1 - 16 years) before B cell lymphoma diagnosis. High levels of sCD30 were associated with B cell lymphoma risk (OR = 2.30, 95% CI = 1.23 – 4.30, P_{trend} = 0.002). Presented ORs represent estimates for the 4th vs. 1st quartile. Risk estimates for sCD27 and lymphoma risk were of similar size (OR = 2.27, 95% CI = 0.85 – 6.05, P_{trend} = 0.05). Subtype-specific analyses revealed that high levels of sCD27 seemed to be associated particularly with risk of chronic lymphocytic leukemia (CLL) (OR = 3.85, 95% CI = 0.97 – 15.29, P_{trend} = 0.02). In contrast, both sCD27 and sCD30 were not associated with myeloma risk.

Subsequently, we performed a meta-analysis including six case-control studies nested in five prospective cohorts to derive pooled estimates for lymphoma risk and levels of sCD27 and sCD30 (Bassig et al., 2015; De Roos et al., 2012; Hosnijeh et al., 2016; Purdue et al., 2011; Purdue et al., 2015; Purdue et al., 2009). Results for sCD30 were consistent across cohorts and lymphoma subtypes. In contrast, the data indicated some heterogeneity for results of sCD27 across considered studies and subtypes. The heterogeneity of results could partially be explained by different distributions of lymphoma subtypes ranging for CLL between 20% and 39% among these studies. In addition, the median time from sample collection to diagnosis ranged from 5 to 12 years for the included studies, potentially influencing the consistency of the results. This theory might be supported by previous observations that sCD27 is produced by various B cell malignancies, particularly by CLL and other indolent lymphoma subtypes (Ho et al., 2008; Molica et al., 1998; van Oers et al., 1993).

Evidence for an association between lymphoma risk and levels of sCD27 and sCD30 was initially provided by studies among HIV infected patients. Subsequently, the association between prediagnostic measures of sCD27 and sCD30 and risk of B cell lymphoma was also observed in the general population (Ambinder et al., 2010). Six previously mentioned and four additional nested case-control studies (Epstein et al., 2018; Purdue et al., 2018; Späth et al., 2017; Vermeulen et al., 2011), including more than 4000 participants, provide proof for this association in the general population. Paper III and one study by Purdue et al. (2015) found an association between high levels of sCD30 and lymphoma risk in 145 samples collected in median 17 years and 20 years before diagnosis, respectively.

The available evidence raises the possibility that elevated sCD30 might be a marker for increased B cell lymphoma susceptibility. Biological mechanisms contributing to this association are not yet clear, but it is plausible that increased immune activation is involved in lymphomagenesis. However, chronic B cell hyper-activation, potentially enhancing the generation of molecular lymphomagenic lesions, has been linked to HIV-related lymphoma (Epeldegui et al., 2006).

3.3 B CELL ACTIVATION MARKERS RISE CLOSER TO LYMPHOMA DIAGNOSIS (PAPER III)

We investigated trajectories of CXCL13, sTNF-R1, sCD23, sCD27, and sCD30 in 170 future lymphoma patients and matched controls. CXCL13, sCD23, sCD27, and sCD30 increased significantly in lymphoma patients closer to diagnosis, whereas markers remained stable in controls. The increase in the marker level between baseline and repeated sample collection was pronounced for indolent lymphoma subtypes, especially for CLL and sCD23 (~100% increase) and sCD27 (~60% increase). Of note, in diffuse large B cell lymphoma (DLBCL) – the main aggressive B cell lymphoma subtype - we observed a significant increase for CXCL13 between the baseline and repeated sample (~50% increase).

3.3.1 Levels at baseline and marker slopes associated with lymphoma risk

B cell lymphoma risk was associated with both immune marker concentration at baseline and change in the marker level (slope) for CXCL13, sCD23, sCD27, and sCD30.

Investigating baseline samples, collected on average 13 years prior diagnosis, CXCL13, sCD23, and sCD30 were significantly associated with risk of aggressive lymphoma. Sub-analyses of samples collected on average 18 years (all > 15 years) before aggressive lymphoma diagnosis revealed similar patterns, although only sCD30 reached significance (OR = 5.73, 95% CI = 1.01 – 32.50, P_{trend} = 0.015) (OR = 4^{th} vs. 1^{st} quartile).

The clinical course among B cell lymphoma subtypes is highly heterogeneous. Heterogeneity is also evident in DLBCL, one of the most common lymphoma subtypes, displaying a potential curable but aggressive disease course in most patients (Martelli et al., 2013). Therefore, high levels of sCD30, measured in samples collected almost two decades before aggressive lymphoma diagnosis, might support an etiologic involvement of sCD30 in the development of these tumors.

Increasing immune marker trajectories might be related to the process of disease progression. Lymphoma progression, in turn, might be driven in part by micro environmental-related factors (Cacciatore et al., 2012). Immune system environment, characterized by elevated levels of B cell stimulatory immune markers such as CXCL13, may contribute to lymphomagenesis by promoting DNA-modifying activities of B cells in the germinal center (Vendrame & Martínez-Maza, 2010). Interestingly, it was recently shown that germinal center activity is reflected by levels of CXCL13 (Havenar-Daughton et al., 2016). Therefore, increasing trajectories of CXCL13 observed in DLBCL could reflect increased germinal center activity related to lymphoma development (Teater et al., 2018).

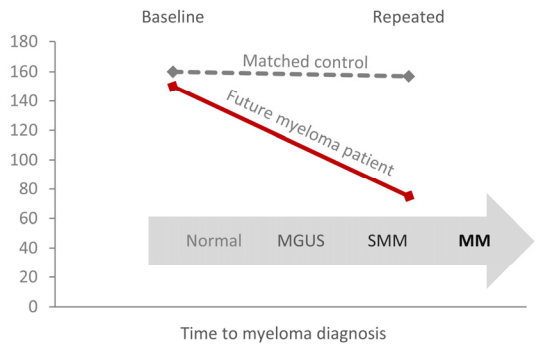
3.4 MCP-3, VEGF, FGF-2, FRACTALKINE, AND TGF-α AND MYELOMA RISK (PAPER IV)

Compared to controls, myeloma patients had significantly lower blood levels of MCP-3 (β = -0.129, *P* = 0.029), VEGF (β = -0.128, *P* = 0.021), FGF-2 (β = -0.101, *P* = 0.024), fractalkine (β = -0.090, *P* = 0.026), and TGF-α (β = -0.206, *P* = 0.029), findings that replicate results from a previous study (Vermeulen et al., 2018).

3.4.1 Marker trajectories indicative of progression to multiple myeloma

Levels of MCP-3, VEGF, FGF-2, fractalkine, and TGF-α decreased in myeloma patients closer to diagnosis, whereas plasma marker levels remained stable in controls. Observed trajectories might indicate myeloma progression (Figure 6).

Figure 6. Potential immune marker trajectory.



The potential to predict progression to MM seemed to be best for TGF-α assessed at time of the prediagnostic repeated sample (ROC). Future myeloma patients with low TGF-α (< 3.53 pg/mL) at this time-point had a shorter median time to MM progression (42 vs. 76 months, *P* = 0.004). This association remained significant in a multivariable model, which included known risk factors of progression (HR = 3.53, 95% CI = 1.54 – 8.10; *P* = 0.003).

Biological mechanisms promoting this decrease in marker level are unclear. Myeloma cells induce expression of several pro-angiogenic molecules in the bone marrow, including VEGF and FGF-2 (Giuliani et al., 2011). Therefore, the association between low levels of VEGF and FGF-2 and myeloma risk seems to be implausible. However, immune marker levels assessed in blood might not be synonymous for conditions met in the bone marrow. Support for our data is found in previous studies. For example, Tsirakis et al. found lower serum levels of TGF-β1, a member of the TGF family, in MM as compared to both controls and MGUS patients (Tsirakis et al., 2012). In addition, lower levels of sVEGFR2, the soluble receptor of VEGF, were observed in blood in MM in contrast to MGUS (Mailankody et al., 2017). Another small study performed by Kara and co-workers suggest lower blood levels of TGF-α in MM than observed in healthy controls, although these differences were not statistically significant (Kara et al., 2006).

4 DISCUSSION

4.1 DO MARKERS REFLECT CAUSALITY, BIOLOGICAL RESPONSE, OR MALIGNANCY?

More than 50 years ago, sir Austin Bradford Hill proposed a set of criteria that need to be fulfilled to provide evidence of causal inference (Hill, 1965). Since then, advancements in all scientific fields have improved the possibilities to evaluate exposure-disease relationships. This progress has led to better understanding of disease complexity, including the discovery of multi-factorial risk determinants of smaller magnitude crucial for disease development and progression (Fedak et al., 2015). Nevertheless, the Bradford Hill criterion of temporality (i.e., causal factors must precede the onset of disease) remains applicable (Rothman & Greenland, 2005). In this context, it is interesting that we observed several associations between marker levels and disease risk in samples collected about 15 years before diagnosis (Table 2).

Table 2. Time to diagnosis. Arrows indicate direction of marker-disease association.

Paper	Median time to diagnosis (range)	Disease (N)	Marker
I	15 years (2 months to 35 years)	Glioblastoma (N = 396)	sEGFR , sERBB2 (↑)
III	17 years (15 to 25 years)	B cell lymphoma (N = 59)	CXCL13, sCD23, sCD27, sCD30 (↑)
IV	14 years (12 to 24 years)	Myeloma (N = 37)	MCP-3 (↓)

The long time-span between sample collection and diagnosis of an aggressive disease such as glioblastoma, DLBCL, and MM makes it unlikely that observed associations can solely be referred to undiagnosed disease (Paper I, III, and IV). For indolent lymphoma subtypes, on the other hand, it might be highly plausible that undiagnosed disease exists up to decades before diagnosis. This assumption is supported by the observation that in CLL a high proportion of patients will never require therapy (i.e., remain asymptomatic) (Hallek et al., 2008). Low plasma levels of MCP-3 were associated with myeloma risk in samples collected 12 to 24 years before diagnosis. This observation was not evident in individuals who had no MGUS at this time-point. Considering the broad immunological effects of MCP-3, including anti-tumor activity (Dempe et al., 2012), low MCP-3 levels could indicate immune impairment involved in early myelomagenesis although without significance for MGUS initiation (Paper IV).

Dynamic immune marker trajectories (Paper III and IV) might be more likely to be reflective of biological processes related to tumorigenesis or the malignancy itself. This situation is particularly applicable for increasing measures of sCD23 and sCD27

observed in indolent lymphoma as both have been associated with tumor load in B cell malignancies (Reinisch et al., 1994; van Oers et al., 1993) (Paper III).

Decreasing levels of angiogenetic factors and growth factors closer to myeloma diagnosis (Paper IV) might be biologically counterintuitive. However, previous studies have observed similar results, such as decreasing blood levels of TGF-β1 among MM as compared to MGUS or controls (Tsirakis et al., 2012). These trajectories might be a measure of a biological response of the disease progression process without etiological involvement given the lack of an association longer before diagnosis (Paper IV).

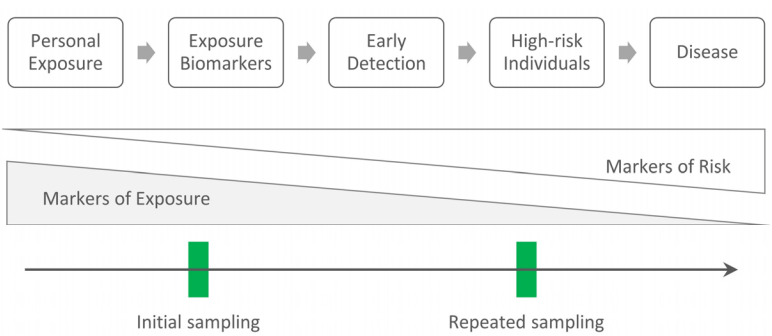
It is important to emphasize that these studies, given their observational character (Paper I to IV), cannot definitely distinguish between early diagnostic markers and true etiological factors.

