Immunoglobulin G glycosylation in patients with colorectal cancer

Thaçi, Kujtim

UNIVERSITY OF ZAGREB
SCHOOL OF MEDICINE

KUJTIM THAÇI

IMMUNOGLOBULIN G GLYCOSYLATION IN PATIENTS WITH COLORECTAL CANCER

DISSERTATION

ZAGREB 2017

UNIVERSITY OF ZAGREB
SCHOOL OF MEDICINE

KUJTIM THAÇI

IMMUNOGLOBULIN G GLYCOSYLATION IN PATIENTS WITH COLORECTAL CANCER

DISSERTATION

ZAGREB 2017

This dissertation was made at Glycobiology Laboratory, Genos Ltd., Zagreb, under the supervision of professor Gordan Lauc, PhD. The research was funded by FP7 grant Glycobiom.

Mentor: Prof.Gordan Lauc, PhD

Acknowledgement:

I wish to acknowledge everyone who has assisted me throughout my doctoral studies over the years. First of all I want to thank my mentor prof.dr. Gordan Lauc who was abundantly helpful and offered invaluable assistance, support and guidance. I can't find enough words to express my gratitude, therefore I oblige myself to continue our work in the future in the best possible way I can.

I thank from my heart to the very kind and supportive staff of ''Genos'', with a special thanks to Dr. Maja Pučić Baković, Jasminka Ilić and Mislav Novokmet for their irreplaceable help and support.

I am forever indebted to my parents and my wife for their understanding, endless patience and encouragement when it was most required.

To my inspiration,my beloved sons, LIRI and LISI.

I. TABLE OF CONTENTS

1. INTRODUCTION AND BACKGROUND FOR THE PROPOSED RESEARCH....1

1.1. Protein Glycosylation....1

1.2. Glycans relevance as biomarkers....3

1.3. Immunoglobulin G....5

1.4. The role of the Fc glycans in IgG activity....8

1.5. IgG glycosylation under physiologic and pathophysiologic conditions....15

1.6. The prognostic importance of colorectal cancer....17

2. HYPOTHESIS....21

3. AIMS AND PURPOSE OF THE RESEARCH....22

3.1. General aim:....22

3.2. Specific aims:....22

4. MATERIALS AND METHODOLOGY....23

4.1. Study population....23

4.2. Ethics statement....24

4.3. IgG N-glycans sample preparation and analyses....24

4.3.1 Purification of IgG....24

4.3.2 Release and labelling of IgG glycans....25

4.3.3 HILIC-SPE....26

4.3.4 HILIC-UPLC....26

4.4. Survival and risk related parameters....27

4.5. Statistical analysis....27

4.5.1 Supplementary Section: Model Selection....31

5. RESULTS....33

5.1. IgG glycan measurements in colorectal cancer prognosis....33

5.2. Survival analysis....40

5.3. IgG glycome composition in CRC patients and matching controls.....78

6. DISCUSSION 87

6.1. Stage differences 87

6.2. Galactosylation, sialylation, GlcNAc and core fucosylation residues present on IgG glycans in cancer 88

7. CONCLUSION 95

8. SAŽETAK / ABSTRACT IN CROATIAN 97

9. ABSTRACT IN ENGLISH 98

10. REFERENCES 99

11. LIST OF FIGURES AND TABLES 123

12. CURRICULUM VITAE 127

II. LIST OF ABBREVIATIONS

2-AB	2-aminobenzamide
ACN	Acetonitrile
ADCC	Antibody Depedent Cellular Cytotoxicity
AHA	Autoimmune Hemolitic Anemia
AJCC	American Joint Committee on Cancer
ASN	Asparagine
AUC	Area Under Roc Curve
BMI	Body Mass Index
CA19-9	Cancer Antigen 19-9
CD4+	Cluster Of Differentiation 4
CEA	Carcinoembryonic Antigen
CH	Constant Heavy Chain
CIDP	Chronic Inflammatory Dimyelinating Polyneuropathy
CL	Constant Light Chain
CRC	Colorectal Cancer
CT	Computed Tomography
DC-SIGN	Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Non-integrin
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
Fab	Antigen Binding Fragment
Fc	Crystallysable Fragment

FcγRIa	Crystallisable Fragment Gamma Receptors Ia
FcγRIIa	Crystallisable Fragment Gamma Receptors IIa
FcγRIIIa	Crystallisable Fragment Gamma Receptors IIIa
FcRs	Crystallisable Fragment Receptors
FcγRsIIB	Crystallisable Fragment Gamma Receptors IIB
FcRn	Crystallisable Fragment Neonatal Receptor
Fuc	Fucose
FUT 8	Fucose Transferase 8
Gal	Galactose
Gal T	Galactose Transferase
GDP Fuc	Guanosine 5'-diphospho-fucose
Glc	Glucose
GlcNAc	N-Acetylglucosamine
GnTIII	N-Acetylglucosaminyltransferase III
GnT-V	N-Acetylglucosaminyltransferase V
GP	Glycan Peak
HILIC-SPE	Hydrophilic Interaction Liquid Chromatography Solid- Phase Extraction
HIV	Human Immunodeficiency Virus
HNF1A	Hepatocyte Nuclear Factor 1A
HNF1A-MODY	Hepatocyte Nuclear Factor 1A- Maturity Onset Diabetes Of Young
HR	Hazard Ratio
IBD	Inflammatory Bowel Disease
IDI	Integrated Discrimination Index

IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
ITP	Immune thrombocytopenia
IVIG	Intravenous Immunoglobulins
LEMS	Lambert-Eaton Myasthenic Syndrome
mAb	Monoclonal antibody
Man	Manose
MBL	Mannose Binding Lectin
MG	Myasthenia Gravis
NK	Natural Killer
OR	Odds Ratio
PBS	Phosphate-Buffered Saline
PD-1	Programmed Cell Death-1
PNGase	Peptide: N-Glycosidase F
RA	Reumathoid Arthritis
RF	Reumathoid Factor
SDS	Sodium Dodecyl Sulfate
Ser/Thr	Serine/Threonine
Sia T	Sialyltransferase
SIGNR1	Specific Intercellular Adhesion Molecule-3-Grabbing Receptor 1
SLE	Systemic Lupus Erithematosis

SLex	Sialyl-Lewis-X
SNPs	Single Nucleotide Polymorphisms
SOCCS	Study of Colorectal Cancer Scotland
ST6GalNAc	N-acetylgalactosaminide α2-6-sialyltransferase
TNF- α	Tumor Necrosis Factor Alpha
TNM	Tumor Nody Metastases
UPLC	Ultra Performance Liquid Chromatography
VH	Variable Heavy Chain
VL	Variable Light Chain

1. INTRODUCTION AND BACKGROUND FOR THE PROPOSED RESEARCH

1.1. Protein Glycosylation

Glycans are oligosaccharides attached to proteins or lipids [1]. A typical glycan is a complex non-linear branched oligosaccharide composed of 10 to 15 monosaccharide residues. Glycan structure is defined by the order of monomeric units, exact position of the glycosidic bond, its anomeric configuration (α or β), the number of branches and the position of branching [2]. The process of oligosaccharides attachement to a protein is called glycosylation and those proteins are referred to as glycoproteins. The complete set of glycans produced by a given cell type or organism is called glycome and comprehensive study of structures and functions of glycans is referred to as glycomics. Mammalian glycome consist of nine monosaccharides which theoretically could make 10^{12} combination of hexasaccharides [3]. Moreover, a broad range of different glycans can be made by altering the number, order and type of monosaccharides units and the attachment of many different glycans to the same protein result in different glycoforms of the same protein molecule [4]. It is estimated that mammalian glycoproteins consist of up to 2000 different glycan determinants [5] and changes in glycans have an essential impact in modifying the structure and function of polypeptides parts of glycoprotein [6], making the glycoproteome much more complex than the proteome itself [7].

Glycosylation of proteins is an important post-translational modification and glycans play an important role in normal physiological processes such as protein folding, degradation and secretion as well as in cell signalling, cell-cell interactions and immune functions [8]. Dysregulation of glycosylation caused by a combination of several individual mutations leads to various forms of congenital disorders of glycosylation [9]. In addition mutations that cause formation of new glycosylation site can also be deleterious. There are several classes of glycans, including Asn (N)-linked and Ser/Thr (O)-linked glycans [1]. N-glycosylation is the oldest pathway that take place when a block of 14 sugars attached to dolichol phosphate (the dolichol phosphate precursor) is transferred to asparagine (Asn) residue in newly synthesized polypeptides in the endoplasmatic reticulum [10]. During passage through the Golgi compartments to their intra and extra-cellular destinations, these N-glycans will be subject to a large-scale alteration [11]. Common mammalian monosccharides comprise glucose (Glc), mannose (Man),galactose (Gal), N-Acetylglucosamine (GlcNAc), fucose (Fuc) and sialic acid. Additionally, multiple residues can be attached to a single monosccharide which can

lead to complex branched structure. Therefore, oligosaccharides are amongst the most structurally diverse biopolymers in the nature.

Almost all membrane and secreted proteins are modified by covalent addition of glycans [12] and the complete absence of glycans is embryologically lethal [13]. Contrary to polypeptide parts of proteins, which are defined by the sequence of nucleotides in the corresponding genes (made by a direct DNA template), glycans are formed by complex network of interactions by hundreds of glycogenes [14] that code for various glycosyltransferases, glycosidases, enzymes for sugar nucleotide biosynthesis, transporters, transcription factors, ion channels and other proteins [15]. Therefore, the final structure of the glycan will be ultimately affected by changes in the abundance and/or localization of any of the enzymes, glycoprotein substrates and activated sugar donors involved in glycan biosynthesis [16]. Thus, it is well established that glycans are a product of a complex biosynthetic pathway affected by hundreds of genetic and environmental factors [14, 17, 18].

Despite their different and complex biosynthetic origin, both polypeptide and glycan parts of glycoproteins participate as a single molecular entity in the function of a glycoprotein [6]. Even though the glycans are not directly encoded in genes, a high heritability of mammalian glycome composition was reported [19, 20] and various large population studies of human plasma glycome revealed a high inter-individual variability in glycan composition [20, 21]. This heterogenity of glycome composition triggered by alteration in glycosylation has resulted in the tremendous complexity in glycan structures and even a minor changes in the glycan structures can have important functional effects. It is evident that this structural variety of glycans is responsible for adapting to changing environment and also in evading from many pathogens [22]. ABO blood groups are the best conspicuous example of variability in monosaccharides at the end of glycan antennas as a result of modification caused by mutation in genes leading to a large part of individual phenotypic variations. The majority of human variability derives from single nucleotide polymorphisms (SNPs) which individually do not indicate any obvious phenotypes, but if present in specific combinations, they can have significant phenotypic effects [23]. In addition, there is a growing evidence that changes in protein glycosylation have been reported to be of importance in the development and progression of different cancers, thus raising the possibility for early detection, imaging and therapy [24].

1.2. Glycans relevance as biomarkers

Glycans appear to play a key role in a wide range of diseases. The complexities of oligosaccharide structures and the lack of analytical methods for elucidating them have hampered the overall process and only recently glycans have been tested as potential disease biomarkers due to the recent advances in the development of novel analytical tools. It is now well established that structure of glycan expressed by diseased cells are different as compared to those of normal cells [25] and these aberrant glycans are widely known to occur as a response to the disease [26, 27, 28]. Since , glycans are not a result of direct genetic template, it appears that changes in protein glycosylation have the potential to be more notable than protein expression during transformation to the diseased condition [26]. Even small changes in the associated protein (glycosyltransferases) expression have a direct effect on glycans and there are various examples that emphasize the role of glycosyltransferase overexpression in tumorigenesis [29, 30]. As a result, the effect of the disease is significantly amplified and easier to detect [1].

On the other side, it is widely reported that abnormality in protein glycosylation is associated with a complex diseases such as cancer, diabetes, cardiovascular, congenital, immunological and infectious disorders [3, 8, 31]. Furtheremore, changes in protein glycosylation in either the level of or type of glycosylation have been shown to be of importance in the development and progression of different cancers. Hence, glycans are engaged in all major physiological events during various stages of tumor progression, from tumor cell proliferation, metastasis and angiogenesis [26, 32, 33], and a number of different studies have made preliminary reports of potentially important glycan biomarkers for cancer and other diseases [1, 33, 34, 35, 36]. Therefore, due to the importance of alteration in glycosylation in health and disease, there has been extensive research effort on the discovery of glyco-biomarkers that could have potential diagnostic or prognostic monitoring capabilities.

Many known biomarkers for cancer are in fact glycoproteins, and specifically those for colorectal cancer are carcinoembryonic antigen (CEA) and CA19-9. However, diagnostic tests typically only measure the protein fraction, despite the fact that in many cases it has been convincingly demonstrated that glycosylation changes significantly improve the diagnostic value of these biomarkers [37, 38]. Increased branching of oligosaccharides has been reported to be associated with cancer metastasis and tumor progression in melanoma, breast and colorectal cancer [39, 40, 41]. The enzyme N-Acetylglucosaminyltransferase V (GnT-V, Mgat 5), which promotes increased branching of glycans has been reported to be associated with cancer [42]. Its expression was subsequently reported to correlate with poor prognosis in

colorectal [43], endometrial [44] and bladder cancer. In mouse models, knockdown of GnT-V was found to inhibit breast cancer cell growth with activation of CD4+ T cells and macrophages [45]. Functional proteomics studies indicated that GnT-V reinforces the invasive/metastatic potential of colon cancer through aberrant glycosylation on tissue inhibitor of metalloproteinase 1 [46]. More recently 1,3 fucosyltransferases were reported to be master regulators of prostate cancer cell trafficking [47], while the ST6GalNAc gene (one of many sialyl-transferases) was found to be associated with brain metastasis in breast cancer [48]. A recent study reported that polymorphisms in several glyco-genes are also associated with increased or decreased risk for epithelial ovarian cancer [49].

Glycosylation found on cell surfaces and in extracellular matrices is very important in multicellular organisms because it makes the first point of contact in cellular interactions [50]. This has significant impact in cancer progression and metastasis because cancer progression and metastasis is a complex process requiring adhesive interactions, many of which are mediated by cell surface glycans and lectins. N-glycosylated cell adhesion molecules such as E-cadherin and integrins play key roles in cell-cell and cell-extracellular matrix interactions, thereby affecting cancer metastasis [51]. A key element of protein glycosylation is the addition of fucose to the non-reducing ends of N-glycans and this has already been linked to cancer and inflammation [52]. Indeed, fucosylated haptoglobin and sialyl-Lewis-x (SLex) have previously been proposed as cancer biomarkers for pancreatic, ovarian, breast and thyroid cancer [39, 40, 41, 53, 54]. Fucosylated alpha-fetoprotein is highly specific tumor marker for hepatocellular carcinoma [55]. Although, changes in fucosylation are associated with progression of prostate cancer [56]. More recently mutation in hepatocyte nuclear factor (HNF1A) gene which finally cause a decrease in plasma protein antennary fucosylation is a promising biomarker for both, HNF1A-MODY(maturity onset diabetes of the young) diagnosis and HNF1A impaired function [57].

Breast cancer patients with lymph node metastasis have been shown to have significantly higher levels of glycans containing the SLex structure than patients without a metastasis, suggesting a basis for a new biomarker for cancer prognosis [40]. In particular, malignant properties of cancer cells including invasivness and metastatic potential is thought to be a characteristic feature associated with aberrant sialylations in these cells. An increase in sialylation is commonly observed in various tumors, which may be due to either an increased activity of sialyltransferases or increased numbers of possible sialylation sites on N-linked carbohydrates [58]. Sialylation of transferrin and total serum proteins coud be used as a new prognostic marker of acute pancreatitis [59]. Moreover, decreases in triantennary

trigalactosylated glycans and/or bisected core fucosylated biantennary monosialylated glycans and increases in tetraantennary tetrasialylated glycans were found to be correlated with perineural invasion in prostate cancer [60]. Triantenary N-glycans are associated with tumor stage in hepatocarcinoma patients [61]. Modulators of protein glycosylation and glycan branches have been considered to be important candidates for novel anticancer drugs [26, 62]. Even though it appears that the reported markers are cancer associated rather than cancer specific, they nevertheless could provide an insight into the molecular mechanisms of carcinogenesis and disease pathogenesis [63].

Finally, Immunoglobulin G (IgG) molecule is a prominent example of changes in protein glycosylation as a response to the disease. Human serum IgG glycosylation is known to change with various physiological and pathological conditions, especially, it is evident a significant change of IgG glycosylation between control and cancer patients [64].

Taking all of these studies in consideration it is predicted that, in the future, it will be possible to not only diagnose diseases, but perhaps even determine disease progression and specific strain based solely on glycan profiling [1].

1.3. Immunoglobulin G

Immunoglobulins (Ig) are glycoprotein molecules made by plasma cells in response to challenge from microbiological agents or cancer cells. Compared to other classes of immunoglobulins (IgM, IgA, IgD, IgE), IgG molecules represent the predominant class of Igs in the serum with typical concentration of almost 10 mg/ml [65], reflecting its role as a major effector molecule of the humoral immune system in defending organisms against many pathogens. IgG molecules are multifunctional glycoproteins which are one of the best studied class of glycoproteins present in serum in four different subclasses (IgG1, IgG2, IgG3, IgG4) based on their decreasing abundance (65), which also differ from each other on the basis of the length of hinge region and the number of interchain disulphide bonds [66, 67].

IgG antibodies consist of two heavy (H) and two light (L) chains linked together by disulphid bonds which form two fragment antigen binding Fab moieties divided into two homologus domains (VL and CL) and one fragment crystallizable Fc moiety composed of four homologous domains (VH, CH1, CH2 and CH3) which are both together linked through a flexible hinge region [68, 69]. While the Fab portion of IgG are responsible for binding antigens and provides the structural basis for the tremendous immunological diversity of antibodies, the Fc portion promotes a variety of effector functions through interactions with specific ligands, e.g. cellular receptors (FcRs), (FcRn) and the C1 component of complement [70, 71, 72].

Each heavy chain of IgG carries a single covalently attached bi-antennary N-glycan at the highly conserved asparagine 297 residue in each of the CH2 domains of the Fc region of the molecule **(Figure 1).** In contrast to other Ig isotypes, the IgG-Fc glycan moiety is not exposed on the IgG surface but rather buried within the hydrophobic core between the two heavy chains of the CH2-domain forming multiple non-covalent bonds with polypeptide chains and impacting Fc structure [73, 74].

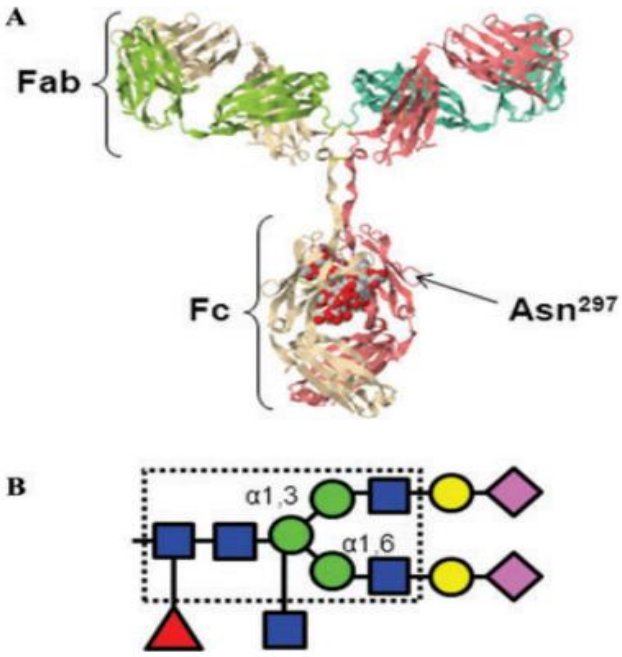


Figure 1. *The structure of an IgG1 molecule. (A) Fab and Fc portion (B). Complex biantennary glycan structure attached to Asn 297 in Fc portion of IgG molecule. Structural schemes are given in terms of blue square (N-acetylglucosamine), red triangle (fucose), green circle (mannose), yellow circle (galactose), and purple diamond (N-acetylneuraminic acid)* [64].

The Fc glycans of human IgG are biantennary complex-type structures which are predominantly core-fucosylated and are in part altered by a bisecting N-acetylglucosamine (GlcNAc) [75, 76], antennae are partially trimmed varying in their degree of galactosylated and may carry a sialic acid residue [77].

More than 95% of the final IgG glycans carry a N-acetylglucosamine on both arms and 85% are fucosylated [78]. The percentage of glycans carrying one galactose (G1 glycan) is 40% and the frequency of non-galactosylated (G0) or bi-galactosylated glycans (G2) ranges between 20-40% [79].

In addition, it appears that there are small amounts of human IgG containing nonfucosylated glycans with or without bisecting GlcNAc residue [80]. In general, it is widely known that the attached oligosaccharide chains are highly heterogenous and over 36 different glycans can be

attached to the conserved Asn 297 of the IgG heavy chains [81]. Since IgG glycoprotein contains two heavy chains, it can lead to hundreds of different IgG isomers that can be generated from this single glycosylation site. Immunoglobulin G (IgG) molecule is one of the most studied glycoprotein in details in sense of structural and functional aspects of glycosylation.

Interestingly it was found that the variability of glycan composition attached to IgG was approximately three times larger than the variability of the total plasma glycome [20], providing IgG with new physiological functions that could not be possible without this complex post-translational process.

Thus, glycosylation is a very complex metabolic process and heterogenity in IgG glycome composition argues the fact that these final products, glycans, are not solely genetically predefined **(Figure 2),** hence glycosylation is known to be affected by factors such as type of glyco-enzymes and their expression levels [82].