4.2 REPEATED SAMPLES IN BIOMARKER STUDIES

4.2.1 How important is this design and what are potential limitations?

Biomarker analyses have been used in nested case-control and other observational studies to improve the understanding of risk factors and mechanisms leading to cancer (Perera, 1996). Prospective studies based on biological single samples per participant are limited in their interpretation of marker-disease associations, as they do not allow the study of within-individual variation and/or associations related to disease progression itself (Figure 7 and Paper III).

Figure 7. The molecular epidemiology paradigm shown in a prospective longitudinal study. Adapted from Groopman & Wang (2011).



The use of biomarkers associated with exposure, disease development, and clinical progression within the same overall design, known as “integrative-epidemiology” (Caporaso, 2007; Spitz et al., 2005), is of high interest. In this context, prospective longitudinal studies shed light on likely early (initiator) or later events (promotor) in tumor development (García-Closas, 2011). Therefore, longitudinal studies provide a comprehensive approach, allowing a better understanding of the degree and direction of change (Caruana et al., 2015). Information about marker changes allows improved

evaluation of the marker-disease association and processes related to disease progression itself.

Limitations of a longitudinal nested case-control design include reduced power due to the restricted number of individuals who have repeated samples available. In NSHDS and Janus, the proportion of participants with longitudinally banked samples is 48% and 22%, respectively (Langseth H, 2009; NSHDS, 2018). To achieve sufficiently powered sample sizes, pooling data across cohorts might be necessary (Kennedy et al., 2016). Despite improved insights on marker-disease associations provided by the use of repeated prediagnostic samples, it should be noted that such studies do not provide proof of underlying biological mechanisms related to marker-disease associations (Boyko, 2013).

4.2.2 How common are prospective longitudinal case-control studies?

Theoretically, prospective cohorts have the possibility to collect serial biological samples for each participant. Sample collection, however, often is limited to individual single banked samples due to logistical and financial constraints (Caruana et al., 2015; García-Closas, 2011). Despite these limitations, some prospective longitudinal cohorts are accessible that allow for the study of multiple disease endpoints (Table 3).

Table 3. Examples of studies that use serially collected prospective blood samples.

Cohort (Reference)	Finding of the study (example)	Study reference
CLUE I/CLUE II (CLUE, 1974 and 1989)	sCD27 is a disease marker in lymphoma, and sCD30 might be of etiologic relevance in lymphomagenesis.	(Purdue et al., 2018)
Janus (Langseth et al., 2017)	Serum levels of HE4 as an early detection biomarker in ovarian cancer.	(Gislefoss et al., 2015)
DoDSR (Rubertone & Brundage, 2002)	Higher levels of IL-15 and IL-16 were independently associated with lower glioma risk.	(Brenner et al., 2018)
PLCO (Prorok et al., 2000)	MGUS consistently precedes multiple myeloma.	(Landgren et al., 2009)

To the best of my knowledge, Purdue and colleagues performed the only study using repeated prediagnostic samples that is directly related to one of the studies included in this thesis (Purdue et al., 2018). The authors reported on repeated measures of sCD27 and sCD30 and future risk of developing lymphoma in the general population. In Purdue et al.'s study, time between sample collections and between repeated sampling and diagnosis were doubled compared to our study (15 and 10 years, respectively) (Paper III). However, results were similar in both studies with increasing marker levels closer to diagnosis among future lymphoma, although not reaching statistical significance for sCD30 in Purdue et al.'s study. Other longitudinal nested

case-control studies investigated associations between blood levels of CXCL13, sCD23, sCD27, and sCD30 and lymphoma risk among HIV infected individuals (Hussain et al., 2013; Yawetz et al., 1995).

Although several prospective longitudinal cohorts exist, some of these might be limited in their applicability for different reasons such as a low number of participants with repeated samples available (Manjer et al., 2001), a restricted number of potential study endpoints (Jacobs et al., 2016), and a short follow-up time (Scholtens et al., 2015).

4.3 WILL THE INVESTIGATED MARKERS BE APPLIED IN THE CLINIC?

Early diagnostic markers allowing timely surgical resection in glioma patients are needed (Hochberg et al., 2014). Estimated effect sizes for sEGFR and sERBB2 and risk of glioblastoma were too small to consider these markers clinically relevant (Paper I).

Considering the marked increase of immune markers in indolent lymphoma closer to diagnosis, particularly for sCD23 and CLL, using these markers to screen for indolent lymphoma might technically be possible. However, screening for early detection is suited for diseases with relatively high prevalence, accessible medical care, or the possibility of an intervention (Pavlou et al., 2013; Socialstyrelsen, 2014). None of these conditions are fulfilled for indolent lymphoma (for asymptomatic patients, the watch-and-wait strategy is preferred). It is therefore unlikely that these markers will be clinically implemented for screening of indolent lymphoma (Paper II and III).

Increasing measures of CXCL13 might reflect biological processes involved in lymphoma development and could be of potential use in early diagnosis and monitoring of DLBCL. Similar to our observations, CXCL13 has been found to be a marker for diagnosis of CNS lymphoma (Rubenstein et al., 2013) and post-transplant lymphoproliferative disorder (Schiffer et al., 2012). Both clinical validity and utility of CXCL13 as an early marker of DLBCL requires further investigation (Paper III).

A population-based randomized trial evaluating the value of MGUS screening and follow-up is ongoing in Iceland (Kristinsson, 2018). Regardless of the results of this study, TGF-α is an interesting biomarker candidate for early detection of MM. However, both sensitivity (67% as indicated by ROC analyses) and specificity (86% as indicated by ROC analyses) should be higher in stand-alone biomarkers (Paper IV).

4.4 LIMITATIONS OF BIOMARKER STUDIES

A brief literature search through PubMed using the search terms “cancer” and “biomarker” yielded 319,571 hits on this topic (November 2018). Only a few of these candidate biomarkers may be implemented in clinical routine (Khleif et al., 2010; Poste, 2011) as difficulties are associated with different stages of biomarker discovery and subsequent steps of biomarker evaluation (Diamandis, 2010). However, robust evidence and extensive evaluation are strictly required before biomarkers can be clinically implemented (Vasan, 2006). Other general biomarker requirements include a

favorable benefit-risk ratio, reasonable cost, acceptability, and convenience (Ioannidis & Tzoulaki, 2010).

One drawback of biomarker studies is the limited number of accessible biological specimens in prospective cohorts. In such situations, sample size considerations might be driven by the number of specimens available and not by design criteria or results of power analyses (McShane et al., 2005). Constraints in sample size particularly applies when using repeated samples, providing only limited power for performing sub analyses (Paper III and IV). Furthermore, biomarker studies are vulnerable to confounding factors (Perera & Weinstein, 2000). For example, the addition of iodoacetate or lyophilization to the sample influences serum protein concentrations (Paper I).

4.4.1 What may help to overcome limitations of biomarker studies?

Compared to meta-analyses, biomarker analyses often overestimate effect sizes of investigated markers (Ioannidis & Panagiotou, 2011; Turakhia & Sabatine, 2017). In this context, we observed heterogeneous results for sCD27 and lymphoma risk in the meta-analyses presented in Paper II, supporting the relevance of this type of investigation in the evaluation of marker-disease associations. Limitations in power due to sample size restrictions could be solved by establishing long-lasting collaborations between population-based cohorts, providing further opportunity to validate results in independent samples.

In addition, it is recommended to follow standards and guidelines that have been developed to advance biomarkers for clinical use. These guidelines include the prospective sample collection and retrospective blinded evaluation (PRoBE) design criteria (Pepe et al., 2008). Detailed recommendations to improve reporting of tumor marker prognostic studies (REMARK) (Altman et al., 2012) and observational studies in epidemiology – and molecular epidemiology (STROBE-ME) (Gallo et al., 2012) have been provided. Moreover, standards for reporting diagnostic accuracy (STARD) such as sensitivity, specificity, predictive values, or AUC are available (Cohen et al., 2016).

5 CONCLUSIONS

5.1 PAPER-SPECIFIC CONCLUSIONS

I. Considering both the clinical course of glioblastoma and the long lag time between sample collection and diagnosis, which was on average 15 years, high blood levels of sEGFR and sERBB2 likely reflect etiologic involvement more than reverse causation (Paper I).

II. The consistency of the marker-disease association for sCD30 across lymphoma subtypes and different cohorts supports an etiologic relationship between elevated circulating sCD30 and lymphomagenesis (Paper II).

III. B cell activation is a dynamic process in future lymphoma patients with differential trajectories across lymphoma subtypes, where both baseline and change in marker levels are associated with lymphoma risk (Paper III).

IV. Decreasing marker levels of MCP-3, FGF-2, fractalkine, and TGF-α indicate disease progression, considering the natural history of MM. In particular, TGF-α could improve prediction of progression to MM (Paper IV).

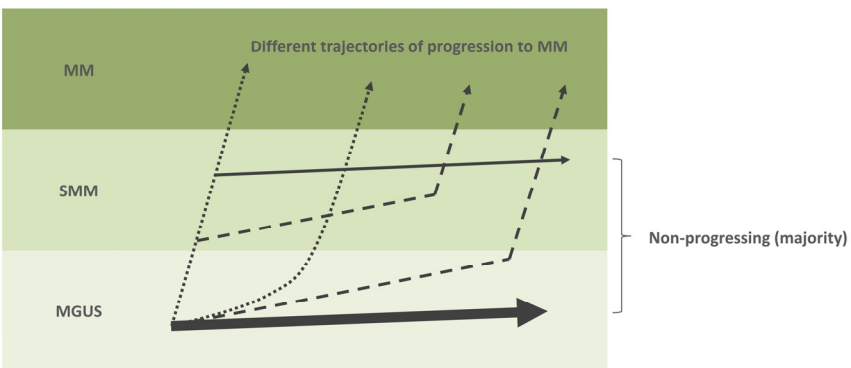
5.2 CONCLUDING REMARKS

The interpretation of marker-disease associations depends on the outcome of interest, the investigated marker itself, and the time between sample collection and diagnosis (Paper I-IV). Prospective longitudinal biomarker studies provide information about individual changes of marker levels that are not predictable by studies using single samples per participant. Changes in individual marker levels could be useful as biomarkers in clinical routine for screening, improved diagnostics and disease monitoring, and/or better prognostication (Paper II-IV).

6 FUTURE DIRECTIONS

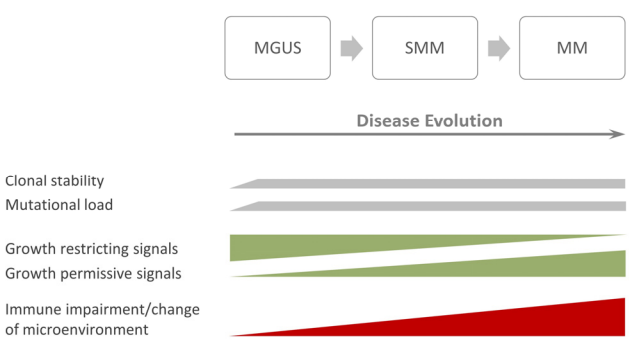
We have learned that reverse causation cannot be safely excluded in such molecular epidemiology studies. However, focusing on prospective longitudinal samples provides a good base to study disease progression. The lack of reliable biomarkers predicting which MGUS and SMM patients will progress to MM and which will not (Cosemans et al., 2018) provides the rationale to study this phenomenon further (Figure 8).

Figure 8. Patterns of disease evolution among MGUS patients. Adapted from Weiss and Kuehl (2010) and Bladé et al. (2008).