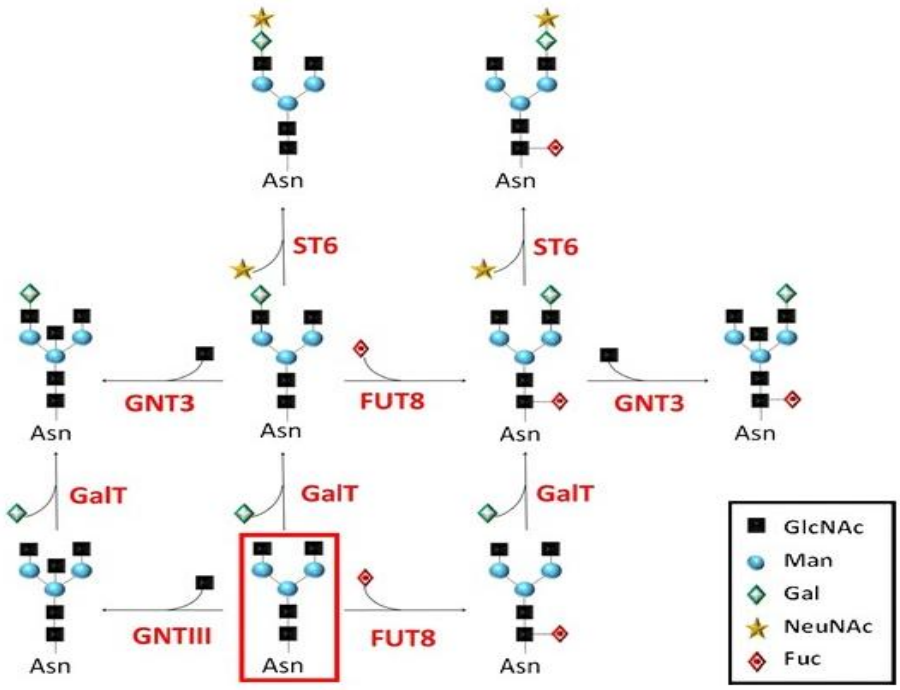


Figure 2. *Structural variations in IgG glycans. Initial GlcNAc2Man3GlcNAc2 structure (red square) can be modified by the addition of bisecting GlcNAc (GnTIII), fucose (FUT8) or galactose (GalT). These resulting structures can further be modified by the activity of the same enzymes or by the addition of the sialic acid (SiaT)* [83].

These attached oligosaccharides are structurally important for the stability of the antibody and its effector functions [84]. In addition, 15–20% of normal IgG molecules also carry complex bi-antennary oligosaccharides attached to the variable regions of the light chain, heavy chain or

both [85, 86]. There are glycosylation site specific differences between the Fab and the Fc. In contrast to Fc portion of IgG, Fab regions are characterised by a high incidence of di- and monosialylated structures and of cores with the bisecting GlcNAc residue [87].

In general, these studies indicate that glycosylation of IgG is essential for the expression of proper effector functions of IgG, and specific physiological and pathological conditions can essentially alter the conformation of Fc region with enormous consequences for IgG effector functions. Therefore, variation of the glycosylation of recombinant monoclonal antibodies in order to derive improved therapeutic effector functions, represents an important research area [88]. Despite the importance of protein glycosylation in all physiological and pathological processes and their potential as diagnostic markers, understanding of the role of IgG N-glycans in various pathological conditions is also a good foundation for the development of novel therapeutics.

1.4. The role of the Fc glycans in IgG activity

Glycosylation appears to be particularly important in the immune system [89] and IgG is the most conspicuous example in terms of how novel IgG functions triggered by alternative glucosylation can provide new adaptive mechanism that allows fight against pathogens or cancer cells. Immunoglobulins and their binding receptors (Fc receptors) are key glycoprotein components of the immune system that link the innate and adaptive arms of immunity. IgG-Fc glycans modulate the biological activities of IgG in completely two opposite directions, thereby acting as a switch from innate anti-inflammatory activity to an adaptive pro-inflammatory response on antigenic challenge [1].

Antibodies communicate with effector cells of the immune system through interactions with membrane bound Fc receptors, which are complex glycoproteins widely expressed throughout haematopoietic system [90] such as natural killer cells, macrophages, eosinophils, neutrophils, limphocytes and dentric cells. Stimulation of cells through FcγRs results in a wide variety of effector functions, including antibody dependent cell mediated cytotoxicity (ADCC) [91] phagocytosis [92], oxidative burst [93] and release of inflammatory mediators [94]. Consequently, alteration of IgG glycosylation patterns changes their respective effector functions [95, 96]. There are two major types of FcγRs: immune reaction activating (FcγRIa, FcγRIIa and FcγRIIIa) and inhibiting receptors (FcγRIIb) [74], also there is an additional Fc receptor known as the neonatal Fc receptor, (FcRn). FcRn is involved in IgG, maintaining the serum concentrations of the antibody due to the enhanced half life of circulating IgG of up to 21 days and regulating IgG homeostasis where high concentrations are required to fight infection [97]. This receptor is

also involved in transcellular transportation of IgG from mother to fetus across the placenta and is critically important for the transfer of humoral immunity to the fetus [98].

Some of the immune cells express both activating and inhibitory FcγR and eventual interaction betwen IgG and these receptors relies on the affinity of the IgG subclasses to distinct FcγRs and the combination of expression level an pattern of FcγRs [99].

IgG glycosylation has important regulatory functions and immune system has possibility to descriminate between different antibody glycoforms, which can also initiate a different impact on the efficacy of IgG effector function [100]. The IgG-Fc glycans maintain the heavy chains in an open conformation required for FcγRs interactions and even a small changes in their composition have enormous consequences for IgG effector functions [85, 101]. However, it is also evident that deglycosylated IgG molecules are unable to elicit *in vivo* an inflammatory response [102] because of their little ability in activating complement and in binding to FcγR [103, 104] as the two heavy chains form a closed conformation, impeding so formation of FcγR binding pocket [105].

Antibody mediated tumor eradication depends on two key mechanisms targeted for activation, ADCC and complement-dependent cytotoxicity (CDC). In order for IgG to induce these two key mechanisms it must rely on proper glycosylation of its Fc region since changes in glycan composition will initiate dramatic consequences for effector functions of IgG **(Figure 3).**

Fc glycans contain a core Fuc residue in a1,6-position linked to the core GlcNAc residue [78]. Bio-synthesis of core-fucosylated glycans is the result of a transfer of a Fuc residue from GDP-Fuc-mediated by a1,6-fucosyltransferase in the trans-Golgi [106]. Core-fucosylation of N-glycans attached to the Fc part of IgG appears to be specifically and separately regulated from the core-fucosylation of glycans attached to Fab fragments of the same protein [107]. While the majority of plasma proteins are not core-fucosylated [19], over 95% of IgG is core-fucosylated in what appears to be a mechanism which modulates binding of IgG to Fc gamma RIIIa [20] and therefore contains “safety switch” which prevents them from eliciting potentially destructive ADCC [108]. FcγRIIIa (CD16) is an activating Fc receptor expressed primarily on natural killer (NK) cells and antibodies initiate ADCC via binding to FcγRIIIa which results in eradication of target cells. Core-fucosylation of IgG plays an important role in mediating ADCC following treatment with therapeutic antibodies because the presence of core-fucose interferes with FcγRIIIa binding and this is closely related to the clinical efficacy of anticancer drugs [109]. Hence, IgG fucosylation is a strong modulator of antigen dependant cell cytoxicity (ADCC) [110, 111] and specific aspects of FcγRIIIa glycosylation have also been reported to be important in this process. Glycans attached to Asn45 and Asn162 of FcγRIIIa are essential for

the regulation of preferential binding of IgG without core-fucose to FcγRIIIa [109]. Despite the role of core fucose in increasing or decreasing ADCC via binding to FcγRIIIa, it was reported that interaction with other activating receptors such as FcγRIIA is uneffected by the presence of the core fucose [112]. Nevertheless, proper regulation of the addition of core-fucose to IgG could play an important role in determining cancer prognosis as the addition of core-fucose decreases the potential of IgG to elicit ADCC by nearly hundred fold [95]. Because of improved binding to activating Fcγ receptor and enhanced ADCC by afucosylated IgG, core-fucosylation of IgG has been the subject of intensive research [95].

Hence, a number of strategies have been developed to reduce fucosylation of IgGs [107] such as: generation of cell lines with genetic modifications that directly influence core fucosylation in a way that either completely lack or have reduced levels of expression of a 1,6-fucosyltransferase [113, 114]. For instance, genetic elimination of fucosyltransferase 8 (FUT8) in Chinese hamster ovary cells has been a successful approach for the prevention of core fucosylated IgG [115]. Besides, as an alternative method is used overexpression of b (1,4)-N-acetylglucosaminyltransferase III (GnTIII) an enzyme that adds bisecting GlcNAc residues to IgG [116] which has been shown to result in reduction of the core fucose content [117]. However, it seems that the lack of core fucose, not the presence of bisecting GlcNAc, has the most critical role in enhanced ADCC [111] as the presence of bisecting GlcNAc is always associated with low fucose content. Biopharmaceutical industry has been focused on generation of afucosylated monoclonal antibody therapeutics (mAb) as well, and the majority of approved (mAb) therapeutics that target ADCC have been engineered for use in anti-cancer therapies. Afucosylated monoclonal antibodies display strong therapeutic potential in anti-cancer therapy [118]. Due to their ability to initiate an enhanced ADCC as a result of their improved FcγRIIIa binding, afucosylated IgG even low doses of afucosylated IgG are enough to overcome the competition with high concentrations of heavily fucosylated serum IgG [110].

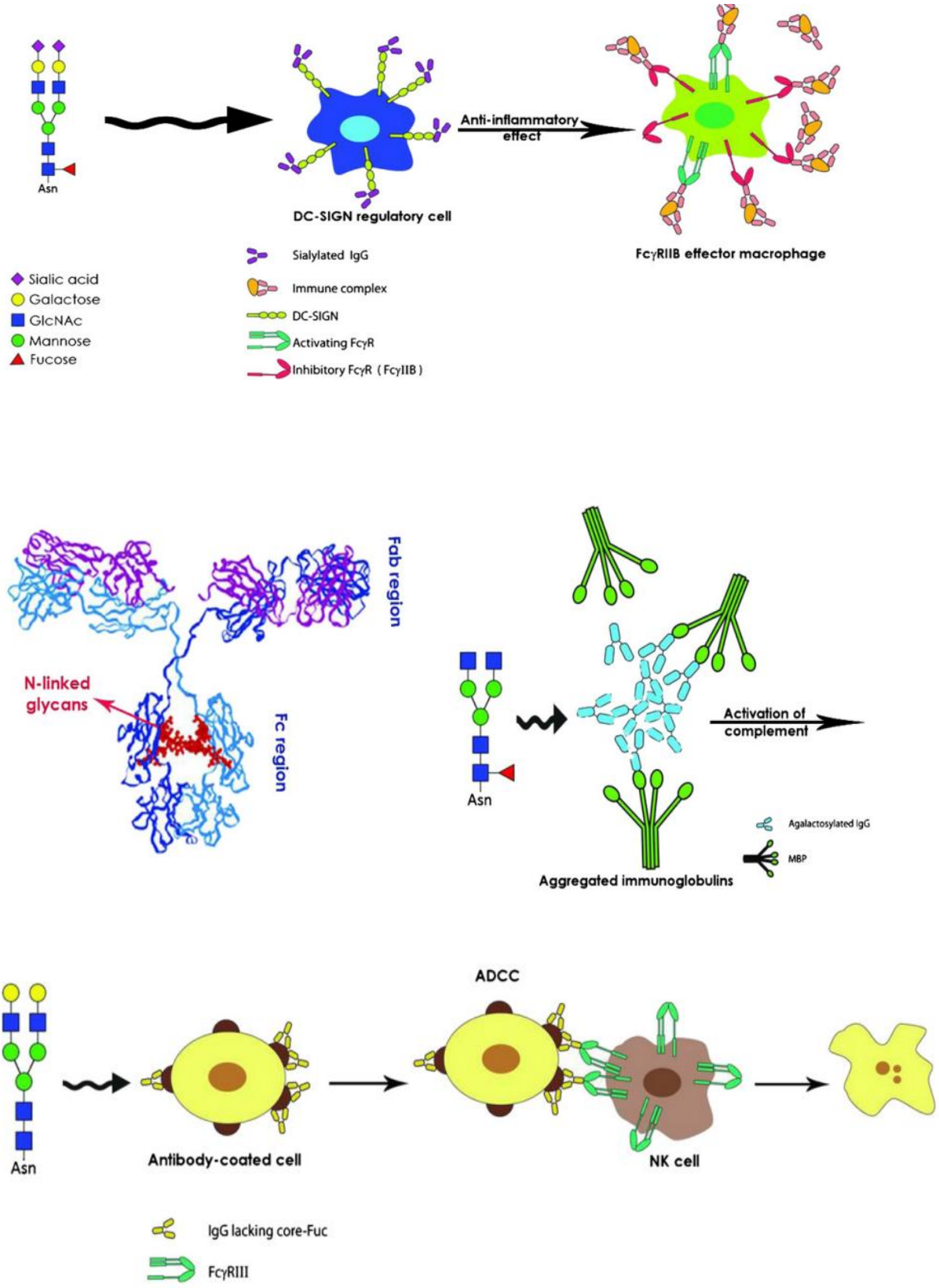


Figure 3. *Modulation of IgG function by alternative glycosylation.Structure of the glycan on IgG Fc part can significantly affect effector function of IgG* [119].

Another structural modification of IgG-Fc glycan is the presence of terminal sialic acid residues and this has received increased attention since sialylation has also been implicated in regulating IgG activity. Increasing sialylation of the Fc glycans results in the decreased ADCC activity of IgGs, as terminal sialylation exhibits low affinity for the Fc gamma RIIIa receptor [120]. It has been reported that increased sialylation makes IgGs an anti-inflamatory agents [81, 121] and this was found to be essential for the function of therapeutic intravenous gamma globulins (IVIG) which together with their Fc fragment are anti-inflammatory [81, 122]. It is reported that one class of IgG Fc glycans through binding to FcγRs perform pro-inflammatory effects of immune complexes and cytotoxic antibodies [81], which are so the key mediators of many autoimmune diseases such as immune thrombocytopenia (ITP), autoimmune hemolitic anemia (AHA), systemic lupus erythematosus (SLE), reumatoid arthritis (RA), type I diabetes and multiple sclerosis [123, 124]. In contrast, therapeutic intravenous gamma globulins and its Fc fragments are anti-inflammatory [81]. Intravenous immunoglobulin (IVIG) is a therapeutic preparation of normal human polyclonal IgG obtained from plasma pooled from a large number of healthy blood donors. Due to its anti-inflammatory effect, IVIG is now widely used as a replacement therapy of patients suffering from autoimmune diseases such as: immune thrombocytopenia (ITP), chronic inflammatory demyelinating polyneuropathy (CIDP), and rheumatoid arthritis (RA) [125, 126]. Because anti-inflammatory activity of IVIG is maintained through enormous quantity of IVIg (1-2 g per kg), it was speculated that it could be possible that only a small sub-fraction of IVIg will mediate advantageus effect. Therefore, Kaneko *et al* [81] suggested that anti-inflammatory activity of IVIG is derived from sialylation of Fc region. Because a reduction in IVIG activity was seen only by desialylated IVIG, it was suggested that sialic acid could be the key sugar performing anti-inflammatory activity [81]. Consistently, in vitro sialylation or acid specific lectin enrichment generated sialylated IgG Fcs which in turn suppressed inflammation at a 30-fold lower dose than IVIG [127]. Also, sialylated IVIG suppressed induced arthritis and nephrotoxic nephritis at a 10 –fold lower dose than whole IVIG [81]. These studies confirmed that anti-inflammatory activity is a property of IgG Fc portion and that Fc portion alone is sufficient to suppress inflammation [128]. Further support for this concept was provided by data showing that only α2,6 linked sialic acid residues have been found to be responsible for the anti-inflammatory activity observed for human IVIG [120]. Moreover Ravetch and colleagues [129] pointed out the effect of sialylated Fc fragments as the anti-inflammatory mediator of IVIG in the K/N experimental arthritis mouse model. Contrary to other studies, it was found that induction of inhibitory IgG cell surface receptor FcγRIIB on effector macrophages by

sialylated IgG has profound effect on anti-inflammmatory activity of these antibodies [121]. Furtheremore, after administration of IVIG in a number of mouse models an increased expression of this inhibitory receptor FcγRIIB was observed and mice lacking this receptor were unresponsive to IVIG [122, 125, 130]. Since, human and mouse IgG glycoforms bearing terminal sialic acid residues show a reduced affinity for activating FcγRs, it seems clear that other receptors might be involved in recognizing sialic acid rich IgG [81]. Indeed, Anthony and colleagues [127] demonstrated that the mechanism that mediate such anti-inflammatory effect does not involve FcγRs and FcRn, but sialylated IgG mediate its anti-inflammatory activity through binding a C-type lectin receptor on mouse splenic macrophages SIGN-R1 or its human orethologue, DC-SIGN. Although, subsequent studies have supported this by suggesting a protective mechanism involving the inhibitory Fc receptor (FcγRIIb) and C-type lectin SIGNR1 or its human ortholog DC-SIGN [121, 131]. Because IVIG lost its therapeutic activity in FcγRIIB knock-out animals in models of ITP, nephrotoxic nephritis, and serum transfer arthritis [130, 132] it is obvious that the inhibitory FcγRIIB is also essential for IVIG activity.

The presence or absence of anntenary galactose residues has also been demonstrated to alter the activity of antibodies. IgG Fc-glycans carrying galactose residues in their teminal antennae bind to the complement component C1q giving rise to classical pathway of complement activation but not ADCC activity because they do not influence binding to the FcγRIIIa receptor [133]. The role of IgG galactosylation was analysed in different inflammatory diseases [77] and also similar observation were made in infectous diseases and cancer [134, 135]. Increased levels of IgG-G0 promots proinflammatory response against HIV virus and this argues for a functional significance of antibody galactosylation in humans [136]. Incomplete galactosylation of IgG exhibits a low affinity to bind to complement component C1q and FcγRs resulting in impaired effector functions [137] and in contrast to galactosylated IgG (G1-IgG and G2-IgG), agalactosylated IgG can activate *in vitro* complement system via the interaction of mannose-binding lectin (MBL) with Fc terminal GlcNAc residues [138] and thus be part of the underlying pathological mechanisms. Furthermore, an increase in terminal GlcNAc content results in decreased binding of antibody to C1q and reduced CDC activity [133]. The MBL is a C-type protein that is involved in clearance of immune complexes and in the lectin pathway of complement activation [139]. It forms a complex with structural homologues of C1 complement components and triggers CDC [138]. Based on the increased binding of MBL to agalactosylated glycans *in vitro*, it has been suggested that the MBL may contribute to an additional inflammation by activating complement [139]. In contrast, recent

studies performed in MBL knock-out mice showed that despite the capacity to bind MBL and activate the complement cascade *in vitro*, there was no significant contribution of MBL to the activity of IgG G0 antibodies *in vivo*, suggesting rather that IgG G0 antibodies work normally via the classical FcγR pathway [140]. Even though this study argues no functional role of interaction between MBL and G0, because of the different role of individual glycans and glycoforms in mouse and human, further studies must be performed in humans to elucidate the exact functional role of agalactosylation IgG antibodies in terms of interaction with MBL.

The pro-inflammatory and anti-inflammatory properties of IgG due to alternative glycosylation is presented in **(Figure 4)**.

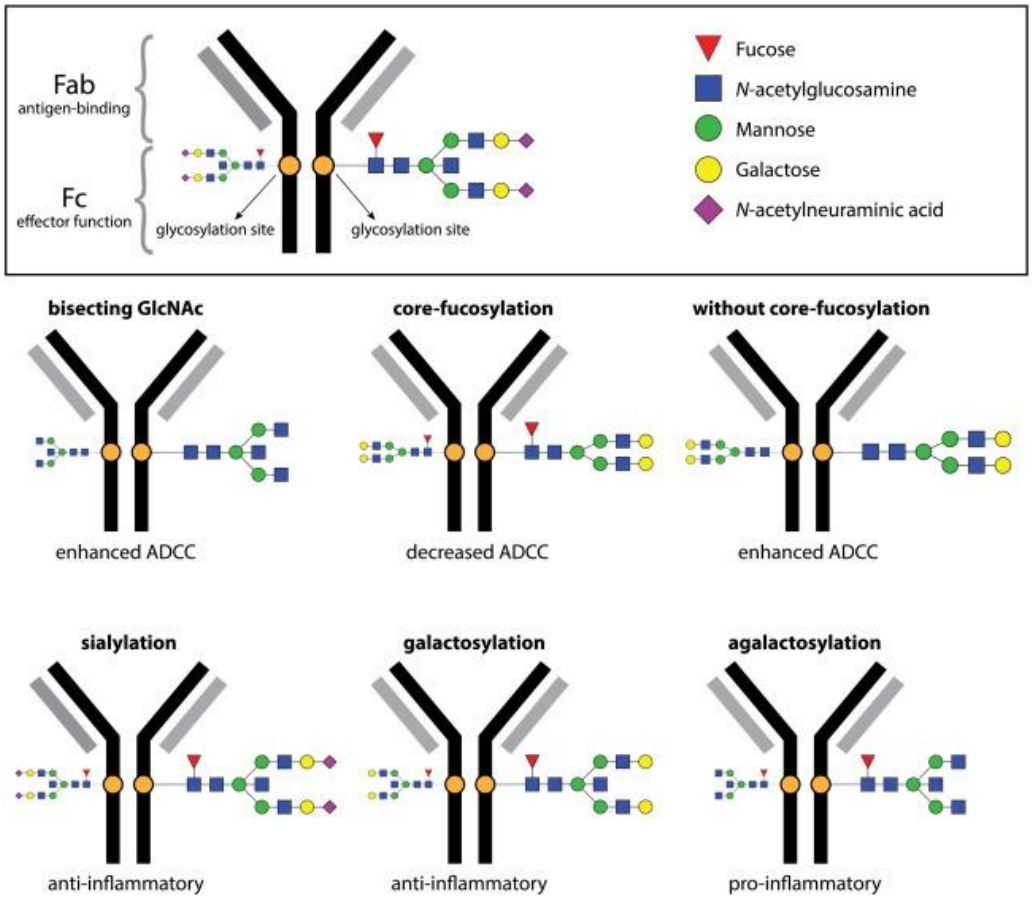


Figure 4. *The schematic representation of the human IgG structure and functional implication of alternative glycosylation* [141].