Over the next 30 years, the number of MGUS patients is expected to double (Go & Rajkumar, 2018) due to the increase in the population's age (Ferrucci et al., 2008). A prospective study in Iceland evaluates the value of screening and follow-up of MGUS (Kristinsson, 2018). Nevertheless, tailored follow-up strategies of MM precursors are of high socio-economic relevance (Go et al., 2015). Improved MGUS stratification could reduce anxiety in affected patients and could improve myeloma-related survival due to timely information on disease progression (Sigurdardottir et al., 2015). Interestingly, data by Dutta et al. (2018) suggest that cell-extrinsic factors such as the microenvironment play an important role in the development of MM (Figure 9).

Figure 9. The natural history of MM development and potentially influencing factors. Adapted from Dhodapkar (2016).



Prospective longitudinal cohorts (e.g., NSHDS and Janus, which includes repeated samples of almost 150,000 individuals) provide a powerful and cost-efficient possibility to identify and replicate biomarker trajectories that indicate disease progression. Based on this, we are preparing two follow-up studies to Paper IV that focus on factors related to myeloma progression.

- A follow-up study seeking independent replication of results from Paper IV is under way. The study uses 882 prospective longitudinal samples from 294 case-control pairs identified within the Janus cohort.

  When confirming marker trajectories observed in Paper IV, it would be of interest to study correlations between levels of these markers in bone marrow and peripheral blood. This could be addressed in the U-CAN cohort, a study collecting samples from both peripheral blood and marrow of same individuals including MGUS, SMM, and MM (Glimelius et al., 2017).

- Extensive analyses of the dynamics of circulating proteins and metabolites might hold the promise to reveal useful biomarker patterns for monitoring MM precursors as previously shown for other diseases using protein and metabolite profiling (Kulasingam et al., 2018; Weiner et al., 2018).

  To this end, 102 participants of NSHDS were identified who donated repeated samples before myeloma diagnosis (December 2017). Matched cancer-free controls and MGUS cases (with a 10-year median myeloma-free follow-up) are available from the same cohort. Potential marker trajectories could be replicated using samples from the Janus cohort.

Considering the availability of buffy-coat within NSHDS, other analytes of interest for investigating factors possibly related to MM development might include potential changes in transcriptomic markers (Chadeau-Hyam et al., 2014) or changes in genomic DNA methylation (Wu et al., 2012).

7 ACKNOWLEDGEMENT

Bea (Beatrice Melin), my supervisor. I really want to thank you for your generous support through all these years. I enjoyed spending this time in your group. Thank you for believing in me!

Calle (Carl Wibom), my co-supervisor. Thank you for generously sharing your thoughts and biostatistical knowledge. You are an incredible teacher and artist in R. Your figures have potential to light up dark winter days.

Roel Vermeulen and Esmeralda Krop, our collaborators at Utrecht University. Roel, thank you for your guidance through QC analyses, linear mixed modeling, and lots of the work behind Paper III and Paper IV. Esmeralda, thank you for teaching me basics in immune marker analyses.

Ulrika, Anna x 2, Benny, Barbro, Sara x 2, Micke, Wendy and all other group members of the “Melin group”. Thank you for sharing your thoughts, advice, and support.

Ingvar Bergdahl, my co-supervisor. Thank you for your support in particular during the initiation of my PhD and for your advice regarding the writing process of this thesis.

Ann-Sofie Johansson, my co-supervisor. Thank you for quickly replying regarding clinical questions related to this work.

Magnus Hultdin, the hematopathologist who always is accessible to questions. Thank you for helping me out regarding disease classification questions.

Furthermore, I want to thank all the individuals who donated blood samples to any of the used biobanks as without your help none of the presented works would have been possible. Also I would like to thank Robert Johansson and Åsa Ågren, both staff at NSHDS, for your technical support. Göran Hallmans your enthusiasm in building up NSHDS has been fundamental to this thesis.

Cecilia Isaksson, Angeliki Vourtsi, Fryderyk Lorenz, Berit Markevärn, Anna Thunström, Annika Burman, Staffan Ericson, Cecilia Björkvall, and Simona Sterbova my colleagues at the hematology department. Thank you for providing such a good working climate!

Antonio Izarra, my clinical supervisor. Thank you for your support and your enthusiasm talking about myeloma diagnostic criteria from almost 30 years ago.

Anders Wahlin, Karin Forsberg, and Maria Liljeholm, former and present bosses at the hematology department. Thank you for your support since day one.

Mum and Dad, thank you for your love and support.

Linnéa und Clara meine geliebten Töchter. Ich bin bewegt euch aufwachsen zu sehen. Danke dafür. Ich habe euch lieb bis zum Himmel und zurück.

Cornelia, love of my life, baseline and slope! ☺
Thank you for encouraging and supporting me since we met and in particular during the last freaking weeks. I love you.

8 REFERENCES

Adamis, A. P. (2010). CHAPTER 4 - Ocular angiogenesis: vascular endothelial growth factor and other factors. In Q. D. Nguyen, E. B. Rodrigues, M. E. Farah & W. F. Mieler (Eds.), *Retinal Pharmacotherapy*, 23-36. Edinburgh: W.B. Saunders.

Aderka, D., Engelmann, H., Maor, Y., Brakebusch, C., & Wallach, D. (1992). Stabilization of the bioactivity of tumor necrosis factor by its soluble receptors. *The Journal of Experimental Medicine,* 175, 323-29.

Akl, M. R., Nagpal, P., Ayoub, N. M., Tai, B., Prabhu, S. A., Capac, C. M., et al. (2016). Molecular and clinical significance of fibroblast growth factor 2 (FGF2 /bFGF) in malignancies of solid and hematological cancers for personalized therapies. *Oncotarget,* 7, 44735-62.

Aloisi, F., & Pujol-Borrell, R. (2006). Lymphoid neogenesis in chronic inflammatory diseases. *Nature Reviews Immunology,* 6, 205.

Altman, D. G., McShane, L. M., Sauerbrei, W., & Taube, S. E. (2012). Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. *PLoS medicine,* 9, e1001216-e16.

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ (Clinical research ed.),* 332, 1080-80.

Ambinder, R. F., Bhatia, K., Martinez-Maza, O., & Mitsuyasu, R. (2010). Cancer biomarkers in HIV patients. *Current opinion in HIV and AIDS,* 5, 531-37.

Amirian, E. S., Zhou, R., Wrensch, M. R., Olson, S. H., Scheurer, M. E., Il'yasova, D., et al. (2016). Approaching a Scientific Consensus on the Association between Allergies and Glioma Risk: A Report from the Glioma International Case-Control Study. *Cancer Epidemiology Biomarkers & Prevention,* 25, 282-90.

Anderson, K., & Carrasco, R. (2011). Pathogenesis of Myeloma. *Annual Review of Pathology: Mechanisms of Disease,* 6, 249-74.

Anderson, L. A., Atman, A. A., McShane, C. M., Titmarsh, G. J., Engels, E. A., & Koshiol, J. (2014). Common infection-related conditions and risk of lymphoid malignancies in older individuals. *British Journal of Cancer,* 110, 2796.

Ansel, K. M., Ngo, V. N., Hyman, P. L., Luther, S. A., Förster, R., Sedgwick, J. D., et al. (2000). A chemokine-driven positive feedback loop organizes lymphoid follicles. *Nature,* 406, 309.

Armitage, J. O., Gascoyne, R. D., Lunning, M. A., & Cavalli, F. (2017). Non-Hodgkin lymphoma. *The Lancet,* 390, 298-310.

Barlow, L., Westergren, K., Holmberg, L., & Talbäck, M. (2009). The completeness of the Swedish Cancer Register – a sample survey for year 1998. *Acta Oncologica,* 48, 27-33.

Bassig, B. A., Shu, X.-O., Koh, W.-P., Gao, Y.-T., Purdue, M. P., Butler, L. M., et al. (2015). Soluble levels of CD27 and CD30 are associated with risk of non-Hodgkin lymphoma in three Chinese prospective cohorts. *International Journal of Cancer,* 137, 2688-95.

Bazan, J. F., Bacon, K. B., Hardiman, G., Wang, W., Soo, K., et al. (1997). A new class of membrane-bound chemokine with a CX3C motif. *Nature,* 385, 640.

Bennett, M., Farrer-Brown, G., Henry, K., Jelliffe, A. M., Gerard-Marchant, R., Hamlin, I., et al. (1974). Classification of Non-Hdogkin's lymphomas. *The Lancet,* 304, 405-08.

Berenson, J. R., Sjak-Shie, N. N., & Vescio, R. A. (2000). The role of human and viral cytokines in the pathogenesis of multiple myeloma. *Seminars in Cancer Biology,* 10, 383-91.

Bianchi, G., & Munshi, N. C. (2015). Pathogenesis beyond the cancer clone(s) in multiple myeloma. *Blood,* 125, 3049-58.

Birmann, B. M., Neuhouser, M. L., Rosner, B., Albanes, D., Buring, J. E., Giles, G. G., et al. (2012). Prediagnosis biomarkers of insulin-like growth factor-1, insulin, and interleukin-6 dysregulation and multiple myeloma risk in the Multiple Myeloma Cohort Consortium. *Blood,* 120, 4929-37.

Bladé, J., Rosiñol, L., Cibeira, M. T., & de Larrea, C. F. (2008). Pathogenesis and progression of monoclonal gammopathy of undetermined significance. *Leukemia,* 22, 1651-57.

Blimark, C. H., Turesson, I., Genell, A., Ahlberg, L., Björkstrand, B., Carlson, K., et al. (2018). Outcome and survival of myeloma patients diagnosed 2008-2015. Real-world data on 4904 patients from the Swedish Myeloma Registry. *Haematologica,* 103, 506-13.

Blombery, P. A., Wall, M., & Seymour, J. F. (2015). The molecular pathogenesis of B-cell non-Hodgkin lymphoma. *European Journal of Haematology,* 95, 280-93.

Bondy, M. L., Scheurer, M. E., Malmer, B., Barnholtz-Sloan, J. S., Davis, F. G., Il'yasova, D., et al. (2008). Brain tumor epidemiology: Consensus from the Brain Tumor Epidemiology Consortium. *Cancer,* 113, 1953-68.

Borst, J., Hendriks, J., & Xiao, Y. (2005). CD27 and CD70 in T cell and B cell activation. *Current Opinion in Immunology,* 17, 275-81.

Boyko, E. J. (2013). Observational research--opportunities and limitations. *Journal of diabetes and its complications,* 27, 642-48.

Boyle, E. M., Davies, F. E., Leleu, X., & Morgan, G. J. (2014). Understanding the multiple biological aspects leading to myeloma. *Haematologica,* 99, 605-12.

Breen, E. C., Fatahi, S., Epeldegui, M., Boscardin, W. J., Detels, R., & Martínez-Maza, O. (2006). Elevated Serum Soluble CD30 Precedes the Development of AIDS-Associated Non-Hodgkin's B Cell Lymphoma. *Tumor Biology,* 27, 187-94.

Brenner, A. V., Inskip, P. D., Rusiecki, J., Rabkin, C. S., et al. (2018). Serially measured pre-diagnostic levels of serum cytokines and risk of brain cancer in active component military personnel. *British Journal of Cancer,* 119, 893-900.

Brizard, A., Morel, F., Lecron, J. C., Dreyfus, B., Brizard, F., Barra, A., et al. (1994). Proliferative Response of B Chronic Lymphocytic Leukemia Lymphocytes Stimulated with IL2 and Soluble CD23. *Leukemia & Lymphoma,* 14, 311-18.