1.5. IgG glycosylation under physiologic and pathophysiologic conditions

Despite the biosynthetic complexity, glycosylation under controlled physiological conditions or during bioprocessing is highly reproducible. Therefore, it has been known for a long time that aberrant changes of IgG glycosylation are associated with specific physiological and pathophysiological conditions [64]. Recently, there has been an increasing interest in the analysis of the N-glycosylation profile of human IgG in health and in a number of diseases such as infections, inflammation and autoimmunity [36, 137, 142, 143, 144, 145, 146]. Although, variations in the glycome in cancerous state have been reported long time ago [29], less is known about the potential role of the Fc glycans in malignancy. Understanding the role of changes in glycosylation in autoimmune diseases and cancer is important for defining both molecular pathology of such conditions and also for the identification of possible targets for subsequent drug development [147, 148]. The mechanisms that cause disease-related glycosylation changes have not yet been clarified, but in order to better understand these mechanisms it is very inportant to define glycosylation changes as they relate to normal physiologic processses. Therefore, a number of studies have reported age and sex dependence of certain IgG glycosylation features in healthy population [149, 150, 79]**.** Parekh *et al.* [149]; Yamada *et al.* [150]; Shikata *et al.* [79]; Knežević *et al.* [17]; Ruhaak *et al.* [151]; Pučić *et al.* [20] in their studies have found a significant decrease of galactosylation and sialylation of IgG with age and an increase of glycan structures with the bisecting GlcNAc, while the core fucosylation of IgG oligosaccharides stayed rather stable over time. Yamada *et al.* [150], found different IgG glycosylation patterns related to gender, with agalactosylated levels of IgG antibodies being higher in both males and females but with males showing higher level of agalactosylated glycans than females in their twenties. Even though bisecting GlcNAc shows higher levels with age, however, differences between males and females were found.

The studies mentioned above provided valuable information regarding changes in glycosylation with respect to chronological age but the question whether IgG Fc glycans would reflect biological age was unknown for a long time, until Krištić *et al.* [152] performed a large-scale popullation study of 5,117 individuals from four European populations. This study revealed very extensive and complex changes in IgG glycosylation with age, and with a combination of several IgG glycans managed to explaine up to 58% of variance in chronological age, much more than other markers of biological age like telomere lengths. This was the first study confirming the close association of IgG glycans with both chronological and biological ages. High galactosylation and sialylation levels of IgG and

decrease of bisecting GlcNAc was reported in pregnacy [153]. In addition, glycosylation of IgG from both fetal and maternal circulation was also analyzed and compared. Interestingly, the incidence of IgG-G0 was found to be approximately 25% higher in maternal serum as compared to fetal umbilical vein serum, which might indicate a glycosylation sensitive transport of maternal IgG via the placenta to the fetal circulation with a selection for highly galactosylated glycoforms [154].

It seems that glycosylation varies considerably between individuals of the same sex and age, but the influence of life time on glycosylation is stroger than differences between sexes.

Among the most notable features of IgG glycosylation in pathologic conditions, galactosylation is the most studied. Increased levels of agalactosylated IgG have been reported in many autoimmune diseases suggesting that the absence of galactose may have a role in the pathogenity of the autoantibodies. In addiion, until now more than 50 different studies have analyzed the role of IgG galactosylation in different inflammatory diseases. Despite this, the molecular significance of these changes is still mostly unknown [64]. In sera of patients with rheumathoid arthritis (RA) a significant change in IgG glycosylation with increase levels of IgG-G0 was reported [143]. However, increased levels of agalactosylated IgG have been also found in sera of patients with juvenile arthritis, Crohn's diseas and tuberculosis [144, 155]. Young *et al* [156] have found that G0 values can be used clinically as an indicator of disease severity and also in the early diagnosis of rheumatoid arthritis. It was suggested that IgG molecules containing less terminal galactose bind better to certain rheumatoid factors (RA) [157]. Also, It was reported that IgG-G0 have pathogenic role in mouse model of arthritis [158] and interesingly, increased content of galactose in IgG carbohydrate chains was significantly increased during pregnancy, when re-mission of RA symptoms occurs [159]. In addition to RA, alteration in IgG glycosylation has been reported in studies of small vessel vasculitis, Wegener's granulomatosis, microscopic polyangitis and Churg-Strauss syndrome [142]. More recent studies described lower levels of IgG Fc galactosylation in well-defined antibody-mediated autoimmune diseases, i.e., Lambert-Eaton myasthenic syndrome (LEMS) and myasthenia gravis (MG) [135]. Changes of IgG glycosylation have been also reported in many other non-autoimmune diseases such as malignant states. Gercel-Taylor *et al.* [160] and Alley *et al.* [161] have found an elevated expression of agalactosylated IgG in the serum of patients with ovarian cancer. Increase in agalactosylation levels was also reported in patients with gastric cancer [148, 162], lung cance [163] and prostate cancer [164]. Recently, it was demonstrated that IgG glycosylation signature might be useful as a predictive marker for gastric cancer [165]. In contrast, Chen *et al.* [166] have reported a

higher levels of galactosylated IgG in female thyroid cancer patients than in female controls. Aberrant IgG Fc glycosylation was also reported in multiple mieloma with decreased galactosylation in comparison to age and sex matched controls [167].

These described changes of the overall IgG-Fc N-glycosylation, could provide insight into potential association of a certain glycosylation feature and antibody pathogenicity [135]. However less is known about the potential role of IgG glycosylation in malignancy and tumor immunity. Therefore, further analysis of IgG-Fc glycans may provide clearer understading of the possible impact of glycosylation on tumor progession and evaluation of the Fc glycans as a potentially predictive biomarker for monitoring of cancer patients. The functions of certain sugar residues for the pro and anti-inflammatory functions of IgG have shed light in sense of the functional importance of these sugar moieties attached to immunoglobulin molecule. The presence or absence of distinct sugar residues such as galactose, sialic acid and core fucose can dramatically alter IgG activity and changes in serum IgG glycosylation during age, pregnacy, autoimmune disease and cancer suggests existance of active regulatory mechanisms that could trigger a molecular switch keeping the humoral immune system in an active pro-inflammatory or a more anti-inflammatory state.

1.6. The prognostic importance of colorectal cancer

Colorectal cancer (CRC) is a malignant neoplasm of the colon and the rectum. It is the third most common malignancy in the world [168]. Unfortunately, nearly 800,000 new colorectal cancer cases are thought to globally occur each year, which account for 10% of all incident cancers, and the mortality from colorectal cancer is estimated at nearly 450,000 per year [46]. This places a major economic burden on the global health care system [169]. CRC develops through a progressive accumulation of genetic, both inherited and somatic,and epigenetic changes, leading to the conversion of normal colonic mucosa into invasive cancer. Almost 70 to 90 % of CRC-s derive from adenomatous polyps (adenomas) [170]. The risk of CRC increases with age especially after the age of 50 years, and the risk is also increased by certain inherited genetic mutations (familial adenomatouspolyposis and hereditary non-polyposis colorectal cancer), a personal or family history of colorectal neoplasia, or a personal history of inflammatory bowel diseases (IBD) [171]. Several modifiable factors are also associated with increased risk of CRC, including obesity, physical inactivity, smoking, heavy alcohol consumption, type II diabetes, and a diet high in red or processed meat.

CRC is still associated with poor prognosis, low survival rate and usually relatively late diagnosis. CRC develops slowly and the risk of recurrence and death from CRC is related to tumor stage at diagnosis. It continues to be such a serious health problem because it is largely asymptomatic until the latter stages oftentimes when the cancer has already metastasized. The growing repertoire of treatments available for CRC, including new chemotherapy approaches, combined with challenging benefit:toxicity ratios and cost, means that it is crucial to target interventions to patients most likely to benefit. Current cancer screening methods for CRC include fecal occult blood test followed by imaging procedures such as sigmoidoscopy, colonoscopy, or potentially computed tomography (CT). Although current screening methods are helpful, they are not optimal and present major concerns regarding sensitivity, specificity, complexity, cost, and compliance. Until now there are no reliable serologic markers available which would allow early diagnosis, monitoring and prognosis of patients. Because CRC develops slowly from removable precancerous lesions, detection of the disease at an early stage during regular health examination can reduce both the incidence and mortality of the disease [168]. Early diagnosis, including detection of adenomas, is considered to be a key aspect for improving patient survival and prognostic or predictive biomarkers are essential for guiding patient therapy or monitoring treatment efficacy. In addition, the importance of detecting colorectal cancer at an early stage is also appreciated by the fact that patients diagnosed with TNM Stage I disease have a 5-year survival rate of 90% following surgical resection [172], whereas when diagnosed at the later stag-es (i.e., Stage IV), the 5-year survival rate is only 5%. Nevertheles, advances in cancer treatment have increased survival for some cancers, yet the main CRC prognostic marker at present is stage at diagnosis as reflected by the AJCC/UICC TNM-classification which describes tumor spread through the bowel wall, number of affected lymph nodes and spread of tumor to distant organs (metastasis) [173]. In spite of this, even with the use of the TNM-classification method, there are some patients of lower TNM-stages that have a worse clinical outcome than patients of higher TNM-stages [174]. This is clinically important, since patients with AJCC stage 2 CRC may be offered adjuvant chemotherapy if their cancer is classified as high risk [175]. Hence, whilst pathological staging (TNM/Dukes') stratifies prognostic groups, it is limited in its ability to categorise poor/good prognosis tumours accurately and direct treatment decisions at the individual level. Also, current biomarkers which are at least partly released into the circulation from the tumor, perform poorly in terms of sensitivity and this greatly limits their value in cancer prognosis [168]. In line with this, currently, unfortunately scientists have failed to develop simple and non-invasive screening test for the early detection of cancers based only on genomic and

proteomic studies [176]. Therefore, the need for early detection and accurate diagnosis, prognosis and efficient therapeutic approach has led scientists to search further into the molecular level in order to identify novel biomarkers that could further refine pathology-based prognostic information and offer much potential for clinical and public health benefit, thus aiding the existing tumor classification systems in determining CRC prognostication.

A recent extensive report endorsed by the US National Academies concluded that "glycans are directly involved in the pathophysiology of every major disease" and that "additional knowledge from glycoscience will be needed to realize the goals of personalized medicine and to take advantage of the substantial investments in human genome and proteome research and its impact on human health" [177]. It is well known since a half century the relevance of alteration in glycosylation in disease states, particularly in the study of cancer [34] but glycans have only recently been introduced as potential disease biomarkers due to the development of a large-scale studies of the glycome. In general, immune response occurs in the early stage of tumor development and it may not only protect organisms against tumor development, but can also promote tumor development and progression by selecting variants with low immunogenicity. Inter-individual differences in immune repertoire and the capacity to process and present antigens is a key element in cancer immunosurveilance [178, 179]. Glycans have an enormous impact in the immune system and inter-individual variation in glycosylation may effect function in immune system in multiple levels [180], leading to inter-individual variability in cancer immunosurveilance and/or response to therapeutic antibodies, as therapeutic antibodies have been set up as 'standard of care' therapy for several highly prevalent human cancers [181]. Even though many serum glycoproteins display changes in glycosylation pattern, little is known about changes in IgG glycosylation in patients with cancer. Therefore understanding the role of altered IgG glycosylation in colorectal cancer offers the potential for the discovery of a new category of glycan based biomarkers which would have high discriminative power to predict cancer survival outcome and would help tailor the most appropriate surgical and chemotherapy regimens to individual patients thereby improving patient outcomes.

Measurement of plasma N-glycan biomarkers thus shows potential as a novel non-invasive approach to determine cancer prognosis. However, published studies have been limited to small scale pilot studies due to technical challenges in identifying and quantifying N-glycan structures. The N-glycome wide approach which was applied in this thesis has only recently been made possible due to recent developments in high-throughput glycan analysis [182]. This

study will be the first to investigate the IgG N-glycome and will employ a sample size which is substantially larger than current published studies.

2. HYPOTHESIS

N-glycomic analysis could reveal significant inter-individual differences and identify several specific glycan biomarkers which are associated with CRC prognosis.

3. AIMS AND PURPOSE OF THE RESEARCH

3.1. General aim:

This research aims to identify IgG N-glycans biomarkers with discriminative power to predict survival in patients with CRC.

3.2. Specific aims:

1. To analyse IgG N-glycans in 1229 patients with CRC and 538 matching controls.
2. To investigate the association between plasma IgG N-glycans and CRC survival and evaluate their potential role as clinical biomarkers for CRC prognosis.
3. To investigate the relationship between plasma levels of IgG N-glycans and CRC mortality.
4. To investigate the discriminative power of specific plasma IgG N-glycans as clinical biomarkers of CRC prognosis.

4. MATERIALS AND METHODOLOGY

4.1. Study population

The SOCCS (Study Of Colorectal Cancer in Scotland) study (1999-2006) is a case-control study designed to identify genetic and environmental factors associated with non-hereditary colorectal cancer risk and survival outcome. This study includes prospective CRC cases from almost all hospitals in Scotland therefore is broadly representative of the colorectal cancer population. The study has been described in details elsewhere [183].

Samples for the analyses were assembled from a large population-based resource including a subset of 1229 patients with pathologically confirmed colorectal adenocarcinoma and 538 matching controls from the University of Edinburg (age/gender/residence area matched) with extensive data about cases and controls including: lifestyle/ behavioural/ diet risk factors, tumor related parameters, ancestry, clinical data, and biological samples (DNA, lymphocytes, plasma). In addition to demographic data, detailed clinical data for each patient's CRC episode were collected and validated, which include: clinical and imaging data, date of diagnosis; presence/date of recurrence; new cancer occurrence data; date of death; cause of death; cancer staging assigned according to the American Joint Committee on Cancer (AJCC) system (all available clinical, imaging and pathology data were incorporated in final AJCC staging).

Cases were recruited as soon as possible after diagnosis to limit survival bias among those recruited and maximize the person-years of follow up IgG glycan composition was analysed in the time period after CRC diagnosis or recruitment. Participants completed one questionnaire with lifestyle and cancer information, reporting their status 1 year prior to diagnosis or recruitment. A semi-quantitative food frequency (http://www.foodfrequency.org) and supplements questionnaire was completed by participants. Participants were asked about their general medical history, physical activity, and smoking status. Additionally, subjects were asked to report any regular intake of aspirin and nonsteroidal anti-inflammatory drugs. Reported height, weight, and waist circumference were recorded. Participants were also asked to report some demographic, socioeconomic, and race/ethnicity data. Finally, women were asked about their menstrual and reproductive history and type of hormone replacement therapy and hormonal contraception, if used.

Blood was collected and transferred to the research centre within 72 h of sampling. Plasma was prepared from whole blood by gentle centrifugation of sodium EDTA tubes through a ficoll-hypaque gradient and 1.5 mL of each participant's plasma was stored at -80°C.

Effects of surgery on the IgG glycome were evaluated in 28 patients sampled before surgery, 24h, 48h and 7 days after surgery. Samples were collected at the Clinical Department for Laboratory Diagnostics at University Hospital Dubrava, Zagreb, Croatia.

Patients who were sampled before the initial diagnosis of CRC, as well as matching controls that did not develop CRC during the same follow-up time were selected from the FINNRISK cohort [184].

4.2. Ethics statement

Approval for the study of 1229 patients with CRC and 538 matching controls was obtained from the MultiCentre Research Ethics Committee for Scotland and Local Research Ethics committee.

The study of 28 patients with CRC samples before surgery was registrated at ClinicalTrials.gov, number NCT01244022 and was approved by the Ethics Committee of University Hospital Dubrava.

All participants gave written informed consent.

4.3. IgG N-glycans sample preparation and analyses

4.3.1 Purification of IgG

The IgG was isolated using protein G monolithic plates (BIA Separations, Ajdovščina, Slovenia) as described previously [185]. Before use, the monolithic plate was washed with 10 column volumes (CV) of ultra pure water and then equilibrated with 10 CV of binding buffer (1X PBS, pH 7.4). Plasma samples (70-100µl) were diluted 10 X with the binding buffer and applied to the Protein G plate. The filtration of the samples was completed in 5 min. The plate was then washed five times with 5 CV of binding buffer to remove unbound proteins. IgG was released from the protein G monoliths using 5 CV of elution solvent (0.1 M formic acid, pH 2.5). Eluates were collected in a 96-deep-well plate and immediately neutralized to pH 7.0

with neutralization buffer (1 M ammonium bicarbonate) to maintain the IgG stability. After each sample application, the monoliths were regenerated with the following buffers: 10 CV of 10 X PBS, followed by 10 CV of 0.1 M formic acid and afterward 10 CV of 1 X PBS to re-equilibrate the monoliths. Each step of the chromatographic procedure was done under vacuum (cca. 60 mmHg pressure reduction while applying the samples, 500 mmHg during elution and washing steps) using a manual set-up consisting of a multichannel pipet, a vacuum manifold (Beckman Coulter, Brea, CA) and a vacuum pump (Pall Life Sciences, Ann Arbor, MI). If the plate was not used for a short period, it was stored in 20% ethanol (v/v) at 4 °C. After repeated use of the plate contaminants present in the sample sometimes did not completely elute from the monolithic stationary phase. A specific cleaning protocol was developed that included washing with 0.1 M NaOH to remove precipitated proteins and with 30% propan-2-ol to remove strongly bound hydrophobic proteins or lipids. This procedure effectively removed all precipitates and did not significantly diminish IgG binding capacity of the immobilized protein G. The purity of the isolated IgG was verified by SDS-PAGE with NuPAGE Novex 4–12% Bis-Tris gels in an Xcell SureLock Mini-Cell (Invitrogen) according to the manufacturer. Precision Plus Protein All Blue Standards (BioRad, Hercules, CA) was used as the molecular weight marker. The gels were run at 180 V for 45 min, stained with GelCode Blue (Pierce) and visualized by a VersaDoc Imaging System (BioRad).

4.3.2 Release and labelling of IgG glycans

IgG samples were dried in a vacum concentrator, denatured with addition of 30 μL 1.33% SDS (w/v) (Invitrogen, Carlsbad, CA, USA) and by incubation at 65 °C for 10 min. After incubation samples were left to cool down to room temperature for 30 minutes. Subsequently, 10 μL of 4% Igepal-CA630 (Sigma-Aldrich) and 1.25 mU of PNGase F (ProZyme, Hayward, CA, USA) in 10 μL 5× PBS were added to the samples. The samples were incubated overnight at 37 °C for N-glycan release. The released N-glycans were labelled with 2-aminobenzamide (2-AB). The labelling mixture was freshly prepared by dissolving 2-AB (Sigma-Aldrich) in DMSO (Sigma-Aldrich) and glacial acetic acid (Merck) mixture (85:15, v/v) to a final concentration of 48 mg/mL. A volume of 25 μL of labelling mixture was added to each N-glycan sample in the 96-well plate. Also, 25 μL of freshly prepared reducing agent solution (106.96 mg/ml 2-picoline borane (Sigma-Aldrich) in DMSO) was added and the plate was sealed using adhesive tape. Mixing was achieved by shaking for 10 min, followed by 2 hour incubation at 65 °C. Samples (in a volume of 100 μL) were brought to 80% ACN (v/v) by adding 400 μL of ACN (J.T. Baker, Phillipsburg, NJ, USA).

4.3.3 HILIC-SPE

Free label and reducing agent were removed from the samples using hydrophilic interaction liquid chromatography-solid phase extraction (HILIC-SPE). An amount of 200 μL of 0.1 g/mL suspension of microcrystalline cellulose (Merck) in water was applied to each well of a 0.45 μm GHP filter plate (Pall Corporation, Ann Arbor, MI, USA). Solvent was removed by application of vacuum using a vacuum manifold (Millipore Corporation, Billerica, MA, USA). All wells were prewashed using 5× 200 μL water, followed by equilibration using 3× 200 μL acetonitrile/water (80:20, v/v). The samples were loaded to the wells. The wells were subsequently washed 7× using 200 μL acetonitrile/water (80:20, v/v). Glycans were eluted 2× with 100 μL of water and combined eluates were stored at –20 °C until usage.

4.3.4 HILIC-UPLC

Fluorescently labelled N-glycans were separated by hydrophilic interaction chromatography on a Waters Acquity UPLC instrument (Milford) consisting of a quaternary solvent manager, sample manager and a FLR fluorescence detector set with excitation and emission wavelengths of 330 and 420 nm, respectively. The instrument was under the control of Empower 2 software, build 2145 (Waters, Milford). Labelled N-glycans were separated on a Waters BEH Glycan chromatography column, 100 × 2.1 mm i.d., 1.7 μm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. Separation method used linear gradient of 75–62% acetonitrile (v/v) at flow rate of 0.4 ml/min in a 25 min analytical run. Samples were maintained at 5 °C before injection, and the separation temperature was 60 °C. The system was calibrated using an external standard of hydrolyzed and 2-AB labelled glucose oligomers from which the retention times for the individual glycans were converted to glucose units. Data processing was performed using an automatic processing method with a traditional integration algorithm after which each chromatogram was manually corrected to maintain the same intervals of integration for all the samples. The chromatograms were all separated in the same manner into 24 peaks and the amount of glycans in each peak was expressed as percentage (%) of total integrated area. In addition to 24 directly measured glycan structures, 53 derived traits were calculated. These derived traits average particular glycosylation features (galactosylation, fucosylation, sialylation) across different individual glycan structures. Consequently, they are more closely related to individual enzymatic activities, and underlying genetic polymorphisms.

4.4. Survival and risk related parameters

Data relevant to the survival analysis were obtained from the Scottish registries General Register Office and the Scottish Cancer Registry (which are known to have high levels of data quality and data completeness) after linkage of our participants with their databases using the Community Health Index number. Mortality outcomes were ascertained through linkage with the National Records of Scotland. Primary cause of death ("CRC" or "other") was assigned from death certificates separately by two researchers (concordance was >99%). Survival follow-up was censored at the date of death or at January, 31 2013, for participants who were not known to have died. Clinicopathological staging data was collected where possible (eg. TNM is not feasible in patients who did not undergo surgery). Clinical records were reviewed and tumor site and multiplicity were determined from clinical and pathological records. Pre-operative staging imaging was collected through participating centres. Using the collated pathology, imaging and clinical data, tumor stage was assigned according TNM staging system and mapped onto the American Joint Committee on Cancer (AJCC) tumor-node-metastasis system (AJCC 1-4).

Blood was collected at various points after the CRC diagnosis and after surgery. Median time to sampling was 5.4 months after the diagnosis (interquartile range, IQR: 3.2 to 8.3 months). Since illness and treatment may acutely affect IgG glycan levels and confound the analysis, a variable describing time from operation to blood collection and a variable determining the type of operation were created.

4.5. Statistical analysis

Data was analysed using STATA (version 12.0) and R.

Initially we examined the association between IgG glycan levels (continuous and quartiles) and CRC/all-cause mortality using cox proportional hazards models. Three models were applied: a crude model (Model I), a model where hazard ratios (HR) were adjusted for age at diagnosis, sex and stage of disease (Model II) and a model where HRs were adjusted for age at diagnosis, sex, stage of disease, body mass index (BMI), time from operation to blood collection, type of operation and CRP (Model III). P-values were adjusted for multiple testing using false discovery rate method (Benjamini–Hochberg procedure).

We estimated the predictive value of a clinical only Cox-regression algorithm for model II (which included age, sex, disease stage) and model III (which was adjusted for age, sex,

disease stage, BMI and CRP level), by calculating the Harrell's C concordance coefficient, the Area under the ROC curve (AUC) and the Integrated Discrimination Index (IDI defined as a difference in discrimination slopes) and compared this to an algorithm that also included glycan predictors.This analysis was ran in the whole data set and after AJCC stage stratification. The glycan variables included in the final model were selected by applying generalised boosted regression, which orders the variables by their relative importance, in 1000 bootstrap samples [186, 187], over the 10 inner training folds, and forward selection of ranked glycans by applying log-likelihood ratio test.

The predictive value of the models was evaluated on independent samples using 10-fold cross-validation for all models except for AJCC stage 1 strata, where cross-validation was not possible due to the small number of events. $P < 0.002$ was considered statistically significant, after applying the Bonferroni correction for 21 independent tests.

Classification analyses were performed to a) predict the 5-year risk of CRC death and b) to predict the rapid progressors within each stage. A rapid progressor was defined as someone who died of CRC and whose follow-up time was in the lower 1/3rd of the patients to die of CRC in that stage of cancer, with the cut-off thresholds at 2.9, 2.4, and 1.3 years for stages 2-4 respectively. We applied several families of classification models (LASSO, nearest neighbours, PAM, Support Vector Machines, naive Bayes, Decision Trees, and boosted stump classifiers), with and without stratification, with and without initial filtering on the training data, with and without log transformations of glycan expressions and clinical factors. The choice of the models was influenced by their popularity in biomarker studies, and their ability to address high-dimensional (large-p, small-n) problems via regularization or an explicit control for model complexity. More information about these estimators is presented in **Supplementary Box 1**, and the motivations for considering multiple classifiers for this problem are discussed in **Supplementary Section** on Model Comparison.

All the results for this analysis were averages over 10 runs using 10-fold cross validation, where the validation folds were used neither for filtering nor for estimation of model parameters. We also used 10 inner folds to estimate the stopping criteria or optimal value of hyperparameters (such as the regularization parameter for LASSO). Then we estimated whether adding glycans to clinical covariates would improve the predictive performance of a model of the same class on independent test data; that is, we compared LASSO using clinical variables with LASSO using clinical variables and glycans, DTs using clinical variables with DTs using clinical variables and glycans, etc. This task is different from the association

analysis, or from identifying specific glyco-clinical models outperforming a known baseline, where corrections for multiple tests are needed to control the probability of false discoveries. We applied the paired Wilcoxon sign-rank test comparing models with and without glycans, testing whether the difference in the cross-validated AUC of the clinical and glycol-clinical models is significantly different from zero.

Comparison of clinical characteristics among 760 patients and 538 matching controls was done by using Wilcoxon and Fischer Exact tests. Data was analyzed and visualized using R programming language (version 3.0.1). To make measurements across samples comparable, normalization by total area was performed where peak area of each of 24 glycan structures was divided by total area of corresponding chromatogram. Batch correction was performed on normalized log-transformed measurements using linear mixed models (R package lme4), where technical source of variation was modelled as random effect.

Association analyses between disease status and glycan traits were performed using a logistic regression model with age and sex included as additional covariates. Prior to analyses, glycan variables were all transformed to standard Normal distribution (mean=0,sd=1) by inverse transformation of ranks to Normality (R package "GenABEL", function rntransform). Using rank transformed variables in the analysis makes estimated odds ratios of different glycans comparable as transformed glycan variables have the same standardized variance, and in that case estimated odds ratios always correspond to one standard deviation change in the measured glycan trait. False discovery rate was controlled using Benjamini-Hochberg procedure [188]. For prediction of CRC status, regularized logistic (elastic net) regression model was applied (R package "glmnet"). For classification, only 24 initial glycan traits were used as predictors. Prior to model training and validation, elastic net regularization parameters (alpha and lambda) were tuned on 20% of samples (260 samples), and optimal parameters chosen by the tuning procedure (alpha = 0, lambda = 0.1) were used in further analysis. To evaluate performance of predictive model 10-cross validation procedure was used on remaining 80% of samples. Predictions from each validation round were merged into one validation set on which model performance was evaluated based on area under the receiver operating characteristic curve (AUC) criteria. The AUCs of different models were compared using a bootstrap test.

LASSO - sparse logisitic regression. This is a standard logistic regression model with a logit link function, and L_1 penalties on the weights. To set the penalty parameter, we run 10-fold cross-validation over a set of 10 penalties in the range 0.001 to 0.1. We choose the penalty resulting in the minimum-deviance model.

k-nearest neighbours - k nearest neighbours with k=1. For a new input, this classifier predicts the class of the nearest input in the training data, measured by Euclidean distance.

PAM - Prediction Analysis for Microarrays. This is the "nearest-shrunken-centroid" classifier. To set the shrinkage parameter, we run 10-fold cross-validation over 31 equally spaced values in the range 0 to 1.5. We choose the parameter that produces the lowest mean error rate.

Support Vector Machines – kernel-based maximum-margin classifiers. We consider linear, square, cubic, and squared exponential (Gaussian) kernel functions with fixed hyperparameters.

Naive Bayes – factorized class-conditional classifiers with normal or kernel density estimator-based marginal distributions. The KDE bandwidth is selected automatically.

Decision Trees – non-parametric tree classifiers, with the internal nodes corresponding to predictors, and leaves encoding classification labels. Binary trees with the Gini impurity splitting criterion were used.

Boosted stump classifiers – aggregations of multiple boosted one-node decision trees (stumps), where each later stump focuses on previously misclassified samples, using a version of Adaboost.

Supplementary Box 1. *Information about the estimators used in the classification analysis.*

4.5.1 Supplementary Section: Model Selection

One classifier vs multiple classifiers

In life and clinical sciences, it is common to analyse the predictive performance by using an arbitrarily chosen single regression or classification model such as linear regression for continuous outcomes or logistic regression for binary outcomes, without motivating the model choice. There are multiple models that may in principle be considered for continuous and binary outcomes, and deeper insights about the utility of biomarkers may potentially be obtained by evaluating many such models. We note that the analysis of the predictive performance based on a single model may be misleading, due to the following observations: (i) It may happen that by considering a single model, researchers observe that biomarkers do not improve the quality of predictions. But this observation may be an artefact of the implied modelling constraints (such as the linear decision surface separating cases from controls in logistic regression). One reason for failing to demonstrate an improvement in predictions may be the fact that the chosen predictive model was limited and inappropriate for the dataset. The biomarkers may still be useful predictors, but the researchers may be making incorrect assumptions about the data and using a wrong model, without trying to evaluate whether the modelling assumptions are correct. (ii) A similar argument may hold for a subset of variables. For example, researchers may be able to demonstrate that a common model such as logistic regression with covariates defined by biomarkers and clinical variables outperforms logistic regression that only uses clinical variables, and may conclude that the biomarkers are generally useful for predicting the considered outcome. However, it may happen that the logistic assumption is particularly unfavourable to the clinical model (for example, when the mapping from the clinical variables to the outcomes is complex, and the classification surface cannot be well modelled by a hyperplane in the subspace of clinical variables). In this case, a clinical model of some other class (for example, an SVM) that does not use the biomarkers could significantly outperform models with biomarkers. In this case, the conclusion that the biomarkers are useful for developing a diagnostics, may be misleading - one may be able to achieve a superior quality of predictions when considering "richer" clinical models (something overlooked by considering model of a single class). We note that the assessment of the predictive performance by considering a single model may often be limited, and the results may need to be interpreted with some care. This work is an empirical attempt to overcome the arbitrariness of a specific model choice. In particular, we considered a larger set of models that make different assumptions about the mapping from glycans to outcomes. We

use (nested) cross-validation to estimate the predictive performance on new previously unseen data. We then compare pairs of models of the same class that use clinical variables only and clinical variables with glycans, and test how likely it is that using glycans for predictions leads to improvements over clinical models independently of the modelling assumptions. (Note that the models are generally not nested even when they belong to the same class – so we cannot use standard tests). As the evaluation criterion, we use the AUC computed by cross-validation over the test folds of data. In some sense, rather than comparing an arbitrary model with or without glycans, we are evaluating how easy it may be to use glycomic biomarkers to construct a superior diagnostic independently of the modelling details.

5. RESULTS

5.1. IgG glycan measurements in colorectal cancer prognosis

IgG glycome composition was analysed in 1229 patients with CRC and 538 matching controls. Total IgG glycan (combined Fc and Fab glycans) compositions were determined by HILIC-UPLC analysis of 2AB labelled glycans as reported recently [189].

1. To examine the potential role of individual variation in IgG glyosylation on CRC prognosis we initially performed detailed analyses of IgG glycome composition in 1229 patients with CRC of all CRC stages.
2. In addition, we explored the prognostic biomarker potential of IgG glycans after stage stratification to account for the different stage prognosis of CRC patients.

Total IgG glycan measurements resulted in 24 chromatogaphic peaks (GP1-GP24) that were directly measured glycan structures, and 53 derived traits that represent common features shared among several measured glycans (galactosylation, sialylation, core fucosylation and the incidence of bisecting GlcNAc. **(Table 1)** [20, 185]. A typical chromatogram showing seperation of the IgG glycome into individual structures is shown in **(Figure 5).**
We restricted our survival analysis to those IgG glycan traits that were found to be robustly analysed. Robustness was calculated as follows. On each plate from the CRC cohort we put 3 standards that were biologically identical. Therefore, differences between measurements of standards are consequence of only experimental noise. We then calculated the variance of standards only and the variance in the whole CRC population. "Robustness" is defined as the ratio of those two variances (Var(Stand)/Var(CRC))*100 (i.e. lower values indicate higher robustness) and represents the contribution of experimental variation in total variation.

Thirty nine of the 77 glycan traits whose percentage of experimental variation was below 20% were included in the analysis **(Table 1).**

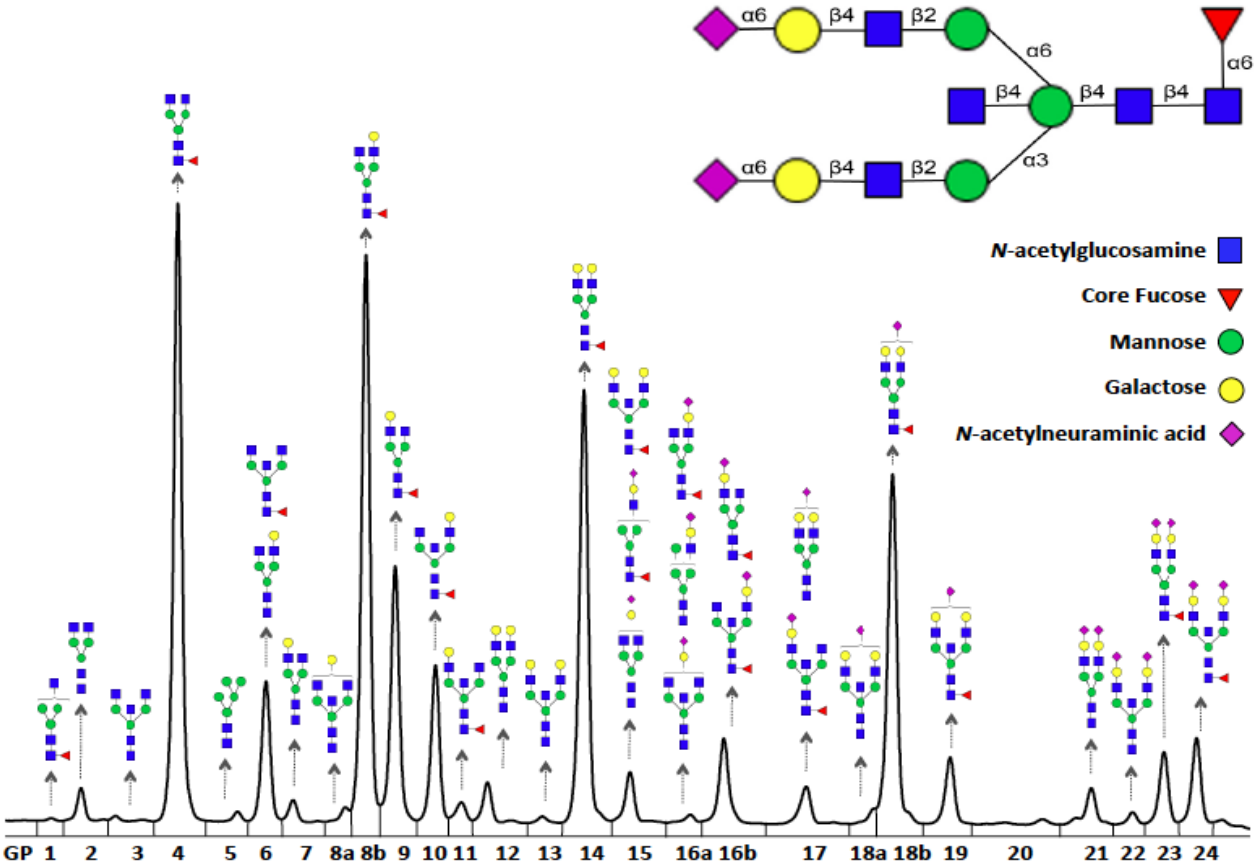


Figure 5. HILIC-UPLC chromatogram of IgG n-glycans. *Each IgG contains one conserved N-glycosylation site on Asn197 of its heavy chains. Different glycans can be attached to this site and the process seems to be highly regulated. UPLC analysis can reveal composition of the glycome attached to a population of IgG molecules by separating total IgG N-glycome into 24 chromatographic glycan peaks (GP1–GP24), mostly corresponding to individual glycan structures.*

The amount of glycans in each peak is expressed as percentage of total integrated area. Glycoforms represented by each peak were analyzed on the basis of their elution positions and measured in glucose units then compared to reference values in the "GlycoBase" database for structure assignment [190]. GU units were assigned according to 2AB labeled standard of glucose oligomers of different sizes. Value for each peak is presented as the area percentage of the total glycome of the analyzed sample.

Table 1. Glycans annotations and experimental variation for each glycan variable

GROUP	Code	Glycan Peaks and derived traits Code	Description	Formula	Variation
Total glycans (neutral + charged)	IGP1	GP1	*The percentage of FA1 glycan in total IgG glycans*	*GP1 / GP * 100*	34.35
	IGP2	GP2	*The percentage of A2 glycan in total IgG glycans*	*GP2 / GP * 100*	34.02
	IGP3	GP4	*The percentage of FA2 glycan in total IgG glycans*	*GP4 / GP * 100*	2.45
	IGP4	GP5	*The percentage of M5 glycan in total IgG glycans*	*GP5 / GP * 100*	32.90
	IGP5	GP6	*The percentage of FA2B glycan in total IgG glycans*	*GP6 / GP * 100*	1.78
	IGP6	GP7	*The percentage of A2G1 glycan in total IgG glycans*	*GP7 / GP * 100*	29.37
	IGP7	GP8	*The percentage of FA2[6]G1 glycan in total IgG glycans*	*GP8 / GP * 100*	10.82
	IGP8	GP9	*The percentage of FA2[3]G1 glycan in total IgG glycans*	*GP9 / GP * 100*	6.71
	IGP9	GP10	*The percentage of FA2[6]BG1 glycan in total IgG glycans*	*GP10 / GP * 100*	2.42
	IGP10	GP11	*The percentage of FA2[3]BG1 glycan in total IgG glycans*	*GP11 / GP * 100*	31.34
	IGP11	GP12	*The percentage of A2G2 glycan in total IgG glycans*	*GP12 / GP * 100*	28.10
	IGP12	GP13	*The percentage of A2BG2 glycan in total IgG glycans*	*GP13 / GP * 100*	95.50
	IGP13	GP14	*The percentage of FA2G2 glycan in total IgG glycans*	*GP14 / GP * 100*	1.05
	IGP14	GP15	*The percentage of FA2BG2 glycan in total IgG glycans*	*GP15 / GP * 100*	12.56
	IGP15	GP16	*The percentage of FA2G1S1 glycan in total IgG glycans*	*GP16 / GP * 100*	33.19
	IGP16	GP17	*The percentage of A2G2S1 glycan in total IgG glycans*	*GP17/ GP * 100*	112.42
	IGP17	GP18	*The percentage of FA2G2S1 glycan in total IgG glycans*	*GP18 / GP * 100*	2.19
	IGP18	GP19	*The percentage of FA2BG2S1 glycan in total IgG glycans*	*GP19 / GP * 100*	15.42
	IGP19	GP20	*Structure not determined*	*GP20 / GP * 100*	104.69
	IGP20	GP21	*The percentage of A2G2S2 glycan in total IgG glycans*	*GP21 / GP * 100*	49.60
	IGP21	GP22	*The percentage of A2BG2S2 glycan in total IgG glycans*	*GP22 / GP * 100*	96.00
	IGP22	GP23	*The percentage of FA2G2S2 glycan in total IgG glycans*	*GP23 / GP * 100*	25.62
	IGP23	GP24	*The percentage of FA2BG2S2 glycan in total IgG glycans*	*GP24 / GP * 100*	30.39

Total IgG glycans - derived parameters	IGP24	FGS/(FG+FGS)	*The percentage of sialylation of fucosylated galactosylated structures without bisecting GlcNAc in total IgG glycans*	*SUM(GP16 + GP18 + GP23) / SUM(GP16 + GP18 + GP23 + GP8 + GP9 + GP14) * 100*	18.10
	IGP25	FBGS/(FBG+FBGS)	*The percentage of sialylation of fucosylated galactosylated structures with bisecting GlcNAc in total IgG glycans*	*SUM(GP19 + GP24) / SUM(GP19 + GP24 + GP10 + GP11 + GP15) * 100*	13.28
	IGP26	FGS/(F+FG+FGS)	*The percentage of sialylation of all fucosylated structures without bisecting GlcNAc in total IgG glycans*	*SUM(GP16 + GP18 + GP23) / SUM(GP16 + GP18 + GP23 + GP4 + GP8 + GP9 + GP14) * 100*	8.10
	IGP27	FBGS/(FB+FBG+FBGS)	*The percentage of sialylation of all fucosylated structures with bisecting GlcNAc in total IgG glycans*	*SUM(GP19 + GP24) / SUM(GP19 + GP24 + GP6 + GP10 + GP11 + GP15) * 100*	13.02
	IGP28	FG1S1/(FG1+FG1S1)	*The percentage of monosialylation of fucosylated monogalactosylated structures without bisecting GlcNAc in total IgG glycans*	*GP16 / SUM(GP16 + GP8 + GP9) * 100*	48.95
	IGP29	FG2S1/(FG2+FG2S1+FG2S2)	*The percentage of monosialylation of fucosylated digalactosylated structures without bisecting GlcNAc in total IgG glycans*	*GP18 / SUM(GP18 + GP14 + GP23) * 100*	9.74
	IGP30	FG2S2/(FG2+FG2S1+FG2S2)	*The percentage of disialylation of fucosylated digalactosylated structures without bisecting GlcNAc in total IgG glycans*	*GP23 / SUM(GP23 + GP14 + GP18) * 100*	22.02
	IGP31	FBG2S1/(FBG2+FBG2S1+FBG2S2)	*The percentage of monosialylation of fucosylated digalactosylated structures with bisecting GlcNAc in total IgG glycans*	*GP19 / SUM(GP19 + GP15 + GP24) * 100*	9.54
	IGP32	FBG2S2/(FBG2+FBG2S1+FBG2S2)	*The percentage of disialylation of fucosylated digalactosylated structures with bisecting GlcNAc in total IgG glycans*	*GP24 / SUM(GP24 + GP15 + GP19) * 100*	44.66
	IGP33	$F^{total}S1/F^{total}S2$	*Ratio of all fucosylated monosialylated and disialylated structures (+/- bisecting GlyNAc) in total IgG glycans*	*SUM(GP16 + GP18 + GP19) / SUM(GP23 + GP24)*	34.85
	IGP34	FS1/FS2	*Ratio of fucosylated monosialylated and disialylated structures (without bisecting GlcNAc) in total IgG glycans*	*SUM(GP16 + GP18) / GP23*	29.15
	IGP35	FBS1/FBS2	*Ratio of fucosylated monosialylated and disialylated structures (with bisecting GlcNAc) in total IgG glycans*	*GP19 / GP24*	45.81