Cacciatore, M., Guarnotta, C., Calvaruso, M., Sangaletti, S., Florena, A. M., Franco, V., et al. (2012). Microenvironment-centred dynamics in aggressive B-cell lymphomas. *Advances in hematology,* 2012.

Caligaris-Cappio, F., Bertero, M., Converso, M., Stacchini, A., et al. (1995). Circulating levels of soluble CD30, a marker of cells producing Th2-type cytokines, are increased in patients with systemic lupus erythematosus and correlate with disease activity. *Clinical and experimental rheumatology,* 13, 339-43.

Caporaso, N. E. (2007). Integrative Study Designs—Next Step in the Evolution of Molecular Epidemiology? *Cancer Epidemiology Biomarkers & Prevention,* 16, 365-66.

Carlsson, A. C., Juhlin, C. C., Larsson, T. E., Larsson, A., Ingelsson, E., Sundström, J., et al. (2014). Soluble tumor necrosis factor receptor 1 (sTNFR1) is associated with increased total mortality due to cancer and cardiovascular causes – Findings from two community based cohorts of elderly. *Atherosclerosis,* 237, 236-42.

Caruana, E. J., Roman, M., Hernández-Sánchez, J., & Solli, P. (2015). Longitudinal studies. *Journal of thoracic disease,* 7, E537-E40.

Cerhan, J. R., & Slager, S. L. (2015). Familial predisposition and genetic risk factors for lymphoma. *Blood,* 126, 2265-73.

Chadeau-Hyam, M., Vermeulen, R. C. H., Hebels, D. G. A. J., Castagné, R., Campanella, G., Portengen, L., et al. (2014). Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. *Annals of Oncology,* 25, 1065-72.

Chiappella, A., Santambrogio, E., Castellino, A., Nicolosi, M., & Vitolo, U. (2017). Integrating novel drugs to chemoimmunotherapy in diffuse large B-cell lymphoma. *Expert review of hematology,* 10, 697-705.

Choi, J.-H., Oh, J. Y., Ryu, S. K., Kim, S. J., Lee, N. Y., Kim, Y. S., et al. (1997). Detection of epidermal growth factor receptor in the serum of gastric carcinoma patients. *Cancer,* 79, 1879-83.

Ciccarelli, B. T., Yang, G., Hatjiharissi, E., Ioakimidis, L., Patterson, C. J., Manning, R. J., et al. (2009). Soluble CD27 Is a Faithful Marker of Disease Burden and Is Unaffected by the Rituximab-Induced IgM Flare, as Well as by Plasmapheresis, in Patients with Waldenström's Macroglobulinemia. *Clinical Lymphoma and Myeloma,* 9, 56-58.

CLUE. (1974 and 1989). Johns Hopkins University, Bloomberg School of Public Health. www.jhsph.edu/research/centers-and-institutes/george-w-comstock-center-for-public-health-research-and-prevention/clue_research_ activities.html

Codony-Servat, J., Albanell, J., Lopez-Talavera, J. C., Arribas, J., & Baselga, J. (1999). Cleavage of the HER2 Ectodomain Is a Pervanadate-activable Process That Is Inhibited by the Tissue Inhibitor of Metalloproteases-1 in Breast Cancer Cells. *Cancer Research,* 59, 1196-201.

Cohen, J. F., Korevaar, D. A., Altman, D. G., Bruns, D. E., Gatsonis, C. A., Hooft, L., et al. (2016). STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open,* 6.

Colbeck, E. J., Ager, A., Gallimore, A., & Jones, G. W. (2017). Tertiary Lymphoid Structures in Cancer: Drivers of Antitumor Immunity, Immunosuppression, or Bystander Sentinels in Disease? *Frontiers in immunology,* 8, 1830-30.

Cooper, A. M., Hobson, P. S., Jutton, M. R., Kao, M. W., Drung, B., Schmidt, B., et al. (2012). Soluble CD23 controls IgE synthesis and homeostasis in human B cells. *The Journal of Immunology*, 188, 3199-207.

Cooper, A. M., Hobson, P. S., Jutton, M. R., Kao, M. W., Drung, B., Schmidt, B., et al. (2012). Soluble CD23 controls IgE synthesis and homeostasis in human B cells. *Journal of immunology (Baltimore, Md. : 1950),* 188, 3199-207.

Cosemans, C., Oben, B., Arijs, I., Daniëls, A., Declercq, J., Vanhees, K., et al. (2018). Prognostic Biomarkers in the Progression From MGUS to Multiple Myeloma: A Systematic Review. *Clinical Lymphoma Myeloma and Leukemia,* 18, 235-48.

Costa, L. J., Brill, I. K., Omel, J., Godby, K., Kumar, S. K., & Brown, E. E. (2017). Recent trends in multiple myeloma incidence and survival by age, race, and ethnicity in the United States. *Blood advances,* 1, 282-87.

Couper, K. N., Blount, D. G., & Riley, E. M. (2008). IL-10: The Master Regulator of Immunity to Infection. *The Journal of Immunology,* 180, 5771-77.

Dahle, N., & Jerkeman, M. (2011). Marked improvement of overall survival in mantle cell lymphoma: a population based study from the Swedish Lymphoma Registry AU - Abrahamsson, Anna. *Leukemia & Lymphoma,* 52, 1929-35.

De Roos, A. J., Mirick, D. K., Edlefsen, K. L., LaCroix, A. Z., Kopecky, K. J., Madeleine, M. M., et al. (2012). Markers of B-Cell Activation in Relation to Risk of Non-Hodgkin Lymphoma. *Cancer Research,* 72, 4733-43.

De Roos, A. J., Spinelli, J., Brown, E. B., Atanackovic, D., Baris, D., Bernstein, L., et al. (2018). Pooled study of occupational exposure to aromatic hydrocarbon solvents and risk of multiple myeloma. *Occupational and Environmental Medicine,* 75, 798-806.

Delespesse, G., Sarfati, M., Wu, C., Fournier, S., & Letellier, M. (1992). The low-affinity receptor for IgE. *Immunological Reviews,* 125, 77-97.

Dempe, S., Lavie, M., Struyf, S., Bhat, R., Verbeke, H., Paschek, S., et al. (2012). Antitumoral activity of parvovirus-mediated IL-2 and MCP-3/CCL7 delivery into human pancreatic cancer: implication of leucocyte recruitment. *Cancer Immunology, Immunotherapy,* 61, 2113-23.

Dhodapkar, M. V. (2016). MGUS to myeloma: a mysterious gammopathy of underexplored significance. *Blood,* 128, 2599-606.

Diamandis, E. P. (2010). Cancer biomarkers: can we turn recent failures into success? *Journal of the National Cancer Institute,* 102, 1462-67.

Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., et al. (2001). Efficacy and Safety of a Specific Inhibitor of the BCR-ABL Tyrosine Kinase in Chronic Myeloid Leukemia. *New England Journal of Medicine,* 344, 1031-37.

Dutta, A. K., Fink, J. L., Grady, J. P., Morgan, G. J., Mullighan, C. G., To, L. B., et al. (2018). Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability. *Leukemia*, doi: 10.1038/s41375-018-0206-x.

Dutta, P. R., & Maity, A. (2007). Cellular responses to EGFR inhibitors and their relevance to cancer therapy. *Cancer letters,* 254, 165-77.

Edlefsen, K. L., Martínez-Maza, O., Madeleine, M. M., Magpantay, L., Mirick, D. K., Kopecky, K. J., et al. (2014). Cytokines in serum in relation to future non-Hodgkin lymphoma risk: Evidence for associations by histologic subtype. *International Journal of Cancer,* 135, 913-22.

Ekström Smedby, K., Vajdic, C. M., Falster, M., Engels, E. A., Martínez-Maza, O., et al. (2008). Autoimmune disorders and risk of non-Hodgkin lymphoma subtypes: a pooled analysis within the InterLymph Consortium. *Blood,* 111, 4029-38.

Enroth, S., Hallmans, G., Grankvist, K., & Gyllensten, U. (2016). Effects of Long-Term Storage Time and Original Sampling Month on Biobank Plasma Protein Concentrations. *EBioMedicine,* 12, 309-14.

Epeldegui, M., Widney, D. P., & Martínez-Maza, O. (2006). Pathogenesis of AIDS lymphoma: role of oncogenic viruses and B cell activation-associated molecular lesions. *Current Opinion in Oncology,* 18, 444-48.

Epstein, M. M., Rosner, B., Breen, E. C., Batista, J. L., Giovannucci, E. L., Magpantay, L., et al. (2018). Pre-diagnosis plasma immune markers and risk of non-Hodgkin lymphoma in two prospective cohort studies. *Haematologica,* 103, 1679-87.

Ernster, V. L. (1994). Nested Case-Control Studies. *Preventive Medicine,* 23, 587-90.

Faderl, S., Talpaz, M., Estrov, Z., O'Brien, S., Kurzrock, R., et al. (1999). The Biology of Chronic Myeloid Leukemia. *New England Journal of Medicine,* 341, 164-72.

Fairfield, H., Falank, C., Avery, L., & Reagan, M. R. (2016). Multiple myeloma in the marrow: pathogenesis and treatments. *Annals of the New York Academy of Sciences,* 1364, 32-51.

Fedak, K. M., Bernal, A., Capshaw & Gross, S. (2015). Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerging themes in epidemiology,* 12, 14-14.

Ferretti, E., Pistoia & Corcione, A. (2014). Role of fractalkine/CX3CL1 and its receptor in the pathogenesis of inflammatory and malignant diseases with emphasis on B cell malignancies. *Mediators of inflammation,* 2014, 480941-41.

Ferrucci, L., Giallauria, F., & Guralnik, J. M. (2008). Epidemiology of Aging. *Radiologic clinics of North America,* 46, 643-v.

Fioretti, F., Fradelizi, D., Stoppacciaro, A., Ramponi, S., et al. (1998). Reduced Tumorigenicity and Augmented Leukocyte Infiltration After Monocyte Chemotactic Protein-3 (MCP-3) Gene Transfer: Perivascular Accumulation of Dendritic Cells in Peritumoral Tissue and Neutrophil Recruitment Within the Tumor. *The Journal of Immunology,* 161, 342-46.

Fontes Jardim, D. L., Schwaederle, M., Wei, C., Lee, J. J., Hong, D. S., Eggermont, A. M., et al. (2015). Impact of a Biomarker-Based Strategy on Oncology Drug Development: A Meta-Analysis of Clinical Trials Leading to FDA Approval. *Journal of the National Cancer Institute,* 107, djv253-djv53.

Foster, M. W., & Sharp, R. R. (2005). Will investments in large-scale prospective cohorts and biobanks limit our ability to discover weaker, less common genetic and environmental contributors to complex diseases? *Environmental Health Perspectives,* 113, 119-22.

Gainor, J. F., Longo, D. L., & Chabner, B. A. (2014). Pharmacodynamic Biomarkers: Falling Short of the Mark? *Clinical Cancer Research,* 20, 2587-94.

Gallo, V., Egger, M., McCormack, V., Farmer, P., Ioannidis, J. A., Kirsch-Volders, M., et al. (2012). STrengthening the Reporting of OBservational studies in Epidemiology - Molecular Epidemiology (STROBE-ME): An extension of the STROBE statement. *European Journal of Clinical Investigation,* 42, 1-16.

Gandolfi, S., Prada, C. P., & Richardson, P. G. (2018). How I treat the young patient with multiple myeloma. *Blood,* 132, 1114-24.

Ganna, A., Reilly, M., de Faire, U., Pedersen, N., Magnusson, P., & Ingelsson, E. (2012). Risk Prediction Measures for Case-Cohort and Nested Case-Control Designs: An Application to Cardiovascular Disease. *American Journal of Epidemiology,* 175, 715-24.