	IGP36	FBS^{total}/FS^{total}	*Ratio of all fucosylated sialylated structures with and without bisecting GlcNAc in total IgG glycans*	*SUM(GP19 + GP24) / SUM(GP16 + GP18 + GP23)*	6.69
	IGP37	FBS1/FS1	*Ratio of fucosylated monosialylated structures with and without bisecting GlcNAc in total IgG glycans*	*GP19 / SUM(GP16 + GP18)*	5.67
	IGP38	FBS1/(FS1+FBS1)	*The incidence of bisecting GlcNAc in all fucosylated monosialylated structures in total IgG glycans in total IgG glycans*	*GP19 / SUM(GP16 + GP18 + GP19)*	5.48
	IGP39	FBS2/FS2	*Ratio of fucosylated disialylated structures with and without bisecting GlcNAc in total IgG glycans*	*GP24 / GP23*	3.89
	IGP40	FBS2/(FS2+FBS2)	*The incidence of bisecting GlcNAc in all fucosylated disialylated structures in total IgG glycans*	*GP24 / SUM(GP23 + GP24)*	4.44
Neutral IgG glycans	IGP41	$GP1^n$	*The percentage of FA1 glycan in total neutral IgG glycans (GP^n)*	*$GP1 / GP^n * 100$*	35.92
	IGP42	$GP2^n$	*The percentage of A2 glycan in total neutral IgG glycans (GP^n)*	*$GP2 / GP^n * 100$*	37.40
	IGP43	$GP4^n$	*The percentage of FA2 glycan in total neutral IgG glycans (GP^n)*	*$GP4 / GP^n * 100$*	1.27
	IGP44	$GP5^n$	*The percentage of M5 glycan in total neutral IgG glycans (GP^n)*	*$GP5 / GP^n * 100$*	36.24
	IGP45	$GP6^n$	*The percentage of FA2B glycan in total neutral IgG glycans (GP^n)*	*$GP6 / GP^n * 100$*	1.02
	IGP46	$GP7^n$	*The percentage of A2G1 glycan in total neutral IgG glycans (GP^n)*	*$GP7 / GP^n * 100$*	32.32
	IGP47	$GP8^n$	*The percentage of FA2[6]G1 glycan in total neutral IgG glycans (GP^n)*	*$GP8 / GP^n * 100$*	2.69
	IGP48	$GP9^n$	*The percentage of FA2[3]G1 glycan in total neutral IgG glycans (GP^n)*	*$GP9 / GP^n * 100$*	2.06
	IGP49	$GP10^n$	*The percentage of FA2[6]BG1 glycan in total neutral IgG glycans (GP^n)*	*$GP10 / GP^n * 100$*	0.91

	IGP50	$GP11^n$	*The percentage of FA2[3]BG1 glycan in total neutral IgG glycans (GP^n)*	*$GP11 / GP^n * 100$*	37.96
	IGP51	$GP12^n$	*The percentage of A2G2 glycan in total neutral IgG glycans (GP^n)*	*$GP12 / GP^n * 100$*	29.22
	IGP52	$GP13^n$	*The percentage of A2BG2 glycan in total neutral IgG glycans (GP^n)*	*$GP13 / GP^n * 100$*	92.53
	IGP53	$GP14^n$	*The percentage of FA2G2 glycan in total neutral IgG glycans (GP^n)*	*$GP14 / GP^n * 100$*	0.58
	IGP54	$GP15^n$	*The percentage of FA2BG2 glycan in total neutral IgG glycans (GP^n)*	*$GP15 / GP^n * 100$*	13.13
Neutral IgG glycans - derived parameters	IGP55	$G0^n$	*The percentage of agalactosylated structures in total neutral IgG glycans*	*$SUM(GP1^n: GP4^n + GP6^n)$*	1.24
	IGP56	$G1^n$	*The percentage of monogalactosylated structures in total neutral IgG glycans*	*$SUM(GP7^n: GP11^n)$*	1.09
	IGP57	$G2^n$	*The percentage of digalactosylated structures in total neutral IgG glycans*	*$SUM(GP12^n: GP15^n)$*	1.51
	IGP58	$F^{n\ total}$	*The percentage of all fucosylated structures (+/- bisecting GlcNAc) in total neutral IgG glycans*	*$SUM(GP1^n + GP4^n + GP6^n + GP8^n + GP9^n + GP10^n + GP11^n + GP14^n + GP15^n)$*	44.04
	IGP59	$FG0^{n\ total}/G0^n$	*The percentage of fucosylation of agalactosylated structures in total neutral IgG glycans*	*$SUM(GP1^n + GP4^n + GP6^n) / G0^n * 100$*	40.29
	IGP60	$FG1^{n\ total}/G1^n$	*The percentage of fucosylation of monogalactosylated structures in total neutral IgG glycans*	*$SUM(GP8^n + GP9^n + GP10^n + GP11^n) / G1^n * 100$*	20.74
	IGP61	$FG2^{n\ total}/G2^n$	*The percentage of fucosylation of digalactosylated structures in total neutral IgG glycans*	*$SUM(GP14^n + GP15) / G2^n * 100$*	81.14
	IGP62	F^n	*The percentage of fucosylated structures (without bisecting GlcNAc) in total neutral IgG glycans*	*$SUM(GP1^n + GP4^n + GP8^n + GP9^n + GP14^n)$*	11.33
	IGP63	$FG0^n/G0^n$	*The percentage of fucosylation of agalactosylated structures (without bisecting GlcNAc) in total neutral IgG glycans*	*$SUM(GP1^n + GP4^n) / G0^n * 100$*	4.70

	IGP64	$FG1^n/G1^n$	*The percentage of fucosylation of monogalactosylated structures (without bisecting GlcNAc) in total neutral IgG glycans*	$SUM(GP8^n + GP9^n) / G1^n * 100$	5.89
	IGP65	$FG2^n/G2^n$	*The percentage of fucosylation of digalactosylated structures (without bisecting GlcNAc) in total neutral IgG glycans*	$GP14^n / G2^n * 100$	57.82
	IGP66	FB^n	*The percentage of fucosylated structures (with bisecting GlcNAc) in total neutral IgG glycans*	$SUM(GP6^n + GP10^n + GP11^n + GP15^n)$	2.27
	IGP67	$FBG0^n/G0^n$	*The percentage of fucosylation of agalactosylated structures (with bisecting GlcNAc) in total neutral IgG glycans*	$GP6^n / G0^n * 100$	1.06
	IGP68	$FBG1^n/G1^n$	*The percentage of fucosylation of monogalactosylated structures (with bisecting GlcNAc) in total neutral IgG glycans*	$SUM(GP10^n + GP11^n) / G1^n * 100$	2.70
	IGP69	$FBG2^n/G2^n$	*The percentage of fucosylation of digalactosylated structures (with bisecting GlcNAc) in total neutral IgG glycans*	$GP15^n / G2^n * 100$	21.03
	IGP70	FB^n/F^n	*Ratio of fucosylated structures with and without bisecting GlcNAc in total neutral IgG glycans*	$FB^n / F^n * 100$	3.39
	IGP71	$FB^n/F^{n\ total}$	*The incidence of bisecting GlcNAc in all fucosylated structures in total neutral IgG glycans*	$FB^n / F^{n\ total} * 100$	3.43
	IGP72	$F^n/(B^n + FB^n)$	*Ratio of fucosylated non-bisecting GlcNAc structures and all structures with bisecting GlcNAc in total neutral IgG glycans*	$F^n/(GP13^n + FB^n)$	6.00
	IGP73	$B^n/(F^n + FB^n)$	*Ratio of structures with bisecting GlcNAc and all fucosylated structures (+/- bisecting GlcNAc) in total neutral IgG glycans*	$GP13^n / (F^n + FB^n) * 1000$	92.61
	IGP74	$FBG2^n/FG2^n$	*Ratio of fucosylated digalactosylated structures with and without bisecting GlcNAc in total neutral IgG glycans*	$GP15^n/GP14^n$	28.12
	IGP75	$FBG2^n/(FG2^n + FBG2^n)$	*The incidence of bisecting GlcNAc in all fucosylated digalactosylated structures in total neutral IgG glycans*	$GP15^n/(GP14^n + GP15^n) * 100$	28.43
	IGP76	$FG2^n/(BG2^n + FBG2^n)$	*Ratio of fucosylated digalactosylated non-bisecting GlcNAc structures and all digalactosylated structures with bisecting GlcNAc in total neutral IgG glycans*	$GP14^n/(GP13^n + GP15^n)$	52.69
	IGP77	$BG2^n/(FG2^n + FBG2^n)$	*Ratio of digalactosylated structures with bisecting GlcNAc and all fucosylated digalactosylated structures (+/- bisecting GlcNAc) in total neutral IgG glycans*	$GP13^n/(GP14^n + GP15^n) * 1000$	106.54

5.2. Survival analysis

Among the 1229 patients, there were 9563 person-years of follow-up. There were 489 deaths, including 385 from CRC.Median follow-up was 9.4 years (IQR: 4.4 to 10.6 years) overall, and 10.3 years (IQR: 9.6 to 11.0 years) for live patients.

Summary statistics and univariate Cox regression analysis for the confounding factors that were included in the subsequent glycan analysis are presented in **(Table 2).** Of them stage at diagnosis and post-surgery CRP levels were strongly associated with all-cause and CRC-specific mortality (all-cause mortality: stage 3 vs. stage 1 OR (95% CI): 2.65 (1.96, 3.59), p-value $3.0x10^{-10}$; stage 4 vs. stage 1 OR (95% CI): 14.32 (10.37, 19.77), p-value $8.1x10^{-19}$; CRP levels >10 mg/l vs. ≤10mg/l OR(95% CI): 2.13 (1.67, 2.72), p-value $1.1x10^{-9}$).

Age at diagnosis, sex and site of cancer (colon or rectum) were not associated with all-cause or CRC-specific mortality.

Table 2. Summary statistics and univariate Cox regression for factors influencing all-cause and CRC mortality.

All-cause mortality	Deceased Cases N=489	Survived/ Censored cases N=740	p-value	HR (95% CI)	p-value
Mean age (SD)	59.94 (10.15)	58.59 (9.87)	0.02	1.01 (1.00, 1.02)	0.09
Sex					
Men	287	416		1.00	
Women	202	324	0.39	0.96 (0.80, 1.15)	0.67
AJCC stage					
1	54	195		1.00	
2	115	306		1.35 (0.98, 1.87)	0.07
3	186	227		2.65 (1.96, 3.59)	$3.0x10^{-10}$
4	134	12	$<10^{-5}$	14.32 (10.37, 19.77)	$8.3x10^{-19}$
Site					
Colon	263	430		1.00	
Rectum	223	304	0.12	1.14 (0.95, 1.36)	0.15
Mean BMI(SD)[1]	26.86 (4.77)	26.32 (4.09)	0.05	1.03 (1.01, 1.05)	0.02
Mean CRP(SD)	5.99 (14.24)	2.54 (8.88)	$<10^{-5}$	1.02 (1.01, 1.02)	$1.2x10^{-11}$
CRP					
≤10mg/l	412	687			
>10 mg/l	77	53	$<10^{-5}$	2.13 (1.67, 2.72)	$1.1x10^{-9}$
CRC mortality	**Deceased Cases N=385**	**Survived/ Censored cases N=844**	**p-value**	**HR (95% CI)**	**p-value**
Mean age (SD)[1]	58.81 (10.23)	59.27 (9.89)	0.46	1.00 (0.99, 1.01)	0.48
Sex					
Men	212	491		1.00	
Women	173	353	0.31	1.12 (0.91, 1.36)	0.28
AJCC stage					
1	22	227		1.00	
2	76	345		2.17 (1.35, 3.49)	0.001
3	159	254		5.40 (3.46, 8.43)	$1.3x10^{-13}$
4	128	18	$<10^{-5}$	30.63 (19.38, 48.40)	$1.2x10^{48}$
Site					
Colon	205	488		1.00	
Rectum	178	349	0.12	1.16 (0.95, 1.42)	0.15

Mean BMI (SD)[2]	27.08 (4.88)	26.29 (4.12)	0.007	1.04 (1.01, 1.06)	0.002
Mean CRP(SD)	6.57 (15.24)	2.71 (8.98)	$<10^{-5}$	1.02 (1.01, 1.03)	$4.5x10^{-12}$
CRP					
≤10mg/l	318	781		1.00	
>10 mg/l	67	63	$<10^{-5}$	2.37 (1.82, 3.09)	$1.4x10^{-10}$

BMI available for 1057 CRC cases (415 CRC deaths and 642 survived / censored).

The univariate glycan HRs for the whole sample are presented in **(Table 3 and Table 4)** for all-cause mortality and in **(Table 5 and Table 6)** for CRC specific mortality.

Table 3. All cause analysis for rank transformed glycans. Q value represents the adjusted p-values using the false discovery rate method (Benjamini–Hochberg procedure).

Glycan	Dead (N=489)	Survived (N=740)	Model II (AJCC, age, sex, time between sample and surgery, operation type, BMI, CRP, n=952)		
	Mean (SD)	Mean (SD)	HR (95% CI)	p-value	q-value
Total IgG glycans (neutral and charged); Measured					
GP4	26.39 (7.33)	23.87 (6.37)	1.28 (1.14, 1.43)	$2.6x10^{-5}$	0.0008
GP6	6.34 (1.88)	5.79 (1.57)	1.19 (1.06, 1.34)	0.002	0.05
GP8	18.26 (2.21)	18.67 (1.81)	0.87 (0.78, 0.97)	0.01	0.18
GP9	9.38 (1.36)	9.83 (1.35)	0.83 (0.75, 0.92)	0.0003	0.009
GP10	5.45 (1.23)	5.48 (1.15)	0.93 (0.84, 1.04)	0.2	0.74
GP11	10.19 (3.13)	11.44 (3.13)	0.77 (0.68, 0.86)	$9.6x10^{-6}$	0.0004
GP15	1.38 (0.41)	1.50 (0.43)	0.86 (0.77, 0.96)	0.007	0.133
GP18	7.78 (2.29)	8.47 (2.34)	0.79 (0.71, 0.89)	$4.9x10^{-5}$	0.002
GP19	1.87 (0.38)	1.90 (0.39)	0.98 (0.89, 1.08)	0.7	0.74
Sialylation; Derived					
FGS/(FG+FGS)	24.78 (3.20)	24.95 (3.10)	0.93 (0.84, 1.03)	0.18	0.74
FBGS/(FBG+FBGS	32.83 (6.17)	32.58 (6.34)	1.04 (0.94, 1.15)	0.47	0.74
FGS/(F+FG+FGS)	16.35 (3.63)	17.35 (3.61)	0.81 (0.72, 0.90)	0.0001	0.003
FBGS/(FB+FBG+FBGS)	21.21 (4.88)	21.82 (4.96)	0.95 (0.86, 1.05)	0.3	0.74
FG2S1/(FG2+FG2S1+FG2S2)	40.18 (2.99)	39.51 (2.77)	1.06 (0.95, 1.18)	0.28	0.74
FBG2S1/(FBG2+FBG2S1+FBG2S2)	37.02 (3.87)	36.41 (3.97)	1.12 (1.01, 1.24)	0.03	0.51
Bisecting GlcNAc; Derived					
FBS^{total}/FS^{total}	0.30 (0.08)	0.28 (0.07)	1.19 (1.06, 1.33)	0.002	0.05
FBS1/FS1	0.17 (0.05)	0.16 (0.05)	1.18 (1.06, 1.32)	0.003	0.07
FBS1/(FS1+FBS1)	0.14 (0.04)	0.14 (0.03)	1.19 (1.06, 1.32)	0.002	0.05
FBS2/FS2	1.35 (0.32)	1.27 (0.30)	1.17 (1.05, 1.31)	0.005	0.11

FBS2/(FS2+FBS2)	0.56 (0.06)	0.55 (0.06)	1.17 (1.05, 1.30)	0.005	0.11
Neutral IgG glycans; Measured					
$GP4^n$	32.33 (8.02)	29.56 (7.02)	1.28 (1.14, 1.43)	$2.6x10^{-5}$	0.0008
$GP6^n$	7.78 (2.07)	7.18 (1.76)	1.17 (1.05, 1.31)	0.006	0.12
$GP8^n$	22.59 (3.09)	23.32 (2.52)	0.83 (0.74, 0.92)	0.001	0.028
$GP9^n$	11.60 (1.81)	12.28 (1.75)	0.79 (0.71, 0.88)	$1.5x10^{-5}$	0.0005
$GP10^n$	6.73 (1.54)	6.85 (1.46)	0.90 (0.81, 1.01)	0.06	0.74
$GP14^n$	12.70 (4.29)	14.39 (4.37)	0.76 (0.68, 0.86)	$7.4x10^{-6}$	0.0003
$GP15^n$	1.72 (0.54)	1.89 (0.59)	0.85 (0.76, 0.95)	0.003	0.07
Galactosylation; Derived					
$G0^n$	41.15 (9.19)	37.69 (8.09)	1.31 (1.16, 1.47)	$5.5x10^{-6}$	0.0002
$G1^n$	42.72 (4.71)	44.26 (3.56)	0.78 (0.69, 0.87)	$1.3x10^{-5}$	0.0005
$G2^n$	15.65 (5.00)	17.61 (5.15)	0.78 (0.69, 0.87)	$2.4x10^{-5}$	0.0008
Core fucosylation and bisecting GlcNAc; Derived					
F^n	79.47 (3.74)	79.77 (3.46)	0.96 (0.86, 1.07)	0.48	0.74
$FG0^n/G0^n$	78.81 (4.29)	78.69 (4.11)	1.04 (0.93, 1.16)	0.46	0.74
$FG1^n/G1^n$	80.03 (3.79)	80.45 (3.56)	0.94 (0.85, 1.05)	0.31	0.74
FB^n	17.29 (3.20)	16.97 (2.90)	1.02 (0.91, 1.14)	0.73	0.74
$FBG0^n/G0^n$	19.15 (3.88)	19.30 (3.65)	0.94 (0.84, 1.05)	0.25	0.74
$FBG1^n/G1^n$	18.20 (3.46)	17.82 (3.25)	1.04 (0.93, 1.16)	0.5	0.74
FB^n/F^n	0.22 (0.05)	0.21 (0.05)	1.02 (0.92, 1.14)	0.68	0.74
$FB^n/F^{n\ total}$	17.88 (3.38)	17.55 (3.07)	1.02 (0.92 , 1.14)	0.68	0.74
$F^n/(B^n + FB^n)$	4.65 (1.09)	4.71 (1.01)	0.98 (0.88, 1.09)	0.74	0.74

Table 4. All cause analysis

Code	Glycan	Dead (N=489)	Survived (N=740)	Crude model (n=1229)		Model I (AJCC, age, sex, n=1229)		Model II (AJCC, age, sex, time between sample and surgery, operation type, bmi, CRP n=952)	
		Mean (SD)	Mean (SD)	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured									
IGP3	Continuous	26.39 (7.33)	23.87 (6.37)	1.04 (1.03, 1.06)	$4.7x10^{-12}$	1.04 (1.03, 1.06)	$7.5x10^{-11}$	1.04 (1.02, 1.05)	$7.5x10^{-6}$
	RT			1.36 (1.24, 1.48)	$5.9x10^{-11}$	1.35 (1.23, 1.48)	$5.0x10^{-10}$	1.28 (1.14, 1.43)	$2.6x10^{-5}$
IGP5	Continuous	6.34 (1.88)	5.79 (1.57)	1.15 (1.09, 1.21)	$1.5x10^{-8}$	1.14 (1.08, 1.20)	$3.5x10^{-7}$	1.11 (1.04, 1.18)	0.001
	RT			1.28 (1.17, 1.40)	$9.7x10^{-8}$	1.27 (1.15, 1.39)	$1.4x10^{-6}$	1.19 (1.06, 1.34)	0.002
IGP7	Continuous	18.26 (2.21)	18.67 (1.81)	0.91 (0.87, 0.96)	0.0001	0.93 (0.88, 0.97)	0.001	0.92 (0.87, 0.97)	0.003
	RT			0.85 (0.77, 0.93)	0.001	0.87 (0.79, 0.95)	0.003	0.87 (0.78, 0.97)	0.01
IGP8	Continuous	9.38 (1.36)	9.83 (1.35)	0.83 (0.78, 0.89)	$3.4x10^{-8}$	0.86 (0.80, 0.91)	$2.6x10^{-6}$	0.87 (0.81, 0.94)	0.0002
	RT			0.78 (0.74, 0.85)	$3.4x10^{-8}$	0.81 (0.74, 0.89)	$4.4x10^{-6}$	0.83 (0.75, 0.92)	0.0003
IGP9	Continuous	5.45 (1.23)	5.48 (1.15)	0.96 (0.89, 1.04)	0.35	0.96 (0.88, 1.03)	0.26	0.95 (0.87, 1.04)	0.29
	RT			0.95 (0.87, 1.04)	0.25	0.94 (0.85, 1.03)	0.16	0.93 (0.84, 1.04)	0.20
IGP13	Continuous	10.19 (3.13)	11.44 (3.13)	0.89 (0.87, 0.92)	$8.9x10^{-13}$	0.90 (0.87, 0.93)	$1.4x10^{-10}$	0.92 (0.89, 0.96)	$4.0x10^{-5}$
	RT			0.70 (0.64, 0.77)	$5.2x10^{-14}$	0.71 (0.65, 0.79)	$1.6x10^{-11}$	0.77 (0.68, 0.86)	$9.6x10^{-6}$
IGP14	Continuous	1.38 (0.41)	1.50 (0.43)	0.55 (0.44, 0.70)	$4.0x10^{-7}$	0.59 (0.46, 0.74)	$5.9x10^{-6}$	0.70 (0.53, 0.92)	0.01
	RT			0.78 (0.71, 0.85)	$9.0x10^{-8}$	0.80 (0.73, 0.88)	$2.0x10^{-6}$	0.86 (0.77, 0.96)	0.007
IGP17	Continuous	7.78 (2.29)	8.47 (2.34)	0.89 (0.85, 0.93)	$9.3x10^{-8}$	0.89 (0.85, 0.93)	$1.3x10^{-7}$	0.91 (0.86, 0.96)	0.0002
	RT			0.76 (0.70, 0.83)	$4.4x10^{-9}$	0.76 (0.69, 0.83)	$8.4x10^{-9}$	0.79 (0.71, 0.89)	$4.9x10^{-5}$
IGP18	Continuous	1.87 (0.38)	1.90 (0.39)	0.85 (0.67, 1.08)	0.18	0.81 (0.65, 1.02)	0.08	0.94 (0.73, 1.22)	0.64
	RT			0.94 (0.86, 1.03)	0.18	0.93 (0.85, 1.01)	0.10	0.98 (0.89, 1.08)	0.70
Sialylation									
IGP24	Continuous	24.78 (3.20)	24.95 (3.10)	0.99 (0.96, 1.01)	0.33	0.97 (0.95, 1.00)	0.06	0.98 (0.95, 1.01)	0.18
	RT			0.95 (0.87, 1.04)	0.29	0.92 (0.84, 1.00)	0.06	0.93 (0.84, 1.03)	0.18
IGP25	Continuous	32.83 (6.17)	32.58 (6.34)	1.01 (0.99, 1.02)	0.38	1.00 (0.99, 1.02)	0.67	1.01 (0.99, 1.02)	0.48
	RT			1.04 (0.95, 1.14)	0.37	1.02 (0.93, 1.12)	0.65	1.04 (0.94, 1.15)	0.47