García-Closas, M., Vermeulen, R., Cox, D., Lan, Q., Caporaso, NE. & Rothman N. (2011). Population-based study designs in molecular epidemiology. *IARC Scientific Publications,* 163, 241-59.

Geng, P., Sa, R., Li, J., Li, H., Liu, C., Liao, Y., et al. (2016). Genetic polymorphisms in C-reactive protein increase cancer susceptibility. *Scientific reports,* 6, 17161-61.

Giacomelli, R., Cipriani, P., Lattanzio, R., Di Franco, M., Locanto, M., Parzanese, I., et al. (1997). Circulating levels of soluble CD30 are increased in patients with systemic sclerosis (SSc) and correlate with serological and clinical features of the disease. *Clinical and experimental immunology,* 108, 42-46.

Giacomelli, R., Passacantando, A., Parzanese, I., Vernia, P., Klidara, N., et al. (1998). Serum levels of soluble CD30 are increased in ulcerative colitis (UC) but not in Crohn's disease (CD). *Clinical and experimental immunology,* 111, 532-35.

Gibson, T. M., Morton, L. M., Shiels, M. S., Clarke, C. A., & Engels, E. A. (2014). Risk of non-Hodgkin lymphoma subtypes in HIV-infected people during the HAART era: a population-based study. *AIDS (London, England),* 28, 2313-18.

Gislefoss, R. E., Langseth, H., Bolstad, N., Nustad, K., & Mørkrid, L. (2015). HE4 as an Early Detection Biomarker of Epithelial Ovarian Cancer: Investigations in Prediagnostic Specimens From the Janus Serumbank. *International Journal of Gynecological Cancer,* 25, 1608-15.

Giuliani, N., Storti, P., Bolzoni, M., Palma, B. D., & Bonomini, S. (2011). Angiogenesis and multiple myeloma. *Cancer microenvironment : official journal of the International Cancer Microenvironment Society,* 4, 325-37.

Glimelius, B., Melin, B., Enblad, G., Alafuzoff, I., Beskow, A., Ahlström, H., et al. (2017). U-CAN: a prospective longitudinal collection of biomaterials and clinical information from adult cancer patients in Sweden. *Acta Oncologica*, 1-8.

Go, R. S., & Rajkumar, S. V. (2018). How I manage monoclonal gammopathy of undetermined significance. *Blood,* 131, 163-73.

Go, R. S., Swanson, K. M., Sangaralingham, L. R., Habermann, E. B., & Shah, N. D. (2015). Clinical prevalence (diagnosed cases) of monoclonal gammopathy of undetermined significance in the US: estimating the burden on health care. *Leukemia,* 30, 1443.

Goodyear, O. C., Essex, S., Seetharam, A., Basu, S., Moss, P., & Pratt, G. (2017). Neoplastic plasma cells generate an inflammatory environment within bone marrow and markedly alter the distribution of T cells between lymphoid compartments. *Oncotarget,* 8, 30383-94.

Goossens, N., Nakagawa, S., Sun, X., & Hoshida, Y. (2015). Cancer biomarker discovery and validation. *Translational cancer research,* 4, 256-69.

Gordon, J. (1992). CD23 and B cell activation. *Clinical & Experimental Allergy,* 22, 199-204.

Greco, C., Vitelli, G., Vercillo, G., Vona, R., Giannarelli, D., Sperduti, I., et al. (2009). Reduction of serum IGF-I levels in patients affected with Monoclonal Gammopathies of undetermined significance or Multiple Myeloma. Comparison with bFGF, VEGF and K-ras gene mutation. *Journal of experimental & clinical cancer research : CR,* 28, 35-35.

Gregorc, V., Ceresoli, G. L., Floriani, I., Spreafico, A., Bencardino, K. B., Ludovini, V., et al. (2004). Effects of Gefitinib on Serum Epidermal Growth Factor Receptor and HER2 in Patients with Advanced Non-Small Cell Lung Cancer. *Clinical Cancer Research,* 10, 6006-12.

Gribben, J. G. (2007). How I treat indolent lymphoma. *Blood,* 109, 4617-26.

Groopman, JD. & Wang JS. (2011). Physical/chemical/immunologic analytical methods. *IARC Scientific Publications,* 163, 43-61.

Hallek, M., Cheson, B. D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Döhner, H., et al. (2008). Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood,* 111, 5446-56.

Hallmans, G., Ågren, Å., Johansson, G., Johansson, A., Stegmayr, B., Jansson, J.-H., et al. (2003). Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort- evaluation of risk factors and their interactions. *Scandinavian Journal of Public Health,* 31, 18-24.

Havenar-Daughton, C., Lindqvist, M., Heit, A., Wu, J. E., Reiss, S. M., Kendric, K., et al. (2016). CXCL13 is a plasma biomarker of germinal center activity. *Proceedings of the National Academy of Sciences of the United States of America,* 113, 2702-07.

Henry, N. L., & Hayes, D. F. (2012). Cancer biomarkers. *Molecular Oncology,* 6, 140-46.

Herbst, R. S. (2004). Review of epidermal growth factor receptor biology. *International Journal of Radiation Oncology*Biology*Physics,* 59, S21-S26.

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine,* 58, 295-300.

Hjalgrim, H., & Engels, E. A. (2008). Infectious aetiology of Hodgkin and non-Hodgkin lymphomas: a review of the epidemiological evidence. *Journal of Internal Medicine,* 264, 537-48.

Hjerkind, K. V., Gislefoss, R. E., Tretli, S., Nystad, W., Bjørge, T., Engeland, A., et al. (2017). Cohort Profile Update: The Janus Serum Bank Cohort in Norway. *International Journal of Epidemiology,* 46, 1101-02f.

Ho, A. W., Hatjiharissi, E., Ciccarelli, B. T., Branagan, A. R., Hunter, Z. R., Leleu, X., et al. (2008). CD27-CD70 interactions in the pathogenesis of Waldenstrom macroglobulinemia. *Blood,* 112, 4683-89.

Ho, V. K. Y., Reijneveld, J. C., Enting, R. H., Bienfait, H. P., Robe, P., Baumert, B. G., et al. (2014). Changing incidence and improved survival of gliomas. *European Journal of Cancer,* 50, 2309-18.

Hochberg, F. H., Atai, N. A., Gonda, D., Hughes, M. S., Mawejje, B., Balaj, L., et al. (2014). Glioma diagnostics and biomarkers: an ongoing challenge in the field of medicine and science. *Expert review of molecular diagnostics,* 14, 439-52.

Horie, R., & Watanabe, T. (1998). CD30: expression and function in health and disease. *Seminars in Immunology,* 10, 457-70.

Hosnijeh, F. S., Portengen, L., Späth, F., Bergdahl, I. A., Melin, B., Mattiello, A., et al. (2016). Soluble B-cell activation marker of sCD27 and sCD30 and future risk of B-cell lymphomas: A nested case-control study and meta-analyses. *International Journal of Cancer*, 138, 2357-67.

Huang, X.-L., Wang, Y.-J., Yan, J.-W., Wan, Y.-N., Chen, B., Li, B.-Z., et al. (2015). Role of anti-inflammatory cytokines IL-4 and IL-13 in systemic sclerosis. *Inflammation Research,* 64, 151-59.

Huang, Y.-s., Zhou, X., Yang, Z.-f., & Lv, Z.-t. (2018). Elevated soluble CD23 level indicates increased risk of B cell non-Hodgkin's lymphomas: evidence from a meta-analysis. *Annals of Hematology,* 97, 1317-25.

Huse, J. T., & Aldape, K. D. (2014). The Evolving Role of Molecular Markers in the Diagnosis and Management of Diffuse Glioma. *Clinical Cancer Research,* 20, 5601-11.

Hussain, S. K., Hessol, N. A., Levine, A. M., Breen, E. C., Anastos, K., Cohen, M., et al. (2013). Serum Biomarkers of Immune Activation and Subsequent Risk of Non-Hodgkin B-Cell Lymphoma among HIV-Infected Women. *Cancer Epidemiology Biomarkers & Prevention,* 22, 2084-93.

Hutchinson, R. A., Adams, R. A., McArt, D. G., Salto-Tellez, M., Jasani, B., & Hamilton, P. W. (2015). Epidermal growth factor receptor immunohistochemistry: new opportunities in metastatic colorectal cancer. *Journal of translational medicine,* 13, 217-17.

IARC. (2018). International Agency for Research on Cancer. http://epic.iarc.fr/

Idriss, H. T., & Naismith, J. H. (2000). TNFα and the TNF receptor superfamily: Structure-function relationship(s). *Microscopy Research and Technique,* 50, 184-95.

Ioannidis, J. A., & Panagiotou, O. A. (2011). Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. *JAMA,* 305, 2200-10.

Ioannidis, J. A., & Tzoulaki, I. (2010). What makes a good predictor? The evidence applied to coronary artery calcium score. *JAMA,* 303, 1646-47.

Jacobs, I. J., Menon, U., Ryan, A., Gentry-Maharaj, A., Burnell, M., Kalsi, J. K., et al. (2016). Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *The Lancet,* 387, 945-56.

Jacquot, S. (2000). CD27/CD70 interactions regulate T dependent B cell differentiation. *Immunologic Research,* 21, 23-30.

Janes, H., & Pepe, M. S. (2008). Matching in Studies of Classification Accuracy: Implications for Analysis, Efficiency, and Assessment of Incremental Value. *Biometrics,* 64, 1-9.

Jones, H. B. (1847). Papers on chemical pathology. *The Lancet,* 50, 511-14.

Jorissen, R. N., Walker, F., Pouliot, N., Garrett, T. P. J., Ward, C. W., & Burgess, A. W. (2003). Epidermal growth factor receptor: mechanisms of activation and signalling. *Experimental Cell Research,* 284, 31-53.

Josimovic-Alasevic, O., Dürkop, H., Schwarting, R., Backé, E., Stein, H., & Diamantstein, T. (1989). Ki-1 (CD30) antigen is released by Ki-1-positive tumor cells in vitro and in vivo. I. Partial characterization of soluble Ki-1 antigen and detection of the antigen in cell culture supernatants and in serum by an enzyme-linked immunosorbent assay. *European Journal of Immunology,* 19, 157-62.

Juliusson, G. & Samuelsson, H. (2011). Hairy cell leukemia: epidemiology, pharmacokinetics of cladribine, and long-term follow-up of subcutaneous therapy. *Leukemia & Lymphoma,* 52, 46-49.

Junlén, H. R., Peterson, S., Kimby, E., Lockmer, S., Lindén, O., Nilsson-Ehle, H., et al. (2014). Follicular lymphoma in Sweden: nationwide improved survival in the rituximab era, particularly in elderly women: a Swedish Lymphoma Registry Study. *Leukemia,* 29, 668.

Jurišić, V., & Čolović, M. (2002). Correlation of sera TNF-α with percentage of bone marrow plasma cells, LDH, β2-microglobulin, and clinical stage in multiple myeloma. *Medical Oncology,* 19, 133-39.

Kara, I. O., Sahin, B., Gunesacar, R., & Unsal, C. (2006). Clinical significance of hepatocyte growth factor, platelet-derived growth factor-AB, and transforming growth factor-α in bone marrow and peripheral blood of patients with multiple myeloma. *Advances in therapy,* 23, 635-45.

Kennedy, A. E., Khoury, M. J., Ioannidis, J. P. A., Brotzman, M., Miller, A., Lane, C., et al. (2016). The Cancer Epidemiology Descriptive Cohort Database: A Tool to Support Population-Based Interdisciplinary Research. *Cancer Epidemiology Biomarkers & Prevention,* 25, 1392-401.

Khleif, S. N., Doroshow, J. H., & Hait, W. N. (2010). AACR-FDA-NCI Cancer Biomarkers Collaborative Consensus Report: Advancing the Use of Biomarkers in Cancer Drug Development. *Clinical Cancer Research,* 16, 3299-318.