IGP26	Continuous	16.35 (3.63)	17.35 (3.61)	0.94 (0.91, 0.96)	$7.3x10^{-7}$	0.93 (0.91, 0.96)	$2.7x10^{-7}$	0.94 (0.92, 0.97)	0.0003
	RT			0.79 (0.72, 0.86)	$1.6x10^{-7}$	0.77 (0.71, 0.85)	$6.2x10^{-8}$	0.81 (0.72, 0.90)	0.0001
IGP27	Continuous	21.21 (4.88)	21.82 (4.96)	0.98 (0.96, 1.00)	0.04	0.98 (0.96, 1.00)	0.03	0.99 (0.97, 1.01)	0.30
	RT			0.91 (0.83, 0.99)	0.04	0.91 (0.83, 0.99)	0.03	0.95 (0.86, 1.05)	0.30
IGP29	Continuous	40.18 (2.99)	39.51 (2.77)	1.07 (1.04, 1.10)	$1.2x10^{-5}$	1.03 (1.00, 1.06)	0.03	1.01 (0.98, 1.05)	0.47
	RT			1.23 (1.12, 1.35)	$9.2x10^{-6}$	1.12 (1.03, 1.23)	0.10	1.06 (0.95, 1.18)	0.28
IGP31	Continuous	37.02 (3.87)	36.41 (3.97)	1.04 (1.01, 1.06)	0.002	1.03 (1.01, 1.05)	0.006	1.03 (1.00, 1.05)	0.04
	RT			1.15 (1.05, 1.26)	0.002	1.14 (1.04, 1.24)	0.004	1.12 (1.01, 1.24)	0.03
Bisecting GlcNAc									
IGP36	Continuous	0.30 (0.08)	0.28 (0.07)	8.74 (2.87, 26.66)	0.0001	7.95 (2.43, 26.01)	0.001	9.40 (2.30, 38.49)	0.002
	RT			1.19 (1.09, 1.30)	0.0001	1.18 (1.07, 1.29)	0.001	1.19 (1.06, 1.33)	0.002
IGP37	Continuous	0.17 (0.05)	0.16 (0.05)	19.00 (3.31, 109.06)	0.001	19.22 (3.12, 118.60)	0.001	30.56 (3.50, 267.02)	0.002
	RT			1.16 (1.06, 1.27)	0.001	1.16 (1.06, 1.27)	0.002	1.18 (1.06, 1.32)	0.003
IGP38	Continuous	0.14 (0.04)	0.14 (0.03)	76.90 (6.38, 927.03)	0.001	73.95 (5.58, 980.16)	0.001	126.7 (5.94, 2701)	0.002
	RT			1.17 (1.07, 1.28)	0.001	1.16 (1.06, 1.28)	0.001	1.19 (1.06, 1.32)	0.002
IGP39	Continuous	1.35 (0.32)	1.27 (0.30)	1.82 (1.39, 2.39)	$1.5x10^{-5}$	1.66 (1.26, 2.19)	0.0003	1.62 (1.16, 2.27)	0.005
	RT			1.22 (1.11, 1.33)	$1.4x10^{-5}$	1.19 (1.09, 1.30)	0.0002	1.17 (1.05, 1.31)	0.005
IGP40	Continuous	0.56 (0.06)	0.55 (0.06)	30.19 (6.24, 146.08)	$2.3x10^{-5}$	19.39 (3.90, 96.34)	0.0003	13.96 (2.03, 96.03)	0.007
	RT			1.22 (1.11, 1.33)	$1.9x10^{-5}$	1.19 (1.08, 1.30)	0.0002	1.17 (1.05, 1.30)	0.005
Neutral IgG glycans									
IGP43	Continuous	32.33 (8.02)	29.56 (7.02)	1.04 (1.03, 1.05)	$3.7x10^{-12}$	1.04 (1.03, 1.05)	$1.2x10^{-10}$	1.03 (1.02, 1.05)	$1.0x10^{-5}$
	RT			1.36 (1.24, 1.49)	$4.3x10^{-11}$	1.35 (1.22, 1.48)	$6.9x10^{-10}$	1.28 (1.14, 1.43)	$2.6x10^{-5}$
IGP45	Continuous	7.78 (2.07)	7.18 (1.76)	1.13 (1.08, 1.18)	$4.9x10^{-8}$	1.12 (1.07, 1.17)	$2.0x10^{-6}$	1.09 (1.03, 1.15)	0.002
	RT			1.27 (1.16, 1.39)	$2.9x10^{-7}$	1.24 (1.13, 1.37)	$9.2x10^{-6}$	1.17 (1.05, 1.31)	0.006
IGP47	Continuous	22.59 (3.09)	23.32 (2.52)	0.92 (0.89, 0.95)	$1.1x10^{-6}$	0.93 (0.90, 0.96)	$5.0x10^{-6}$	0.93 (0.90, 0.97)	0.0002
	RT			0.81 (0.74, 0.88)	$4.7x10^{-6}$	0.81 (0.74, 0.89)	$1.4x10^{-5}$	0.83 (0.74, 0.92)	0.001
IGP48	Continuous	11.60 (1.81)	12.28 (1.75)	0.85 (0.81, 0.89)	$1.6x10^{-10}$	0.86 (0.82, 0.91)	$8.8x10^{-9}$	0.88 (0.83, 0.93)	$1.3x10^{-5}$
	RT			0.74 (0.68, 0.81)	$1.5x10^{-10}$	0.76 (0.70, 0.84)	$1.0x10^{-8}$	0.79 (0.71, 0.88)	$1.5x10^{-5}$
IGP49	Continuous	6.73 (1.54)	6.85 (1.46)	0.95 (0.89, 1.01)	0.08	0.94 (0.88, 1.00)	0.04	0.94 (0.87, 1.01)	0.09
	RT			0.91 (0.83, 1.00)	0.05	0.90 (0.82, 0.98)	0.02	0.90 (0.81, 1.01)	0.06
IGP53	Continuous	12.70 (4.29)	14.39 (4.37)	0.92 (0.90, 0.94)	$3.9x10^{-12}$	0.93 (0.91, 0.95)	$1.9x10^{-10}$	0.94 (0.92, 0.97)	$4.7x10^{-5}$
	RT			0.70 (0.64, 0.77)	$8.5x10^{-14}$	0.71 (0.65, 0.79)	$1.0x10^{-11}$	0.76 (0.68, 0.86)	$7.4x10^{-6}$

IGP54	Continuous	1.72 (0.54)	1.89 (0.59)	0.64 (0.54, 0.75)	$1.7x10^{-7}$	0.66 (0.55, 0.78)	$1.5x10^{-6}$	0.75 (0.61, 0.92)	0.005
	RT			0.77 (0.71, 0.85)	$3.4x10^{-8}$	0.79 (0.72, 0.87)	$4.5x10^{-7}$	0.85 (0.76, 0.95)	0.003
Galactosylation									
IGP55	Continuous	41.15 (9.19)	37.69 (8.09)	1.04 (1.03, 1.05)	$1.2x10^{-13}$	1.04 (1.03, 1.05)	$5.9x10^{-12}$	1.03 (1.02, 1.04)	$2.0x10^{-6}$
	RT			1.39 (1.27, 1.53)	$1.0x10^{-12}$	1.38 (1.26, 1.52)	$2.7x10^{-11}$	1.31 (1.16, 1.47)	$5.5x10^{-6}$
IGP56	Continuous	42.72 (4.71)	44.26 (3.56)	0.93 (0.91, 0.95)	$1.7x10^{-12}$	0.93 (0.92, 0.95)	$2.2x10^{-11}$	0.94 (0.92, 0.96)	$2.1x10^{-7}$
	RT			0.75 (0.69, 0.83)	$2.2x10^{-9}$	0.76 (0.69, 0.83)	$8.3x10^{-9}$	0.78 (0.69, 0.87)	$1.3x10^{-5}$
IGP57	Continuous	15.65 (5.00)	17.61 (5.15)	0.94 (0.92, 0.95)	$5.9x10^{-12}$	0.94 (0.92, 0.96)	$4.2x10^{-10}$	0.95 (0.93, 0.98)	0.0001
	RT			0.71 (0.65, 0.78)	$2.5x10^{-13}$	0.72 (0.66, 0.80)	$3.6x10^{-11}$	0.78 (0.69, 0.87)	$2.4x10^{-5}$
Core fucosylation and bisecting GlcNAc									
IGP62	Continuous	79.47 (3.74)	79.77 (3.46)	0.99 (0.96, 1.01)	0.27	0.99 (0.96, 1.02)	0.42	0.99 (0.96, 1.02)	0.41
	RT			0.96 (0.87, 1.05)	0.32	0.97 (0.88, 1.06)	0.52	0.96 (0.86, 1.07)	0.48
IGP63	Continuous	78.81 (4.29)	78.69 (4.11)	1.01 (0.99, 1.03)	0.34	1.01 (0.99, 1.03)	0.29	1.01 (0.98, 1.04)	0.51
	RT			1.05 (0.96, 1.15)	0.30	1.06 (0.96, 1.16)	0.24	1.04 (0.93, 1.16)	0.46
IGP64	Continuous	80.03 (3.79)	80.45 (3.56)	0.98 (0.96, 1.00)	0.10	0.98 (0.96, 1.01)	0.18	0.98 (0.95, 1.01)	0.25
	RT			0.93 (0.85, 1.02)	0.12	0.95 (0.86, 1.04)	0.24	0.94 (0.85, 1.05)	0.31
IGP66	Continuous	17.29 (3.20)	16.97 (2.90)	1.02 (0.99, 1.05)	0.14	1.01 (0.98, 1.05)	0.38	1.01 (0.97, 1.05)	0.64
	RT			1.06 (0.97, 1.16)	0.19	1.03 (0.94, 1.13)	0.50	1.02 (0.91, 1.14)	0.73
IGP67	Continuous	19.15 (3.88)	19.30 (3.65)	0.99 (0.96, 1.01)	0.26	0.98 (0.96, 1.01)	0.17	0.98 (0.96, 1.01)	0.28
	RT			0.95 (0.86, 1.03)	0.22	0.93 (0.85, 1.02)	0.13	0.94 (0.84, 1.05)	0.25
IGP68	Continuous	18.20 (3.46)	17.82 (3.25)	1.02 (0.99, 1.05)	0.12	1.02 (0.99, 1.04)	0.27	1.01 (0.98, 1.05)	0.44
	RT			1.07 (0.98, 1.17)	0.13	1.05 (0.95, 1.15)	0.34	1.04 (0.93, 1.16)	0.50
IGP70	Continuous	0.22 (0.05)	0.21 (0.05)	4.33 (0.71, 26.41)	0.11	2.73 (0.41, 18.15)	0.30	2.05 (0.23, 18.63)	0.52
	RT			1.06 (0.97, 1.16)	0.19	1.03 (0.94, 1.14)	0.49	1.02 (0.92, 1.14)	0.68
IGP71	Continuous	17.88 (3.38)	17.55 (3.07)	1.02 (0.99, 1.05)	0.16	1.01 (0.98, 1.04)	0.38	1.01 (0.98, 1.04)	0.59
	RT			1.06 (0.97, 1.16)	0.21	1.03 (0.94, 1.14)	0.49	1.02 (0.92 , 1.14)	0.68
IGP72	Continuous	4.65 (1.09)	4.71 (1.01)	0.97 (0.89, 1.06)	0.51	1.00 (0.91, 1.09)	0.93	0.99 (0.89, 1.10)	0.85
	RT			0.96 (0.87, 1.05)	0.32	0.98 (0.89, 1.07)	0.64	0.98 (0.88, 1.09)	0.74

Table 5. CRC-specific analysis for rank transformed glycans. Q value represents the adjusted p-values using the false discovery rate method (Benjamini– Hochberg procedure).

Glycan	Dead (N=385)	Survived (N=844)	Model II (AJCC, age, sex, time between sample and surgery, operation type, CRP, bmi, n=971)		
	Mean (SD)	Mean (SD)	HR (95% CI)	p-value	q-value
Total IgG glycans (neutral and charged); Measured					
GP4	26.20 (7.31)	24.27 (6.59)	1.23 (1.09, 1.40)	0.001	0.04
GP6	6.23 (1.85)	5.91 (1.63)	1.10 (0.97, 1.25)	0.14	0.99
GP8	18.33 (2.13)	18.59 (1.91)	0.92 (0.81, 1.04)	0.16	0.99
GP9	9.46 (1.35)	9.74 (1.38)	0.90 (0.80, 1.01)	0.06	0.99
GP10	5.40 (1.19)	5.50 (1.18)	0.90 (0.80, 1.02)	0.11	0.99
GP11	10.30 (3.19)	11.24 (3.15)	0.81 (0.71, 0.92)	0.002	0.06
GP15	1.38 (0.40)	1.49 (0.43)	0.86 (0.76, 0.98)	0.02	0.46
GP18	7.87 (2.31)	8.35 (2.34)	0.82 (0.72, 0.93)	0.002	0.06
GP19	1.87 (0.37)	1.90 (0.39)	1.00 (0.90, 1.12)	0.98	0.99
Sialylation; Derived					
FGS/(FG+FGS)	24.84 (3.24)	24.90 (3.09)	0.90 (0.80, 1.01)	0.08	0.99
FBGS/(FBG+FBGS	32.93 (6.13)	32.56 (6.34)	1.05 (0.94, 1.19)	0.38	0.99
FGS/(F+FG+FGS)	16.47 (3.64)	17.17 (3.63)	0.82 (0.72, 0.92)	0.001	0.04
FBGS/(FB+FBG+FBGS)	21.38 (4.93)	21.66 (4.94)	0.99 (0.88, 1.11)	0.8	0.99
FG2S1/(FG2+FG2S1+FG2S2)	40.22 (3.05)	39.57 (2.77)	1.00 (0.89., 1.13)	0.99	0.99
FBG2S1/(FBG2+FBG2S1+FBG2S2)	37.11 (3.80)	36.44 (3.99)	1.17 (1.04, 1.31)	0.009	0.25
Bisecting GlcNAc; Derived					
FBS^{total}/FS^{total}	0.30 (0.07)	0.29 (0.08)	1.17 (1.03, 1.33)	0.02	0.46
FBS1/FS1	0.17 (0.05)	0.16 (0.05)	1.19 (1.05, 1.35)	0.006	0.17
FBS1/(FS1+FBS1)	0.14 (0.03)	0.14 (0.03)	1.19 (1.06, 1.35)	0.005	0.15
FBS2/FS2	1.33 (0.32)	1.29 (0.30)	1.10 (0.98, 1.25)	0.12	0.99

FBS2/(FS2+FBS2)	0.56 (0.06)	0.55 (0.06)	1.10 (0.97, 1.25)	0.13	0.99
Neutral IgG glycans; Measured					
$GP4^n$	32.14 (8.00)	29.98 (7.24)	1.24 (1.09, 1.41)	0.001	0.04
$GP6^n$	7.65 (2.07)	7.31 (1.83)	1.08 (0.94, 1.22)	0.27	0.99
$GP8^n$	22.71 (3.01)	23.18 (2.66)	0.87 (0.77, 0.98)	0.02	0.46
$GP9^n$	11.72 (1.81)	12.14 (1.79)	0.85 (0.76, 0.97)	0.01	0.26
$GP10^n$	6.68 (1.49)	6.85 (1.50)	0.88 (0.77, 0.99)	0.03	0.66
$GP14^n$	12.86 (4.38)	14.11 (4.38)	0.80 (0.71, 0.92)	0.001	0.04
$GP15^n$	1.72 (0.53)	1.87 (0.59)	0.85 (0.75, 0.97)	0.01	0.26
Galactosylation; Derived					
$G0^n$	40.83 (9.18)	38.26 (8.36)	1.24 (1.09, 1.41)	0.001	0.04
$G1^n$	42.90 (4.63)	43.99 (3.83)	0.82 (0.72, 0.94)	0.003	0.09
$G2^n$	15.79 (5.09)	17.30 (5.16)	0.81 (0.71, 0.93)	0.002	0.06
Core fucosylation and bisecting GlcNAc; Derived					
F^n	79.68 (3.70)	79.63 (3.52)	1.02 (0.90, 1.15)	0.77	0.99
$FG0^n/G0^n$	79.01 (4.23)	78.62 (4.16)	1.10 (0.97, 1.24)	0.13	0.99
$FG1^n/G1^n$	80.24 (3.73)	80.30 (3.62)	1.00 (0.89, 1.14)	0.94	0.99
FB^n	17.10 (3.15)	17.09 (2.97)	0.95 (0.84, 1.08)	0.44	0.99
$FBG0^n/G0^n$	18.98 (3.81)	19.36 (3.70)	0.89 (0.79, 1.01)	0.07	0.99
$FBG1^n/G1^n$	18.01 (3.39)	17.96 (3.32)	0.97 (0.86, 1.10)	0.64	0.99
FB^n/F^n	0.21 (0.05)	0.21 (0.05)	0.95 (0.84, 1.08)	0.47	0.99
$FB^n/F^{n\ total}$	17.68 (3.33)	17.68 (3.14)	0.96 (0.85, 1.09)	0.5	0.99
$F^n/(B^n + FB^n)$	4.71 (1.06)	4.68 (1.03)	1.05 (0.92, 1.18)	0.47	0.99

Table 6. CRC-specific analysis

Code	Glycan	Dead (N=385)	Survived (N=844)	Crude model (n=1229)		Model I (AJCC, age, sex, n=1229)		Model II (AJCC, age, sex, time between sample and surgery, operation type, CRP, bmi, n=971)	
		Mean (SD)	Mean (SD)	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured									
IGP3	Continuous	26.20 (7.31)	24.27 (6.59)	1.04 (1.02, 1.05)	$5.9x10^{-8}$	1.04 (1.03, 1.06)	$2.2x10^{-8}$	1.03 (1.01, 1.05)	0.001
	RT			1.31 (1.18, 1.45)	$2.8x10^{-7}$	1.33 (1.20, 1.48)	$5.5x10^{-8}$	1.23 (1.09, 1.40)	0.001
IGP5	Continuous	6.23 (1.85)	5.91 (1.63)	1.10 (1.04, 1.17)	0.0004	1.12 (1.05, 1.18)	0.0002	1.06 (0.99, 1.14)	0.10
	RT			1.18 (1.07, 1.31)	0.001	1.22 (1.09, 1.35)	0.0003	1.10 (0.97, 1.25)	0.14
IGP7	Continuous	18.33 (2.13)	18.59 (1.91)	0.94 (0.89, 0.99)	0.01	0.94 (0.89, 0.99)	0.01	0.95 (0.89, 1.01)	0.10
	RT			0.89 (0.80, 0.99)	0.03	0.89 (0.80, 0.98)	0.02	0.92 (0.81, 1.04)	0.16
IGP8	Continuous	9.46 (1.35)	9.74 (1.38)	0.87 (0.81, 0.94)	0.003	0.90 (0.84, 0.97)	0.006	0.92 (0.84, 1.00)	0.05
	RT			0.83 (0.75, 0.92)	0.0003	0.88 (0.79, 0.97)	0.009	0.90 (0.80, 1.01)	0.06
IGP9	Continuous	5.40 (1.19)	5.50 (1.18)	0.93 (0.85, 1.01)	0.10	0.93 (0.85, 1.02)	0.11	0.92 (0.83, 1.03)	0.15
	RT			0.91 (0.82, 1.01)	0.08	0.91 (0.82, 1.01)	0.07	0.90 (0.80, 1.02)	0.11
IGP13	Continuous	10.30 (3.19)	11.24 (3.15)	0.91 (0.88, 0.94)	$3.5x10^{-8}$	0.91 (0.88, 0.94)	$5.3x10^{-8}$	0.94 (0.90, 0.98)	0.004
	RT			0.73 (0.66, 0.81)	$5.3x10^{-9}$	0.73 (0.66, 0.81)	$1.5x10^{-8}$	0.81 (0.71, 0.92)	0.002
IGP14	Continuous	1.38 (0.40)	1.49 (0.43)	0.54 (0.42, 0.70)	$2.9x10^{-6}$	0.56 (0.43, 0.72)	$1.2x10^{-5}$	0.70 (0.51, 0.96)	0.03
	RT			0.77 (0.70, 0.86)	$9.7x10^{-7}$	0.79 (0.71, 0.87)	$5.7x10^{-6}$	0.86 (0.76, 0.98)	0.02
IGP17	Continuous	7.87 (2.31)	8.35 (2.34)	0.91 (0.87, 0.96)	0.0001	0.90 (0.85, 0.94)	$5.2x10^{-6}$	0.92 (0.87, 0.97)	0.003
	RT			0.80 (0.72, 0.89)	$2.4x10^{-5}$	0.77 (0.70, 0.86)	$1.4x10^{-6}$	0.82 (0.72, 0.93)	0.002
IGP18	Continuous	1.87 (0.37)	1.90 (0.39)	0.85 (0.65, 1.10)	0.22	0.78 (0.60, 1.02)	0.07	0.99 (0.74, 1.33)	0.97
	RT			0.94 (0.85, 1.04)	0.24	0.92 (0.83, 1.01)	0.08	1.00 (0.90, 1.12)	0.98
Sialylation; Derived									
IGP24	Continuous	24.84 (3.24)	24.90 (3.09)	0.99 (0.96, 1.02)	0.63	0.97 (0.94, 1.00)	0.06	0.97 (0.93, 1.00)	0.09
	RT			0.97 (0.88, 1.07)	0.56	0.91 (0.83, 1.01)	0.07	0.90 (0.80, 1.01)	0.08
IGP25	Continuous	32.93 (6.13)	32.56 (6.34)	1.01 (0.99, 1.03)	0.27	1.00 (0.99, 1.02)	0.67	1.01 (0.99, 1.03)	0.37

	RT			1.06 (0.96, 1.17)	0.26	1.02 (0.92, 1.13)	0.65	1.05 (0.94, 1.19)	0.38
IGP26	Continuous	16.47 (3.64)	17.17 (3.63)	0.95 (0.92, 0.98)	0.0003	0.94 (0.91, 0.96)	$7.4x10^{-6}$	0.95 (0.92, 0.98)	0.002
	RT			0.82 (0.74, 0.91)	0.0001	0.78 (0.71, 0.87)	$3.4x10^{-6}$	0.82 (0.72, 0.92)	0.001
IGP27	Continuous	21.38 (4.93)	21.66 (4.94)	0.99 (0.97, 1.01)	0.29	0.98 (0.96, 1.00)	0.08	1.00 (0.97, 1.02)	0.79
	RT			0.95 (0.86, 1.05)	0.30	0.92 (0.83, 1.01)	0.09	0.99 (0.88, 1.11)	0.80
IGP29	Continuous	40.22 (3.05)	39.57 (2.77)	1.07 (1.04, 1.11)	$4.0x10^{-5}$	1.02 (0.99, 1.06)	0.13	0.99 (0.96, 1.03)	0.75
	RT			1.24 (1.12, 1.38)	$3.7x10^{-5}$	1.10 (1.00, 1.22)	0.05	1.00 (0.89., 1.13)	0.99
IGP31	Continuous	37.11 (3.80)	36.44 (3.99)	1.04 (1.01, 1.07)	0.002	1.04 (1.01, 1.06)	0.006	1.04 (1.01, 1.07)	0.01
	RT			1.18 (1.06, 1.30)	0.001	1.15 (1.05, 1.27)	0.004	1.17 (1.04, 1.31)	0.009
Bisecting GlcNAc; Derived									
IGP36	Continuous	0.30 (0.07)	0.29 (0.08)	4.42 (1.23, 15.89)	0.02	5.06 (1.30, 19.60)	0.02	7.14 (1.41, 36.18)	0.02
	RT			1.13 (1.03, 1.25)	0.01	1.14 (1.03, 1.27)	0.01	1.17 (1.03, 1.33)	0.02
IGP37	Continuous	0.17 (0.05)	0.16 (0.05)	7.92 (1.06, 58.84)	0.04	11.26 (1.42, 89.56)	0.02	30.51 (2.53, 367.34)	0.007
	RT			1.12 (1.01, 1.23)	0.03	1.13 (1.02, 1.26)	0.02	1.19 (1.05, 1.35)	0.006
IGP38	Continuous	0.14 (0.03)	0.14 (0.03)	24.79 (1.45, 423.15)	0.03	40.48 (2.19, 748.37)	0.01	138.34 (4.24, 4509)	0.006
	RT			1.12 (1.02, 1.24)	0.02	1.14 (1.03, 1.27)	0.01	1.19 (1.06, 1.35)	0.005
IGP39	Continuous	1.33 (0.32)	1.29 (0.30)	1.53 (1.12, 2.08)	0.008	1.42 (1.04, 1.94)	0.03	1.35 (0.92, 1.98)	0.13
	RT			1.15 (1.04, 1.27)	0.006	1.13 (1.02, 1.25)	0.02	1.10 (0.98, 1.25)	0.12
IGP40	Continuous	0.56 (0.06)	0.55 (0.06)	10.50 (1.78, 61.78)	0.009	7.90 (1.32, 47.39)	0.02	4.87 (0.55, 43.15)	0.16
	RT			1.15 (1.04, 1.27)	0.008	1.13 (1.02, 1.25)	0.02	1.10 (0.97, 1.25)	0.13
Neutral IgG glycans; Derived									
IGP43	Continuous	32.14 (8.00)	29.98 (7.24)	1.04 (1.02, 1.05)	$3.5x10^{-8}$	1.02 (1.03, 1.05)	$2.8x10^{-8}$	1.03 (1.01, 1.05)	0.001
	RT			1.32 (1.19, 1.46)	$1.6x10^{-7}$	1.33 (1.20, 1.48)	$6.4x10^{-8}$	1.24 (1.09, 1.41)	0.001
IGP45	Continuous	7.65 (2.07)	7.31 (1.83)	1.09 (1.04, 1.14)	0.0009	1.10 (1.04, 1.15)	0.001	1.04 (0.98, 1.11)	0.19
	RT			1.17 (1.06, 1.30)	0.002	1.19 (1.07, 1.33)	0.001	1.08 (0.94, 1.22)	0.27
IGP47	Continuous	22.71 (3.01)	23.18 (2.66)	0.94 (0.91, 0.98)	0.001	0.93 (0.90, 0.97)	0.0002	0.95 (0.91, 0.99)	0.02
	RT			0.85 (0.77, 0.95)	0.003	0.83 (0.75, 0.92)	0.0003	0.87 (0.77, 0.98)	0.02
IGP48	Continuous	11.72 (1.81)	12.14 (1.79)	0.89 (0.84, 0.94)	$2.2x10^{-5}$	0.90 (0.85, 0.95)	0.0002	0.92 (0.86, 0.98)	0.01
	RT			0.80 (0.72, 0.89)	$2.1x10^{-5}$	0.82 (0.74, 0.91)	0.0002	0.85 (0.76, 0.97)	0.01
IGP49	Continuous	6.68 (1.49)	6.85 (1.50)	0.92 (0.86, 0.99)	0.03	0.91 (0.85, 0.98)	0.01	0.92 (0.84, 1.00)	0.04
	RT			0.89 (0.80, 0.98)	0.02	0.87 (0.78, 0.97)	0.009	0.88 (0.77, 0.99)	0.03
IGP53	Continuous	12.86 (4.38)	14.11 (4.38)	0.93 (0.91, 0.96)	$1.1x10^{-7}$	0.93 (0.91, 0.96)	$5.6x10^{-8}$	0.95 (0.93, 0.98)	0.003

	RT			0.74 (0.67, 0.82)	$9.1x10^{-9}$	0.73 (0.66, 0.81)	$9.7x10^{-9}$	0.80 (0.71, 0.92)	0.001
IGP54	Continuous	1.72 (0.53)	1.87 (0.59)	0.63 (0.52, 0.76)	$2.4x10^{-6}$	0.63 (0.52, 0.77)	$3.9x10^{-6}$	0.76 (0.60, 0.95)	0.02
	RT			0.77 (0.70, 0.86)	$8.5x10^{-11}$	0.78 (0.70, 0.86)	$2.0x10^{-6}$	0.85 (0.75, 0.97)	0.01
Galactosylation; Derived									
IGP55	Continuous	40.83 (9.18)	38.26 (8.36)	1.03 (1.02, 1.04)	$1.7x10^{-8}$	1.03 (1.02, 1.05)	$1.0x10^{-8}$	1.02 (1.01, 1.04)	0.001
	RT			1.33 (1.20, 1.47)	$7.0x10^{-8}$	1.35 (1.22, 1.50)	$2.3x10^{-8}$	1.24 (1.09, 1.41)	0.001
IGP56	Continuous	42.90 (4.63)	43.99 (3.83)	0.94 (0.92, 0.96)	$2.5x10^{-7}$	0.94 (0.92, 0.96)	$1.1x10^{-7}$	0.95 (0.92, 0.98)	0.0004
	RT			0.80 (0.72, 0.88)	$1.5x10^{-5}$	0.78 (0.70, 0.86)	$2.4x10^{-6}$	0.82 (0.72, 0.94)	0.003
IGP57	Continuous	15.79 (5.09)	17.30 (5.16)	0.94 (0.92, 0.96)	$6.5x10^{-8}$	0.94 (0.92, 0.96)	$6.1x10^{-8}$	0.96 (0.94, 0.99)	0.005
	RT			0.74 (0.66, 0.82)	$7.0x10^{-9}$	0.73 (0.66, 0.82)	$1.2x10^{-8}$	0.81 (0.71, 0.93)	0.002
Core fucosylation and bisecting GlcNAc; Derived									
IGP62	Continuous	79.68 (3.70)	79.63 (3.52)	1.00 (0.98, 1.03)	0.79	1.00 (0.97, 1.03)	0.94	1.00 (0.97, 1.04)	0.88
	RT			1.02 (0.92, 1.13)	0.71	1.01 (0.91, 1.13)	0.80	1.02 (0.90, 1.15)	0.77
IGP63	Continuous	79.01 (4.23)	78.62 (4.16)	1.02 (1.00, 1.05)	0.07	1.02 (1.00, 1.05)	0.10	1.02 (0.99, 1.05)	0.17
	RT			1.10 (1.00, 1.22)	0.06	1.10 (0.99, 1.22)	0.07	1.10 (0.97, 1.24)	0.13
IGP64	Continuous	80.24 (3.73)	80.30 (3.62)	1.00 (0.97, 1.02)	0.84	1.00 (0.97, 1.02)	0.75	1.00 (0.96, 1.03)	0.94
	RT			0.99 (0.90, 1.10)	0.92	0.99 (0.89, 1.11)	0.90	1.00 (0.89, 1.14)	0.94
IGP66	Continuous	17.10 (3.15)	17.09 (2.97)	1.00 (0.97, 1.03)	0.99	1.00 (0.96, 1.03)	0.95	0.99 (0.95, 1.03)	0.51
	RT			0.99 (0.90, 1.10)	0.89	0.99 (0.89, 1.10)	0.80	0.95 (0.84, 1.08)	0.44
IGP67	Continuous	18.98 (3.81)	19.36 (3.70)	0.97 (0.95, 1.00)	0.06	0.97 (0.95, 1.00)	0.06	0.97 (0.94, 1.00)	0.08
	RT			0.90 (0.82, 1.00)	0.05	0.90 (0.81, 1.00)	0.04	0.89 (0.79, 1.01)	0.07
IGP68	Continuous	18.01 (3.39)	17.96 (3.32)	1.00 (0.97, 1.03)	0.87	1.00 (0.97, 1.03)	0.93	0.99 (0.96, 1.03)	0.73
	RT			1.00 (0.91, 1.11)	0.94	1.00 (0.90, 1.11)	0.93	0.97 (0.86, 1.10)	0.64
IGP70	Continuous	0.21 (0.05)	0.21 (0.05)	1.08 (0.13, 8.62)	0.94	1.06 (0.12, 9.53)	0.30	0.50 (0.04, 6.76)	0.60
	RT			0.99 (0.90, 1.10)	0.87	0.99 (0.89, 1.10)	0.79	0.95 (0.84, 1.08)	0.47
IGP71	Continuous	17.68 (3.33)	17.68 (3.14)	1.00 (0.97, 1.03)	0.96	1.00 (0.97, 1.03)	0.96	0.99 (0.95, 1.03)	0.58
	RT			0.99 (0.89, 1.10)	0.85	0.99 (0.89, 1.10)	0.80	0.96 (0.85, 1.09)	0.50
IGP72	Continuous	4.71 (1.06)	4.68 (1.03)	1.03 (0.93, 1.13)	0.61	1.03 (0.93, 1.14)	0.54	1.05 (0.93, 1.18)	0.42
	RT			1.02 (0.92, 1.13)	0.68	1.02 (0.92, 1.14)	0.67	1.05 (0.92, 1.18)	0.47

These results show that IgG glycans linked to mainly galactosylation are strongly associated with all-cause mortality and CRC mortality.
In particular an increase in the percentage of agalactosylated structures ($G0^n$) and a decrease in mono- and di-galactosylated structures ($G1^n$, $G2^n$) was associated with poorer all-cause and CRC-specific mortality. Statistically significant associations were also observed for decreased sialylation and increase in the incidence of bisecting GlcNAc **(Tables 3 and 5).**

Results were similar when AJCC stage 4 patients were excluded from the analysis **(Tables 7 and 8)** . The minus logarithm of the q-values (FDR corrected p-values) of all 39 glycan traits for all-cause mortality and CRC-specific model III are presented in a Manhattan-like plot **(Figure 6)**.

Table 7. All-cause mortality analysis for stages 1-3.

Code	Glycan	Dead (N=355)	Survived (N=728)	Crude model (n=1083)		Model I (AJCC, age, sex, n=1083)		Model II (AJCC, age, sex, time between sample and surgery, operation type, bmi, CRP n=850)	
		Mean (SD)	Mean (SD)	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured									
IGP3	Continuous	26.11 (7.30)	23.90 (6.34)	1.04 (1.03, 1.06)	$9.5x10^{-8}$	1.04 (1.02, 1.05)	$1.6x10^{-6}$	1.04 (1.02, 1.06)	$9.6x10^{-5}$
	RT			1.32 (1.18, 1.47)	$4.4x10^{-7}$	1.29 (1.16, 1.45)	$8.13x10^{-6}$	1.28 (1.11, 1.46)	0.0004
IGP5	Continuous	6.43 (1.86)	5.80 (1.86)	1.18 (1.12, 1.25)	$3.3x10^{-9}$	1.16 (1.10, 1.24)	$7.6x10^{-7}$	1.16 (1.08, 1.25)	$3.8x10^{-5}$
	RT			1.35 (1.22, 1.51)	$3.0x10^{-8}$	1.30 (1.16, 1.47)	$7.2x10^{-6}$	1.30 (1.13, 1.49)	0.0002
IGP7	Continuous	18.25 (2.18)	18.66 (1.80)	0.91 (0.86, 0.96)	0.001	0.92 (0.87, 0.98)	0.006	0.91 (0.85, 0.97)	0.002
	RT			0.84 (0.75, 0.93)	0.001	0.87 (0.78, 0.97)	0.01	0.84 (0.74, 0.96)	0.008
IGP8	Continuous	9.35 (1.39)	9.84 (1.34)	0.81 (0.75, 0.87)	$5.9x10^{-8}$	0.81 (0.75, 0.88)	$2.1x10^{-7}$	0.80 (0.73, 0.87)	$7.9x10^{-7}$
	RT			0.75 (0.67, 0.83)	$6.9x10^{-8}$	0.75 (0.68, 0.84)	$2.8x10^{-7}$	0.73 (0.65, 0.83)	$1.0x10^{-6}$
IGP9	Continuous	5.57 (1.27)	5.48 (1.16)	1.05 (0.96, 1.14)	0.32	1.02 (0.93, 1.12)	0.65	1.00 (0.91, 1.11)	0.94
	RT			1.04 (0.94, 1.16)	0.44	1.01 (0.91, 1.13)	0.81	1.00 (0.88, 1.13)	0.95
IGP13	Continuous	10.29 (3.07)	11.42 (3.11)	0.90 (0.87, 0.93)	$8.1x10^{-9}$	0.91 (0.87, 0.94)	$7.9x10^{-7}$	0.92 (0.88, 0.96)	0.0003
	RT			0.71 (0.64, 0.80)	$1.6x10^{-9}$	0.73 (0.65, 0.82)	$1.6x10^{-7}$	0.75 (0.65, 0.87)	$7.5x10^{-5}$
IGP14	Continuous	1.41 (0.41)	1.50 (0.43)	0.64 (0.49, 0.83)	0.001	0.67 (0.51, 0.87)	0.003	0.73 (0.53, 1.00)	0.05
	RT			0.83 (0.74, 0.92)	0.0004	0.84 (0.75, 0.94)	0.002	0.87 (0.76, 0.99)	0.03
IGP17	Continuous	7.83 (2.30)	8.45 (2.28)	0.90 (0.86, 0.94)	$2.6x10^{-5}$	0.91 (0.86, 0.96)	0.0005	0.92 (0.87, 0.99)	0.02
	RT			0.77 (0.70, 0.86)	$3.2x10^{-6}$	0.79 (0.70, 0.89)	$5.7x10^{-5}$	0.81 (0.71, 0.93)	0.003
IGP18	Continuous	1.88 (0.39)	1.90 (0.39)	0.89 (0.67, 1.16)	0.38	0.87 (0.67, 1.13)	0.31	0.89 (0.66, 1.20)	0.43
	RT			0.95 (0.86, 1.05)	0.33	0.95 (0.86, 1.05)	0.33	0.95 (0.85, 1.07)	0.40
Sialylation									
IGP24	Continuous	24.78 (3.18)	24.93 (3.04)	0.99 (0.95, 1.02)	0.49	0.99 (0.96, 1.03)	0.58	1.00 (0.96, 1.04)	0.88
	RT			0.96 (0.86, 1.06)	0.40	0.96 (0.86, 1.07)	0.48	1.00 (0.88, 1.13)	0.99
IGP25	Continuous	32.49 (6.30)	32.58 (6.38)	1.00 (0.98, 1.02)	0.91	1.00 (0.98, 1.02)	0.94	1.00 (0.98, 1.02	0.90
	RT			0.99 (0.90, 1.10)	0.91	1.00 (0.90, 1.10)	0.96	1.01 (0.90, 1.13)	0.89
IGP26	Continuous	16.43 (3.64)	17.32 (3.55)	0.94 (0.91, 0.97)	0.0001	0.95 (0.92, 0.98)	0.001	0.95 (0.92, 0.99)	0.02
	RT			0.80 (0.72, 0.89)	$4.0x10^{-5}$	0.81 (0.73, 0.91)	0.0002	0.83 (0.73, 0.95)	0.007
IGP27	Continuous	21.01 (5.01)	21.81 (4.97)	0.97 (0.95, 1.00)	0.02	0.98 (0.96, 1.00)	0.05	0.98 (0.96, 1.01)	0.14
	RT			0.87 (0.79, 0.97)	0.01	0.90 (0.81, 1.00)	0.05	0.91 (0.81, 1.02)	0.12