Kieran, M. W., Kalluri, R., & Cho, Y.-J. (2012). The VEGF pathway in cancer and disease: responses, resistance, and the path forward. *Cold Spring Harbor perspectives in medicine,* 2, a006593-a93.

Kinnersley, B., Labussière, M., Holroyd, A., Di Stefano, A.-L., Broderick, P., Vijayakrishnan, J., et al. (2015). Genome-wide association study identifies multiple susceptibility loci for glioma. *Nature communications,* 6, 8559-59.

Kraus, V. B. (2018). Biomarkers as drug development tools: discovery, validation, qualification and use. *Nature Reviews Rheumatology,* 14, 354-62.

Kristinsson, S. Y. (2018). iStopMM. https://www.blodskimun.is/login

Kristinsson, S. Y., Goldin, L. R., Björkholm, M., Koshiol, J., et al. (2009). Genetic and immune-related factors in the pathogenesis of lymphoproliferative and plasma cell malignancies. *Haematologica,* 94, 1581-89.

Kuehl, W. M., & Bergsagel, P. L. (2012). Molecular pathogenesis of multiple myeloma and its premalignant precursor. *The Journal of clinical investigation,* 122, 3456-63.

Kugler, K. G., Hackl, W. O., Mueller, L. A., Fiegl, H., Graber, A., & Pfeiffer, R. M. (2011). The Impact of Sample Storage Time on Estimates of Association in Biomarker Discovery Studies. *Journal of clinical bioinformatics,* 1, 9-9.

Kulasingam, A., Hvas, A.-M., Grove, E. L., Funck, K. L., & Kristensen, S. D. (2018). Detection of biomarkers using a novel proximity extension assay in patients with ST-elevation myocardial infarction. *Thrombosis Research,* 172, 21-28.

Kyle, R. A., Larson, D. R., Therneau, T. M., Dispenzieri, A., Kumar, S., Cerhan, J. R., et al. (2018). Long-Term Follow-up of Monoclonal Gammopathy of Undetermined Significance. *New England Journal of Medicine,* 378, 241-49.

Kyle , R. A., Remstein , E. D., Therneau , T. M., Dispenzieri , A., Kurtin , P. J., Hodnefield , J. M., et al. (2007). Clinical Course and Prognosis of Smoldering (Asymptomatic) Multiple Myeloma. *New England Journal of Medicine,* 356, 2582-90.

Kyrtsonis, M.-C., Vassilakopoulos, T. P., Siakantaris, M. P., Kokoris, S. I., Gribabis, D. A., Dimopoulou, M. N., et al. (2004). Serum syndecan-1, basic fibroblast growth factor and osteoprotegerin in myeloma patients at diagnosis and during the course of the disease. *European Journal of Haematology,* 72, 252-58.

Landgren, O., Kyle, R. A., Pfeiffer, R. M., Katzmann, J. A., Caporaso, N. E., Hayes, R. B., et al. (2009). Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. *Blood,* 113, 5412-17.

Landgren, O., Zeig-Owens, R., Giricz, O., & et al. (2018). Multiple myeloma and its precursor disease among firefighters exposed to the world trade center disaster. *JAMA Oncology,* 4, 821-27.

Landgren, O., Zhang, Y., Zahm, S. H., Inskip, P., Zheng, T., & Baris, D. (2006). Risk of Multiple Myeloma following Medication Use and Medical Conditions: A Case-

Control Study in Connecticut Women. *Cancer Epidemiology Biomarkers & Prevention,* 15, 2342-47.

Langholz, B., & Richardson, D. (2009). Are Nested Case-Control Studies Biased? *Epidemiology,* 20, 321-29.

Langseth, H., Gislefoss, R. E., Martinsen, J. I., Dillner, J., & Ursin, G. (2017). Cohort Profile: The Janus Serum Bank Cohort in Norway. *International Journal of Epidemiology,* 46, 403-04g.

Langseth H, G. R., Martinsen JI, Stornes A, Lauritzen M, Andersen Aa, Jellum E, Dillner J. (2009). The Janus Serum Bank - From sample collection to cancer research. Oslo: Cancer Registry of Norway.

Larsen, I. K., Småstuen, M., Johannesen, T. B., Langmark, F., Parkin, D. M., Bray, F., et al. (2009). Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. *European Journal of Cancer,* 45, 1218-31.

Lee, D. C., Fenton, S. E., Berkowitz, E. A., & Hissong, M. A. (1995). Transforming growth factor alpha: expression, regulation, and biological activities. *Pharmacological Reviews,* 47, 51-85.

Lee, Y. S., Kim, S.-Y., Song, S. J., Hong, H. K., Lee, Y., Oh, B. Y., et al. (2016). Crosstalk between CCL7 and CCR3 promotes metastasis of colon cancer cells via ERK-JNK signaling pathways. *Oncotarget,* 7, 36842-53.

Lemmon, M. A., & Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell,* 141, 1117-34.

Lens, S. M. A., Tesselaar, K., van Oers, M. H. J., & van Lier, R. A. W. (1998). Control of lymphocyte function through CD27–CD70 interactions. *Seminars in Immunology,* 10, 491-99.

Linos, E., Raine, T., Alonso, A., & Michaud, D. (2007). Atopy and Risk of Brain Tumors: A Meta-analysis. *Journal of the National Cancer Institute,* 99, 1544-50.

Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., et al. (2016). The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica,* 131, 803-20.

Lubin, J. H., Colt, J. S., Camann, D., Davis, S., Cerhan, J. R., Severson, R. K., et al. (2004). Epidemiologic Evaluation of Measurement Data in the Presence of Detection Limits. *Environmental Health Perspectives,* 112, 1691-96.

Mailankody, S., Devlin, S. M., Korde, N., Lendvai, N., Lesokhin, A., Landau, H., et al. (2017). Proteomic profiling in plasma cell disorders: a feasibility study. *Leukemia & Lymphoma,* 58, 1757-59.

Manjer, J., Carlsson, S., Elmstahl, S., Gullberg, B., Janzon, L., et al. (2001). The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in

participants and non-participants. *European Journal of Cancer Prevention,* 10, 489-99.

Maramotti, S., Paci, M., Manzotti, G., Rapicetta, C., Gugnoni, M., Galeone, C., et al. (2016). Soluble Epidermal Growth Factor Receptors (sEGFRs) in Cancer: Biological Aspects and Clinical Relevance. *International journal of molecular sciences,* 17, 593.

Marinac, C., Birmann, B., Lee, I. M., Rosner, B., Townsend, M., Giovannucci, E., et al. (2018). Body mass index throughout adulthood, physical activity, and risk of multiple myeloma: a prospective analysis in three large cohorts. *British Journal of Cancer,* 118, 1013-19.

Martelli, M., Ferreri, A. J. M., Agostinelli, C., Di Rocco, A., Pfreundschuh, M., & Pileri, S. A. (2013). Diffuse large B-cell lymphoma. *Critical Reviews in Oncology/Hematology,* 87, 146-71.

Martínez-Maza, O., Widney, D., van der Meijden, M., Knox, R., Echeverri, A., Breen, E. C., et al. (1998). Immune Dysfunction and the Pathogenesis of AIDS-associated non-Hodgkin's Lymphoma. *Memórias do Instituto Oswaldo Cruz,* 93, 373-81.

Matsumoto, K., & Ema, M. (2014). Roles of VEGF-A signalling in development, regeneration, and tumours. *The Journal of Biochemistry,* 156, 1-10.

Maurer, D., Holter, W., Majdic, O., Fischer, G. F., & Knapp, W. (1990). CD27 expression by a distinct subpopulation of human B lymphocytes. *European Journal of Immunology,* 20, 2679-84.

McShane, L. M., Altman, D. G., & Sauerbrei, W. (2005). Identification of Clinically Useful Cancer Prognostic Factors: What Are We Missing? *Journal of the National Cancer Institute,* 97, 1023-25.

Melin, B. S., Barnholtz-Sloan, J. S., Wrensch, M. R., Johansen, C., Il'yasova, D., Kinnersley, B., et al. (2017). Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nature genetics,* 49, 789-94.

Menten, P., Wuyts, A., & Van Damme, J. (2002). Macrophage inflammatory protein-1. *Cytokine & Growth Factor Reviews,* 13, 455-81.

Molica, S., Vitelli, G., Levato, D., Crispino, G., Dell'Olio, M., Dattilo, A., et al. (1998). CD27 in B-cell chronic lymphocytic leukemia. Cellular expression, serum release and correlation with other soluble molecules belonging to nerve growth factor receptors (NGFr) superfamily. *Haematologica,* 83, 398-402.

Mordente, A., Meucci, E., Martorana, G. E., & Silvestrini, A. (2015). Cancer Biomarkers Discovery and Validation: State of the Art, Problems and Future Perspectives. In R. Scatena (Ed.), *Advances in Cancer Biomarkers: From biochemistry to clinic for a critical revision*, 9-26. Dordrecht: Springer Netherlands.

Moreno-Aspitia, A., Hillman, D. W., Dyar, S. H., Tenner, K. S., Gralow, J., Kaufman, P. A., et al. (2013). Soluble human epidermal growth factor receptor 2 (HER2) levels in patients with HER2-positive breast cancer receiving chemotherapy with or without trastuzumab: results from North Central Cancer Treatment Group adjuvant trial N9831. *Cancer,* 119, 2675-82.

Mukhtar, F., Boffetta, P., Dabo, B., Park, J. Y., Tran, C. T. D. et al. (2018). Disparities by race, age, and sex in the improvement of survival for lymphoma: Findings from a population-based study. *PLoS ONE,* 13, e0199745-e45.

Nieters, A., Łuczyńska, A., Becker, S., Becker, N. et al. (2014). Prediagnostic immunoglobulin E levels and risk of chronic lymphocytic leukemia, other lymphomas and multiple myeloma-results of the European Prospective Investigation into Cancer and Nutrition. *Carcinogenesis,* 35, 2716-22.

NIH. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics,* 69, 89-95.

Nocturne, G., Seror, R., Fogel, O., Belkhir, R., Boudaoud, S., Saraux, A., et al. (2015). CXCL13 and CCL11 Serum Levels and Lymphoma and Disease Activity in Primary Sjögren's Syndrome. *Arthritis & Rheumatology,* 67, 3226-33.

Nogai, H., Dörken, B., & Lenz, G. (2011). Pathogenesis of Non-Hodgkin's Lymphoma. *Journal of Clinical Oncology,* 29, 1803-11.

Nolte, M., van Olffen, R., van Gisbergen, K. J. M., & van Lier, R. W. (2009). Timing and tuning of CD27-CD70 interactions: the impact of signal strength in setting the balance between adaptive responses and immunopathology. *Immunological Reviews,* 229, 216-31.

NSHDS. (2018). Northern Sweden Health and Disease Study. https://www.umu.se/en/department-of-biobank-research/

Oberheim Bush, N., & Chang, S. (2016). Treatment Strategies for Low-Grade Glioma in Adults. *Journal of Oncology Practice,* 12, 1235-41.

Oh, M.-J., Choi, J.-H., Kim, I. H., Lee, Y. H., Huh, J. Y., Park, Y. K., et al. (2000). Detection of Epidermal Growth Factor Receptor in the Serum of Patients with Cervical Carcinoma. *Clinical Cancer Research,* 6, 4760-63.

Ohmatsu, H., Sugaya, M., Kadono, T., & Tamaki, K. (2007). CXCL13 and CCL21 Are Expressed in Ectopic Lymphoid Follicles in Cutaneous Lymphoproliferative Disorders. *Journal of Investigative Dermatology,* 127, 2466-68.

Opdenakker, G., Froyen, G., Fiten, P., Proost, P., & Vandamme, J. (1993). Human Monocyte Chemotactic Protein-3 (MCP-3): Molecular Cloning of the cDNA and Comparison with Other Chemokines. *Biochemical and Biophysical Research Communications,* 191, 535-42.

Ostrom, Q. T., Cote, D. J., Ascha, M., Kruchko, C., & Barnholtz-Sloan, J. S. (2018). Adult glioma incidence and survival by race or ethnicity in the united states from 2000 to 2014. *JAMA Oncology,* 4, 1254-62.

Ostrom, Q. T., Gittleman, H., Farah, P., Ondracek, A., Chen, Y., Wolinsky, Y., et al. (2013). CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro-Oncology,* 15 Suppl 2, ii1-ii56.

Ozsan, G. H., & Dispenzieri, A. (2011). Serum free light chain analysis in multiple myeloma and plasma cell dyscrasias. *Expert Review of Clinical Immunology,* 7, 65-73.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., et al. (2004). EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proceedings of the National Academy of Sciences of the United States of America,* 101, 13306-11.

Pavlou, M. P., Diamandis, E. P., & Blasutig, I. M. (2013). The Long Journey of Cancer Biomarkers from the Bench to the Clinic. *Clinical Chemistry,* 59, 147-57.

Pepe, M., Longton, G., & Janes, H. (2009). Estimation and Comparison of Receiver Operating Characteristic Curves. *The Stata journal,* 9, 1-1.

Pepe, M. S., Etzioni, R., Feng, Z., Potter, J. D., Thompson, M. L., Thornquist, M., et al. (2001). Phases of Biomarker Development for Early Detection of Cancer. *Journal of the National Cancer Institute,* 93, 1054-61.

Pepe, M. S., Feng, Z., Janes, H., Bossuyt, P. M., & Potter, J. D. (2008). Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *Journal of the National Cancer Institute,* 100, 1432-38.

Pepe, M. S., Janes, H., Li, C. I., Bossuyt, P. M., Feng, Z., & Hilden, J. (2016). Early-Phase Studies of Biomarkers: What Target Sensitivity and Specificity Values Might Confer Clinical Utility? *Clinical Chemistry,* 62, 737-42.

Perera, F. P. (1996). Molecular Epidemiology: Insights Into Cancer Susceptibility, Risk Assessment, and Prevention. *Journal of the National Cancer Institute,* 88, 496-509.

Perera, F. P., & Weinstein, I. B. (2000). Molecular epidemiology: recent advances and future directions. *Carcinogenesis,* 21, 517-24.

Perez-Witzke, D., Miranda-García, M. A., Suárez, N., Becerra, R., Duque, K., Porras, V., et al. (2016). CTLA4Fcε, a novel soluble fusion protein that binds B7 molecules and the IgE receptors, and reduces human in vitro sCD23 production and lymphocyte proliferation. *Immunology*, 148, 40-55.

Pizzolo, G., Vinante, F., Nadali, G., Krampera, M., Morosato, L., Chilosi, M., et al. (1997). High serum level of soluble CD30 in acute primary HIV-1 infection. *Clinical & Experimental Immunology,* 108, 251-53.

Poste, G. (2011). Bring on the biomarkers. *Nature,* 469, 156-57.

Proost, P., Wuyts, A., & van Damme, J. (1996). Human monocyte chemotactic proteins-2 and -3: structural and functional comparison with MCP-1. *Journal of Leukocyte Biology,* 59, 67-74.

Prorok, P. C., Andriole, G. L., Bresalier, R. S., Buys, S. S., Chia, D., David Crawford, E., et al. (2000). Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Controlled Clinical Trials,* 21, 273S-309S.

Pukkala, E., Engholm, G., Højsgaard Schmidt, L. K., Storm, H., Khan, S., Lambe, M., et al. (2018). Nordic Cancer Registries – an overview of their procedures and data comparability. *Acta Oncologica,* 57, 440-55.

Purdue, M. P., Lan, Q., Bagni, R., Hocking, W. G., Baris, D., Reding, D. J., et al. (2011). Prediagnostic Serum Levels of Cytokines and Other Immune Markers and Risk of Non-Hodgkin Lymphoma. *Cancer Research,* 71, 4898-907.

Purdue, M. P., Lan, Q., Hoffman-Bolton, J., Hildesheim, A., Callahan, C. L., Strickland, P., et al. (2018). Circulating sCD27 and sCD30 in pre-diagnostic samples collected fifteen years apart and future non-Hodgkin lymphoma risk. *International Journal of Cancer,* doi: 10.1002/ijc.31879.

Purdue, M. P., Lan, Q., Kemp, T. J., Hildesheim, A., Weinstein, S. J., Hofmann, J. N., et al. (2015). Elevated serum sCD23 and sCD30 up to two decades prior to diagnosis associated with increased risk of non-Hodgkin lymphoma. *Leukemia*, 29, 1429-31.

Purdue, M. P., Lan, Q., Martinez-Maza, O., Oken, M. M., Hocking, W., Huang, W.-Y., et al. (2009). A prospective study of serum soluble CD30 concentration and risk of non-Hodgkin lymphoma. *Blood,* 114, 2730-32.

Quezada, H., Guzmán-Ortiz, A. L., Díaz-Sánchez, H., Valle-Rios, R., & Aguirre-Hernández, J. (2017). Omics-based biomarkers: current status and potential use in the clinic. *Boletín Médico del Hospital Infantil de México,* 74, 219-26.

Rajkumar, S. V., Dimopoulos, M. A., Palumbo, A., Blade, J., Merlini, G., Mateos, M.-V., et al. (2014). International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *The Lancet Oncology,* 15, e538-e48.

Ravi, P., Kumar, S. K., Cerhan, J. R., Maurer, M. J., Dingli, D., Ansell, S. M., et al. (2018). Defining cure in multiple myeloma: a comparative study of outcomes of young individuals with myeloma and curable hematologic malignancies. *Blood Cancer Journal,* 8, 26.

Reinisch, W., Willheim, M., Hilgarth, M., Gasché, C., Mader, R., Szepfalusi, S., et al. (1994). Soluble CD23 reliably reflects disease activity in B-cell chronic lymphocytic leukemia. *Journal of Clinical Oncology,* 12, 2146-52.

Riboli, E., Hunt, K. J., Slimani, N., Ferrari, P., Norat, T., Fahey, M., et al. (2002). European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutrition,* 5, 1113-24.

Riboli, E., & Kaaks, R. (1997). The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology,* 26, S6-S14.

Rice, T., Lachance, D. H., Molinaro, A. M., Eckel-Passow, J. E., Walsh, K. M., Barnholtz-Sloan, J., et al. (2016). Understanding inherited genetic risk of adult glioma – a review. *Neuro-Oncology Practice,* 3, 10-16.

Rifai, N., Gillette, M. A., & Carr, S. A. (2006). Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature Biotechnology,* 24, 971.

Rothman, K., & Greenland, S. (2005). Causation and Causal Inference in Epidemiology. *American journal of public health,* 95, S144-S50.

Rothman, N., Hainaut, P., Schulte, P., Smith, M., Boffetta, P. & Perera F (2011). Molecular Epidemiology: Principles and Practices. *IARC Scientific Publications*.

Rubenstein, J. L., Wong, V. S., Kadoch, C., Gao, H.-X., Barajas, R., Chen, L., et al. (2013). CXCL13 plus interleukin 10 is highly specific for the diagnosis of CNS lymphoma. *Blood,* 121, 4740-48.

Rubertone, M. V., & Brundage, J. F. (2002). The Defense Medical Surveillance System and the Department of Defense serum repository: glimpses of the future of public health surveillance. *American journal of public health,* 92, 1900-04.

Rupprecht, T. A., Plate, A., Adam, M., Wick, M., Kastenbauer, S., Schmidt, C., et al. (2009). The chemokine CXCL13 is a key regulator of B cell recruitment to the cerebrospinal fluid in acute Lyme neuroborreliosis. *Journal of neuroinflammation,* 6, 42-42.

Rusch, V., Klimstra, D., Venkatraman, E., Pisters, P. W., Langenfeld, J., & Dmitrovsky, E. (1997). Overexpression of the epidermal growth factor receptor and its ligand transforming growth factor alpha is frequent in resectable non-small cell lung cancer but does not predict tumor progression. *Clinical Cancer Research,* 3, 515-22.

Sallusto, F., Lanzavecchia, A., & Mackay, C. R. (1998). Chemokines and chemokine receptors in T-cell priming and Th1/Th2-mediated responses. *Immunology Today,* 19, 568-74.

Salven, P., Orpana, A., Teerenhovi, L., & Joensuu, H. (2000). Simultaneous elevation in the serum concentrations of the angiogenic growth factors VEGF and bFGF is

an independent predictor of poor prognosis in non-Hodgkin lymphoma: a single-institution study of 200 patients. *Blood,* 96, 3712-18.

Schaller, T. H., Batich, K. A., Suryadevara, C. M., Desai, R., & Sampson, J. H. (2017). Chemokines as adjuvants for immunotherapy: implications for immune activation with CCL3. *Expert Review of Clinical Immunology,* 13, 1049-60.

Scheurer, M. E., Etzel, C. J., Liu, M., Barnholtz-Sloan, J., Wiklund, F., Tavelin, B., et al. (2010). Familial aggregation of glioma: a pooled analysis. *American Journal of Epidemiology,* 172, 1099-107.

Schiffer, L., Henke-Gendo, C., Wilsdorf, N., Hussein, K., Pape, L., Schmitt, C., et al. (2012). CXCL13 as a Novel Marker for Diagnosis and Disease Monitoring in Pediatric PTLD. *American Journal of Transplantation,* 12, 1610-17.

Schlehofer, B., Siegmund, B., Linseisen, J., Schüz, J., Rohrmann, S., Becker, S., et al. (2011). Primary brain tumours and specific serum immunoglobulin E: a case–control study nested in the European Prospective Investigation into Cancer and Nutrition cohort. *Allergy,* 66, 1434-41.

Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J. L., Dotinga, A., Vonk, J. M., et al. (2015). Cohort Profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology,* 44, 1172-80.

Schwaederle, M., Zhao, M., Lee, J., & et al. (2016). Association of biomarker-based treatment strategies with response rates and progression-free survival in refractory malignant neoplasms: A meta-analysis. *JAMA Oncology,* 2, 1452-59.

Schwaederle, M., Zhao, M., Lee, J. J., Eggermont, A. M., Schilsky, R. L., Mendelsohn, J., et al. (2015). Impact of Precision Medicine in Diverse Cancers: A Meta-Analysis of Phase II Clinical Trials. *Journal of Clinical Oncology,* 33, 3817-25.

Schwartzbaum, J., Wang, M., Root, E., Pietrzak, M., Rempala, G. A., Huang, R.-P., et al. (2017). A nested case-control study of 277 prediagnostic serum cytokines and glioma. *PLoS ONE,* 12, e0178705.

Sedgwick, P. (2014). Nested case-control studies: advantages and disadvantages. *BMJ : British Medical Journal,* 348.

SEER. (2013). Cancer Stat Facts: Myeloma. *NIH, National Cancer Institute, Surveillance, Epidemiology, and End Results program*.

SEER. (2015). Cancer Stat Facts: Non-Hodgkin Lymphoma. *NIH, National Cancer Institute*, *Surveillance, Epidemiology, and End Results program*.

Sergentanis, T., Zagouri, F., Tsilimidos, G., Tsagianni, A., Tseliou, M., Dimopoulos, M., et al. (2015). Risk Factors for Multiple Myeloma: A Systematic Review of Meta-Analyses. *Clinical Lymphoma Myeloma and Leukemia,* 15, 563-77.e3.

Sezer, O., Jacob, C., Eucker, J., Niemöller, K., Gatz, F., Wernecke, K.-D., et al. (2001). Serum levels of the angiogenic cytokines basic fibroblast growth factor (bFGF),

vascular endothelial growth factor (VEGF) and hepatocyte growth factor (HGF) in multiple myeloma. *European Journal of Haematology,* 66, 83.

Sigurdardottir, E., Turesson, I., Lund, S., & et al. (2015). The role of diagnosis and clinical follow-up of monoclonal gammopathy of undetermined significance on survival in multiple myeloma. *JAMA Oncology,* 1, 168-74.

Simon, R. (2015). Sensitivity, Specificity, PPV, and NPV for Predictive Biomarkers. *Journal of the National Cancer Institute,* 107, djv153.

Singh, B., & Coffey, R. J. (2014). From wavy hair to naked proteins: the role of transforming growth factor alpha in health and disease. *Seminars in cell & developmental biology,* 28, 12-21.

Smith, A., Crouch, S., Lax, S., Li, J., Painter, D., Howell, D., et al. (2015). Lymphoma incidence, survival and prevalence 2004–2014: sub-type analyses from the UK's Haematological Malignancy Research Network. *British Journal of Cancer,* 112, 1575.

Smith, C. J., Ambs, S., & Landgren, O. (2018). Biological determinants of health disparities in multiple myeloma. *Blood Cancer Journal,* 8, 85-85.

Socialstyrelsen. (2014). Nationella screeningprogram. Modell för bedömning, införande och uppföljning.

Spitz, M. R., Wu, X., & Mills, G. (2005). Integrative Epidemiology: From Risk Assessment to Outcome Prediction. *Journal of Clinical Oncology,* 23, 267-75.

Späth, F., Wibom, C., Krop, E. J. M., Johansson, A.-S., Bergdahl, I. A., Vermeulen, R., et al. (2017). Biomarker dynamics in B-cell lymphoma: a longitudinal prospective study of plasma samples up to 25 years before diagnosis. *Cancer Research,* 77, 1408-1415.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., et al. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ,* 338.

Stupp, R., on behalf of the, E. G. W. G., Brada, M., van den Bent, M. J., Tonn, J. C., & Pentheroudakis, G. (2014). High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Annals of Oncology,* 25, iii93-iii101.

Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R., et al. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood,* 127, 2375-90.

Szklo, M. (1998). Population-based Cohort Studies. *Epidemiologic Reviews,* 20, 81-90.

Teater, M., Dominguez, P. M., Redmond, D., Chen, Z., Ennishi, D., Scott, D. W., et al. (2018). AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. *Nature communications,* 9, 222-22.

Terpos, E., Politou, M., Szydlo, R., Goldman, J. M., Apperley, J. F., & Rahemtulla, A. (2003). Serum levels of macrophage inflammatory protein-1 alpha (MIP-1α) correlate with the extent of bone disease and survival in patients with multiple myeloma. *British Journal of Haematology,* 123, 106-09.

Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., et al. (2013). A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Medical Research Methodology,* 13, 92.

The Cancer Genome Atlas Research, N., McLendon, R., Friedman, A., Bigner, D., Van Meir, E. G., Brat, D. J., et al. (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature,* 455, 1061.

Tsirakis, G., Pappa, C. A., Spanoudakis, M., Chochlakis, D., Alegakis, A., Psarakis, F. E., et al. (2012). Clinical significance of sCD105 in angiogenesis and disease activity in multiple myeloma. *European Journal of Internal Medicine,* 23, 368-73.

Tsou, C.-L., Peters, W., Si, Y., Slaymaker, S., Aslanian, A. M., Weisberg, S. P., et al. (2007). Critical roles for CCR2 and MCP-3 in monocyte mobilization from bone marrow and recruitment to inflammatory sites. *The Journal of clinical investigation,* 117, 902-09.

Turakhia, M. P., & Sabatine, M. S. (2017). How we evaluate biomarker studies. *JAMA Cardiology,* 2, 524-24.

Turesson, I., Bjorkholm, M., Blimark, C., Kristinsson, S., Velez, R., & Landgren, O. (2018). Rapidly changing myeloma epidemiology in the general population: Increased incidence, older patients, and longer survival. *European Journal of Haematology,* 101, 237-44.

Vajdic, C. M., Falster, M. O., de Sanjose, S., Martínez-Maza, O., Becker, N., Bracci, P. M., et al. (2009). Atopic Disease and Risk of Non–Hodgkin Lymphoma: An InterLymph Pooled Analysis. *Cancer Research,* 69, 6482-89.

van de Schans, S. A. M., Issa, D. E., Visser, O., Nooijen, P., Huijgens, P. C., Karim-Kos, H. E., et al. (2011). Diverging trends in incidence and mortality, and improved survival of non-Hodgkin's lymphoma, in the Netherlands, 1989-2007. *Annals of Oncology,* 23, 171-82.

van der Weyden, C. A., Pileri, S. A., Feldman & Prince, H. M. (2017). Understanding CD30 biology and therapeutic targeting: a historical perspective providing insight into future directions. *Blood Cancer Journal,* 7, e603.

van Leeuwen, M. T., Grulich, A. E., Webster, A. C., McCredie, M. R. E., Stewart, J. H., McDonald, S. P., et al. (2009). Immunosuppression and other risk factors for early and late non-Hodgkin lymphoma after kidney transplantation. *Blood,* 114, 630-37.

van Oers, M., Pals, S., Evers, L., van der Schoot, C., Koopman, G., Bonfrer, J., et al. (1993). Expression and release of CD27 in human B-cell malignancies. *Blood,* 82, 3430-36.

Wang, H., Wang, L., Chi, P.-D., Wang, W.-d., Chen, X.-Q., Geng, Q.-R., et al. (2016). High level of interleukin-10 in serum predicts poor prognosis in multiple myeloma. *British Journal of Cancer,* 114, 463-68.

Vargas, A. J., & Harris, C. C. (2016). Biomarker development in the precision medicine era: lung cancer as a case study. *Nature Reviews Cancer,* 16, 525.

Vasan, R. S. (2006). Biomarkers of Cardiovascular Disease. *Circulation,* 113, 2335-62.

Wedemeyer, J., Lorentz, A., Göke, M., Meier, P. N., Flemming, P., Dahinden, C. A., et al. (1999). Enhanced production of monocyte chemotactic protein 3 in inflammatory bowel disease mucosa. *Gut,* 44, 629-35.

Weiner, J., 3rd, Maertzdorf, J., Sutherland, J. S., Duffy, F. J., Thompson, E., Suliman, S., et al. (2018). Metabolite changes in blood predict the onset of tuberculosis. *Nature communications,* 9, 5208-08.

Weinstein, A. M., & Storkus, W. J. (2015). Therapeutic Lymphoid Organogenesis in the Tumor Microenvironment. *Advances in cancer research,* 128, 197-233.

Weiss, B. M., Abadie, J., Verma, P. & Kuehl, W. M. (2009). A monoclonal gammopathy precedes multiple myeloma in most patients. *Blood,* 113, 5418-22.

Weiss, B. M., & Kuehl, W. M. (2010). Advances in understanding monoclonal gammopathy of undetermined significance as a precursor of multiple myeloma. *Expert review of hematology,* 3, 165-74.

Vendrame, E., & Martínez-Maza, O. (2010). Assessment of Pre-Diagnosis Biomarkers of Immune Activation and Inflammation: Insights on the Etiology of Lymphoma. *Journal of Proteome Research,* 10, 113-19.

Vermeulen, R., Hosnijeh, F. S., Portengen, L., Krogh, V., Palli, D., Panico, S., et al. (2011). Circulating Soluble CD30 and Future Risk of Lymphoma; Evidence from Two Prospective Studies in the General Population. *Cancer Epidemiology Biomarkers & Prevention,* 20, 1925-27.

Vermeulen, R., Saberi Hosnijeh, F., Bodinier, B., Portengen, L., Liquet, B., Garrido-Manriquez, J., et al. (2018). Pre-diagnostic blood immune markers, incidence and progression of B-cell lymphoma and multiple myeloma: Univariate and functionally informed multivariate analyses. *International Journal of Cancer,* 143, 1335-47.

Wieduwilt, M. J., & Moasser, M. M. (2008). The epidermal growth factor receptor family: biology driving targeted therapeutics. *Cellular and molecular life sciences : CMLS,* 65, 1566-84.

Wiemels, J. L., Wilson, D., Patil, C., Patoka, J., McCoy, L., et al. (2009). IgE, allergy, and risk of glioma: Update from the San Francisco Bay Area Adult Glioma Study in the Temozolomide era. *International Journal of Cancer,* 125, 680-87.

Vincent, L., Jin, D. K., Karajannis, M. A., Shido, K., Hooper, A. T., Rashbaum, W. K., et al. (2005). Fetal Stromal–Dependent Paracrine and Intracrine Vascular

Endothelial Growth Factor-A/Vascular Endothelial Growth Factor Receptor-1 Signaling Promotes Proliferation and Motility of Human Primary Myeloma Cells. *Cancer Research,* 65, 3185-92.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv:1308.5499.*

Wu, H.-C., Wang, Q., Delgado-Cruzata, L., Santella, R. M., & Terry, M. B. (2012). Genomic methylation changes over time in peripheral blood mononuclear cell DNA: differences by assay type and baseline values. *Cancer Epidemiology, Biomarkers & Prevention,* 21, 1314-18.

Wynn, T. A. (2003). IL-13 Effector Functions. *Annual Review of Immunology,* 21, 425-56.

Xu, L. L., McVicar, D. W., Ben-Baruch, A., Kuhns, D. B., Johnston, J., Oppenheim, J. J., et al. (1995). Monocyte chemotactic protein-3 (MCP3) interacts with multiple leukocyte receptors: binding and signaling of MCP3 through shared as well as unique receptors on monocytes and neutrophils. *European Journal of Immunology,* 25, 2612-17.

Yahalom, J. (2014). Radiotherapy of follicular lymphoma: updated role and new rules. *Current treatment options in oncology,* 15, 262-68.

Yang, J., Xu, H. e., Liang, X., Lv, S., & Jia, Y. (2017). Allergic conditions are not associated with the risk of non-Hodgkin's lymphoma or Hodgkin's lymphoma: a systematic review and meta-analysis. *OncoTargets and therapy,* 10, 2189-98.

Yarden, Y., & Pines, G. (2012). The ERBB network: at last, cancer therapy meets systems biology. *Nature Reviews Cancer,* 12, 553.

Yawetz, S., Cumberland, W., Van der Meyden, M., & Martinez-Maza, O. (1995). Elevated serum levels of soluble CD23 (sCD23) precede the appearance ofacquired immunodeficiency syndrome--associated non-Hodgkin's lymphoma. *Blood,* 85, 1843-49.

Yotsukura, S., & Mamitsuka, H. (2015). Evaluation of serum-based cancer biomarkers: A brief review from a clinical and computational viewpoint. *Critical Reviews in Oncology/Hematology,* 93, 103-15.

Zingone, A., & Kuehl, W. M. (2011). Pathogenesis of Monoclonal Gammopathy of Undetermined Significance (MGUS) and Progression to Multiple Myeloma. *Seminars in hematology,* 48, 4-12.

Zintzaras, E., Voulgarelis, M., & Moutsopoulos, H. M. (2005). The risk of lymphoma development in autoimmune diseases: A meta-analysis. *Archives of Internal Medicine,* 165, 2337-44.

Zinzani, P. L., Federico, M., Oliva, S., Pinto, A., Rigacci, et al. (2015). The more patients you treat, the more you cure: managing cardiotoxicity in the treatment of aggressive non-Hodgkin lymphoma. *Leukemia & Lymphoma,* 56, 12-25.