IGP29	Continuous	40.09 (2.90)	39.49 (2.70)	1.07 (1.03, 1.11)	0.0001	1.06 (1.02, 1.10)	0.001	1.06 (1.01, 1.10)	0.009
	RT			1.23 (1.10, 1.37)	0.0002	1.20 (1.07, 1.34)	0.001	1.19 (1.05, 1.35)	0.007
IGP31	Continuous	36.78 (3.86)	36.40 (3.97)	1.02 (1.00, 1.05)	0.10	1.02 (0.99, 1.05)	0.12	1.01 (0.98, 1.04)	0.58
	RT			1.10 (0.99, 1.21)	0.08	1.09 (0.98, 1.21)	0.10	1.04 (0.92, 1.17)	0.53
Bisecting GlcNAc									
IGP36	Continuous	0.30 (0.08)	0.28 (0.07)	10.21 (2.80, 37.30)	0.0004	7.18 (1.82, 28.32)	0.005	7.26 (1.42, 37.22)	0.02
	RT			1.20 (1.09, 1.34)	0.0004	1.17 (1.05, 1.30)	0.006	1.16 (1.02, 1.32)	0.02
IGP37	Continuous	0.17 (0.05)	0.16 (0.05)	20.26 (2.63, 155.9)	0.004	13.32 (1.57, 112.9)	0.02	11.89 (0.94, 150.5)	0.06
	RT			1.16 (1.05, 1.29)	0.005	1.13 (1.02, 1.26)	0.02	1.12 (0.99, 1.28)	0.08
IGP38	Continuous	0.14 (0.04)	0.14 (0.03)	84.20 (4.58, 1549)	0.003	43.03 (2.04, 909.9)	0.02	33.95 (0.92, 12.46)	0.06
	RT			1.17 (1.05, 1.30)	0.003	1.14 (1.02, 1.27)	0.02	1.13 (0.99, 1.28)	0.07
IGP39	Continuous	1.36 (0.33)	1.27 (0.30)	2.04 (1.49, 2.81)	1.0×10^{-5}	1.86 (1.34, 2.58)	0.0002	1.84 (1.24, 2.73)	0.002
	RT			1.26 (1.13, 1.40)	2.0×10^{-5}	1.22 (1.10, 1.36)	0.0003	1.21 (1.06, 1.37)	0.005
IGP40	Continuous	0.56 (0.06)	0.55 (0.06)	55.05 (8.57, 353.6)	2.4×10^{-5}	32.58 (4.82, 220.4)	0.0004	24.66 (2.52, 241.26)	0.006
	RT			1.26 (1.13, 1.40)	2.2×10^{-5}	1.22 (1.09, 1.36)	0.0003	1.20 (1.06, 1.37)	0.005
Neutral IgG glycans									
IGP43	Continuous	31.97 (7.93)	29.59 (6.98)	1.04 (1.02, 1.05)	1.3×10^{-7}	1.04 (1.02, 1.05)	2.4×10^{-6}	1.03 (1.02, 1.05)	0.0001
	RT			1.32 (1.18, 1.47)	5.3×10^{-7}	1.29 (1.15, 1.45)	1.0×10^{-5}	1.28 (1.12, 1.47)	0.0004
IGP45	Continuous	7.88 (2.08)	7.18 (1.75)	1.16 (1.10, 1.22)	5.4×10^{-9}	1.14 (1.08, 1.21)	1.6×10^{-6}	1.14 (1.07, 1.22)	5.8×10^{-5}
	RT			1.35 (1.21, 1.50)	4.8×10^{-8}	1.29 (1.15, 1.45)	1.5×10^{-5}	1.29 (1.12, 1.48)	0.0003
IGP47	Continuous	22.58 (3.07)	23.30 (2.50)	0.92 (0.88, 0.95)	1.3×10^{-5}	0.93 (0.89, 0.97)	0.0004	0.92 (0.88, 0.97)	0.0004
	RT			0.79 (0.71, 0.89)	3.8×10^{-5}	0.82 (0.73, 0.92)	0.001	0.81 (0.71, 0.92)	0.002
IGP48	Continuous	11.56 (1.84)	12.29 (1.74)	0.83 (0.78, 0.88)	8.8×10^{-10}	0.84 (0.79, 0.89)	6.5×10^{-9}	0.83 (0.77, 0.89)	1.1×10^{-7}
	RT			0.72 (0.64, 0.80)	8.0×10^{-10}	0.72 (0.65, 0.81)	7.0×10^{-9}	0.71 (0.63, 0.81)	1.2×10^{-7}
IGP49	Continuous	6.88 (1.60)	6.84 (1.47)	1.01 (0.94, 1.08)	0.80	0.99 (0.93, 1.07)	0.88	0.99 (0.91, 1.07)	0.73
	RT			1.00 (0.90, 1.12)	0.95	0.98 (0.88, 1.09)	0.72	0.97 (0.86, 1.10)	0.63
IGP53	Continuous	12.82 (4.23)	14.36 (4.34)	0.93 (0.90, 0.95)	2.0×10^{-8}	0.93 (0.91, 0.96)	1.5×10^{-6}	0.94 (0.91, 0.97)	0.001
	RT			0.72 (0.64, 0.80)	1.8×10^{-9}	0.73 (0.65, 0.82)	1.5×10^{-7}	0.75 (0.65, 0.87)	7.8×10^{-5}
IGP54	Continuous	1.76 (0.55)	1.88 (0.59)	0.70 (0.58, 0.85)	0.0004	0.73 (0.60, 0.89)	0.002	0.78 (0.62, 0.99)	0.04
	RT			0.82 (0.73, 0.91)	0.0002	0.83 (0.75, 0.93)	0.001	0.86 (0.76, 0.98)	0.02
Galactosylation									
IGP55	Continuous	40.89 (9.04)	37.72 (8.04)	1.04 (1.03, 1.05)	1.7×10^{-9}	1.04 (1.02, 1.05)	7.1×10^{-8}	1.04 (1.02, 1.05)	8.4×10^{-6}
	RT			1.38 (1.24, 1.54)	6.7×10^{-9}	1.35 (1.21, 1.52)	2.9×10^{-7}	1.34 (1.17, 1.54)	2.9×10^{-5}
IGP56	Continuous	42.83 (4.68)	42.25 (3.54)	0.93 (0.91, 0.96)	6.5×10^{-9}	0.94 (0.91, 0.96)	6.9×10^{-8}	0.93 (0.90, 0.96)	3.3×10^{-7}

	RT			0.76 (0.68, 0.85)	$9.3x10^{-7}$	0.77 (0.69, 0.87)	$1.0x10^{-5}$	0.75 (0.65, 0.86)	$2.6x10^{-5}$
IGP57	Continuous	15.81 (4.93)	17.58 (5.12)	0.94 (0.92, 0.96)	$4.1x10^{-8}$	0.94 (0.92, 0.97)	$2.5x10^{-6}$	0.95 (0.93, 0.98)	0.001
	RT			0.72 (0.65, 0.81)	$5.4x10^{-9}$	0.74 (0.66, 0.83)	$3.6x10^{-7}$	0.76 (0.66, 0.88)	0.0002
Core fucosylation and bisecting GlcNAc									
IGP62	Continuous	79.17 (3.89)	79.77 (3.47)	0.97 (0.94, 0.99)	0.02	0.98 (0.95, 1.01)	0.11	0.97 (0.94, 1.01)	0.12
	RT			0.89 (0.80, 0.99)	0.03	0.92 (0.83, 1.03)	0.15	0.91 (0.81, 1.03)	0.15
IGP63	Continuous	78.42 (4.42)	78.69 (4.14)	0.99 (0.97, 1.01)	0.41	1.00 (0.97, 1.02)	0.74	0.99 (0.96, 1.02)	0.54
	RT			0.96 (0.86, 1.07)	0.46	0.99 (0.89, 1.10)	0.81	0.96 (0.85, 1.09)	0.57
IGP64	Continuous	79.71 (3.94)	80.45 (3.57)	0.96 (0.93, 0.99)	0.003	0.97 (0.94, 1.00)	0.03	0.97 (0.94, 1.00)	0.04
	RT			0.86 (0.77, 0.96)	0.005	0.89 (0.80, 0.99)	0.04	0.88 (0.78, 1.00)	0.05
IGP66	Continuous	17.58 (3.31)	16.97 (2.91)	1.05 (1.02, 1.09)	0.002	1.04 (1.00, 1.07)	0.04	1.04 (1.00, 1.08)	0.08
	RT			1.16 (1.05, 1.29)	0.005	1.11 (1.00, 1.24)	0.06	1.11 (0.98, 1.26)	0.10
IGP67	Continuous	19.52 (3.98)	19.29 (3.67)	1.01 (0.98, 1.04)	0.44	1.00 (0.98, 1.03)	0.83	1.01 (0.97, 1.04)	0.67
	RT			1.04 (0.93, 1.15)	0.51	1.00 (0.90, 1.12)	0.93	1.02 (0.90, 1.16)	0.71
IGP68	Continuous	18.52 (3.59)	17.82 (3.27)	1.05 (1.02, 1.08)	0.002	1.04 (1.01, 1.07)	0.02	1.04 (1.00, 1.08)	0.04
	RT			1.17 (1.05, 1.30)	0.003	1.13 (1.01, 1.25)	0.03	1.13 (1.00, 1.28)	0.05
IGP70	Continuous	0.22 (0.05)	0.21 (0.05)	27.42 (3.53, 213.1)	0.002	10.93 (1.32, 90.70)	0.03	10.14 (0.88, 116.8)	0.06
	RT			1.16 (1.05, 1.29)	0.005	1.11 (0.99, 1.24)	0.06	1.11 (0.98, 1.26)	0.10
IGP71	Continuous	18.18 (3.50)	17.55 (3.08)	1.05 (1.02, 1.08)	0.003	1.03 (1.00, 1.07)	0.04	1.03 (1.00, 1.07)	0.08
	RT			1.16 (1.04, 1.29)	0.006	1.10 (0.99, 1.23)	0.07	1.11 (0.98, 1.26)	0.11
IGP72	Continuous	4.57 (1.10)	4.72 (1.01)	0.90 (0.81, 1.00)	0.04	0.94 (0.84, 1.05)	0.25	0.92 (0.81, 1.04)	0.19
	RT			0.88 (0.79, 0.97)	0.01	0.92 (0.82, 1.02)	0.12	0.91 (0.80, 1.03)	0.15

Table 8. **CRC-specific mortality analysis for stages 1-3.**

Code	Glycan	Dead (N=257)	Survived (N=826)	Crude model (n=1083)		Model I (AJCC, age, sex, n=1083)		Model II (AJCC, age, sex, time between sample and surgery, operation type, bmi, CRP n=850)	
		Mean (SD)	Mean (SD)	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured									
IGP3	Continuous	25.76 (7.28)	24.27 (6.53)	1.03 (1.01, 1.05)	0.0004	1.04 (1.02, 1.05)	0.0001	1.03 (1.01, 1.06)	0.004
	RT			1.24 (1.10, 1.40)	0.001	1.28 (1.12, 1.46)	0.002	1.24 (1.05, 1.45)	0.01
IGP5	Continuous	6.29 (1.85)	5.91 (1.63)	1.13 (1.05, 1.21)	0.005	1.13 (1.05, 1.22)	0.001	1.10 (1.01, 1.20)	0.04
	RT			1.23 (1.09, 1.40)	0.001	1.23 (1.08, 1.42)	0.002	1.17 (0.99, 1.37)	0.06
IGP7	Continuous	18.33 (2.09)	18.59 (1.90)	0.93 (0.88, 1.00)	0.04	0.94 (0.88, 1.00)	0.05	0.93 (0.86, 1.00)	0.07
	RT			0.89 (0.78, 1.01)	0.07	0.89 (0.78, 1.02)	0.09	0.88 (0.76, 1.03)	0.11
IGP8	Continuous	9.46 (1.38)	9.75 (1.37)	0.86 (0.79, 0.94)	0.001	0.86 (0.79, 0.95)	0.002	0.84 (0.75, 0.93)	0.001
	RT			0.82 (0.72, 0.92)	0.001	0.82 (0.73, 0.93)	0.002	0.79 (0.68, 0.91)	0.001
IGP9	Continuous	5.54 (1.25)	5.50 (1.18)	1.02 (0.92, 1.13)	0.67	0.99 (0.89, 1.10)	0.83	0.97 (0.85, 1.10)	0.59
	RT			1.02 (0.90, 1.15)	0.77	0.98 (0.86, 1.11)	0.73	0.95 (0.82, 1.10)	0.52
IGP13	Continuous	10.48 (3.14)	11.23 (3.13)	0.92 (0.89, 0.96)	0.0001	0.91 (0.87, 0.96)	$8.3x10^{-5}$	0.93 (0.88, 0.99)	0.02
	RT			0.77 (0.68, 0.87)	$6.0x10^{-5}$	0.74 (0.65, 0.85)	$3.1x10^{-5}$	0.79 (0.67, 0.93)	0.006
IGP14	Continuous	1.41 (0.41)	1.49 (0.43)	0.66 (0.48, 0.89)	0.007	0.64 (0.47, 0.87)	0.004	0.73 (0.50, 1.06)	0.10
	RT			0.84 (0.74, 0.95)	0.005	0.83 (0.73, 0.94)	0.003	0.87 (0.75, 1.01)	0.07
IGP17	Continuous	7.99 (2.33)	8.33 (2.29)	0.93 (0.88, 0.99)	0.02	0.92 (0.87, 0.98)	0.007	0.95 (0.88, 1.02)	0.16
	RT			0.84 (0.74, 0.95)	0.007	0.81 (0.71, 0.93)	0.003	0.86 (0.73, 1.01)	0.07
IGP18	Continuous	1.88 (0.38)	1.89 (0.39)	0.89 (0.65, 1.23)	0.49	0.86 (0.63, 1.17)	0.34	0.94 (0.66, 1.33)	0.71
	RT			0.96 (0.85, 1.08)	0.47	0.95 (0.85, 1.07)	0.40	0.97 (0.85, 1.11)	0.69
Sialylation									
IGP24	Continuous	24.87 (3.25)	24.88 (3.03)	1.00 (0.96, 1.04)	0.93	0.99 (0.95, 1.03)	0.66	1.00 (0.96, 1.05)	0.85
	RT			0.98 (0.87, 1.12)	0.81	0.96 (0.85, 1.09)	0.56	1.00 (0.86, 1.16)	0.99
IGP25	Continuous	32.52 (6.35)	32.56 (6.35)	1.00 (0.98, 1.02)	0.98	1.00 (0.98, 1.02)	0.91	1.01 (0.98, 1.03)	0.62
	RT			1.00 (0.89, 1.13)	0.99	1.01 (0.90, 1.13)	0.90	1.04 (0.90, 1.19)	0.62
IGP26	Continuous	16.64 (3.67)	17.15 (3.58)	0.96 (0.93, 0.99)	0.02	0.95 (0.92, 0.99)	0.007	0.96 (0.92, 1.01)	0.11
	RT			0.85 (0.75, 0.97)	0.01	0.83 (0.73, 0.94)	0.004	0.86 (0.74, 1.01)	0.06
IGP27	Continuous	21.19 (5.12)	21.66 (4.95)	0.98 (0.96, 1.01)	0.17	0.98 (0.96, 1.01)	0.20	0.99 (0.96, 1.02)	0.63
	RT			0.91 (0.81, 1.03)	0.14	0.92 (0.82, 1.04)	0.20	0.96 (0.84, 1.11)	0.59

IGP29	Continuous	40.13 (2.97)	39.55 (2.71)	1.08 (1.03, 1.12)	0.001	1.06 (1.02, 1.10)	0.008	1.05 (1.00, 1.10)	0.05
	RT			1.24 (1.09, 1.41)	0.001	1.19 (1.05, 1.36)	0.007	1.17 (1.01, 1.36)	0.04
IGP31	Continuous	36.84 (3.79)	36.43 (3.97)	1.03 (1.00, 1.06)	0.10	1.03 (1.00, 1.06)	0.06	1.02 (0.98, 1.05)	0.32
	RT			1.11 (0.98, 1.26)	0.09	1.13 (1.00, 1.27)	0.05	1.08 (0.94, 1.24)	0.28
Bisecting GlcNAc									
IGP36	Continuous	0.29 (0.07)	0.29 (0.08)	4.00 (0.84, 19.03)	0.08	3.86 (0.75, 19.74)	0.11	4.17 (0.58, 29.87)	0.16
	RT			1.12 (0.99, 1.27)	0.06	1.12 (0.99, 1.27)	0.08	1.12 (0.96, 1.31)	0.14
IGP37	Continuous	0.17 (0.05)	0.16 (0.05)	6.17 (0.53, 72.39)	0.15	6.98 (0.55, 89.35)	0.14	7.93 (0.37, 168.5)	0.18
	RT			1.10 (0.97, 1.24)	0.13	1.11 (0.98, 1.26)	0.11	1.11 (0.95, 1.29)	0.19
IGP38	Continuous	0.14 (0.03)	0.14 (0.03)	17.82 (0.55, 577.2)	0.11	20.69 (0.56, 761.0)	0.10	21.73 (0.29, 1602)	0.16
	RT			1.11 (0.98, 1.26)	0.10	1.12 (0.98, 1.27)	0.09	1.12 (0.96, 1.30)	0.16
IGP39	Continuous	1.34 (0.33)	1.29 (0.30)	1.61 (1.10, 2.35)	0.02	1.52 (1.03, 2.25)	0.04	1.45 (0.90, 2.35)	0.13
	RT			1.16 (1.03, 1.32)	0.02	1.15 (1.01, 1.30)	0.03	1.11 (0.95, 1.30)	0.17
IGP40	Continuous	0.56 (0.06)	0.55 (0.06)	13.15 (1.49, 116.1)	0.02	10.21 (1.10, 95.12)	0.04	5.86 (0.39, 88.93)	0.20
	RT			1.16 (1.03, 1.31)	0.02	1.14 (1.01, 1.30)	0.04	1.11 (0.95, 1.30)	0.18
Neutral IgG glycans									
IGP43	Continuous	31.61 (7.90)	29.98 (7.18)	1.03 (1.01, 1.05)	0.0004	1.03 (1.02, 1.05)	0.0001	1.03 (1.01, 1.05)	0.004
	RT			1.24 (1.10, 1.41)	0.001	1.28 (1.12, 1.46)	0.0002	1.24 (1.06, 1.46)	0.008
IGP45	Continuous	7.73 (2.07)	7.31 (1.83)	1.11 (1.05, 1.18)	0.001	1.11 (1.04, 1.19)	0.001	1.08 (1.00, 1.17)	0.05
	RT			1.23 (1.08, 1.39)	0.001	1.22 (1.07, 1.40)	0.004	1.16 (0.98, 1.36)	0.08
IGP47	Continuous	22.73 (2.97)	23.17 (2.63)	0.94 (0.90, 0.99)	0.01	0.94 (0.90, 0.98)	0.009	0.94 (0.89, 1.00)	0.04
	RT			0.86 (0.75, 0.97)	0.02	0.85 (0.74, 0.97)	0.02	0.86 (0.74, 1.00)	0.06
IGP48	Continuous	11.72 (1.84)	12.15 (1.78)	0.88 (0.82, 0.94)	0.0002	0.88 (0.82, 0.94)	0.0003	0.86 (0.79, 0.94)	0.0004
	RT			0.79 (0.70, 0.90)	0.0002	0.79 (0.70, 0.90)	0.0003	0.76 (0.66, 0.89)	0.0004
IGP49	Continuous	6.86 (1.56)	6.85 (1.50)	1.00 (0.92, 1.08)	0.96	0.97 (0.89, 1.05)	0.47	0.96 (0.87, 1.06)	0.42
	RT			0.99 (0.87, 1.12)	0.88	0.95 (0.84, 1.07)	0.41	0.93 (0.81, 1.08)	0.36
IGP53	Continuous	13.10 (4.34)	14.10 (4.35)	0.94 (0.92, 0.97)	0.0002	0.94 (0.91, 0.97)	0.0001	0.95 (0.92, 0.99)	0.02
	RT			0.77 (0.68, 0.88)	7.4×10^{-5}	0.75 (0.65, 0.86)	3.2×10^{-5}	0.79 (0.67, 0.94)	0.007
IGP54	Continuous	1.76 (0.54)	1.87 (0.59)	0.72 (0.58, 0.91)	0.005	0.70 (0.56, 0.88)	0.003	0.78 (0.59, 1.04)	0.09
	RT			0.83 (0.73, 0.94)	0.004	0.82 (0.73, 0.93)	0.002	0.87 (0.74, 1.01)	0.07
Galactosylation									
IGP55	Continuous	40.35 (9.00)	38.27 (8.29)	1.03 (1.01, 1.04)	8.3×10^{-5}	1.03 (1.02, 1.05)	2.4×10^{-5}	1.03 (1.01, 1.05)	0.002
	RT			1.28 (1.12, 1.45)	0.0002	1.32 (1.15, 1.51)	5.3×10^{-5}	1.27 (1.08, 1.50)	0.005
IGP56	Continuous	43.11 (4.56)	43.99 (3.79)	0.95 (0.92, 0.98)	0.0003	0.94 (0.92, 0.97)	6.8×10^{-5}	0.94 (0.91, 0.97)	0.0005

	RT			0.82 (0.72, 0.93)	0.003	0.80 (0.70, 0.91)	0.001	0.79 (0.67, 0.93)	0.004
IGP57	Continuous	16.08 (5.03)	17.29 (5.12)	0.95 (0.93, 0.98)	0.0002	0.95 (0.92, 0.97)	0.0001	0.96 (0.93, 0.99)	0.02
	RT			0.77 (0.68, 0.88)	$6.2x10^{-5}$	0.75 (0.65, 0.86)	$2.9x10^{-5}$	0.80 (0.67, 0.94)	0.008
Core fucosylation and bisecting GlcNAc									
IGP62	Continuous	79.40 (3.90)	79.63 (3.54)	0.98 (0.95, 1.02)	0.34	0.99 (0.96, 1.03)	0.68	0.99 (0.95, 1.04)	0.77
	RT			0.95 (0.84, 1.08)	0.43	0.98 (0.87, 1.12)	0.80	0.99 (0.85, 1.15)	0.87
IGP63	Continuous	78.59 (4.39)	78.61 (4.18)	1.00 (0.97, 1.03)	0.99	1.01 (0.98, 1.04)	0.57	1.01 (0.97, 1.04)	0.67
	RT			1.01 (0.89, 1.14)	0.93	1.04 (0.92, 1.18)	0.51	1.04 (0.90, 1.21)	0.61
IGP64	Continuous	79.92 (3.92)	80.30 (3.65)	0.98 (0.94, 1.01)	0.14	0.98 (0.95, 1.02)	0.33	0.98 (0.95, 1.01)	0.43
	RT			0.92 (0.81, 1.04)	0.18	0.95 (0.84, 1.08)	0.41	0.95 (0.82, 1.10)	0.50
IGP66	Continuous	17.41 (3.29)	17.09 (2.98)	1.03 (0.99, 1.07)	0.12	1.02 (0.98, 1.06)	0.37	1.01 (0.96, 1.06)	0.68
	RT			1.09 (0.96, 1.23)	0.17	1.05 (0.92, 1.19)	0.49	1.02 (0.88, 1.19)	0.77
IGP67	Continuous	19.40 (3.95)	19.36 (3.72)	1.00 (0.97, 1.03)	0.94	0.99 (0.96, 1.02)	0.57	0.99 (0.95, 1.03)	0.62
	RT			1.00 (0.88, 1.13)	0.99	0.96 (0.85, 1.09)	0.51	0.96 (0.83, 1.11)	0.58
IGP68	Continuous	18.35 (3.54)	17.95 (3.34)	1.03 (1.00, 1.07)	0.08	1.02 (0.98, 1.06)	0.26	1.02 (0.97, 1.06)	0.44
	RT			1.10 (0.98, 1.25)	0.12	1.06 (0.94, 1.20)	0.34	1.05 (0.91, 1.22)	0.52
IGP70	Continuous	0.22 (0.05)	0.21 (0.05)	7.65 (0.66, 88.88)	0.10	3.42 (0.27, 42.66)	0.34	2.01 (0.10, 39.38)	0.65
	RT			1.09 (0.96, 1.23)	0.19	1.04 (0.92, 1.18)	0.53	1.02 (0.88, 1.19)	0.79
IGP71	Continuous	18.00 (3.49)	17.68 (3.15)	1.03 (0.99, 1.07)	0.14	1.02 (0.98, 1.06)	0.42	1.01 (0.96, 1.06)	0.68
	RT			1.08 (0.96, 1.23)	0.21	1.04 (0.92, 1.18)	0.55	1.03 (0.88, 1.19)	0.78
IGP72	Continuous	4.62 (1.06)	4.68 (1.04)	0.95 (0.84, 1.07)	0.40	0.99 (0.87, 1.12)	0.82	1.00 (0.86, 1.15)	0.95
	RT			0.94 (0.83, 1.06)	0.30	0.97 (0.86, 1.11)	0.69	0.99 (0.85, 1.15)	0.88

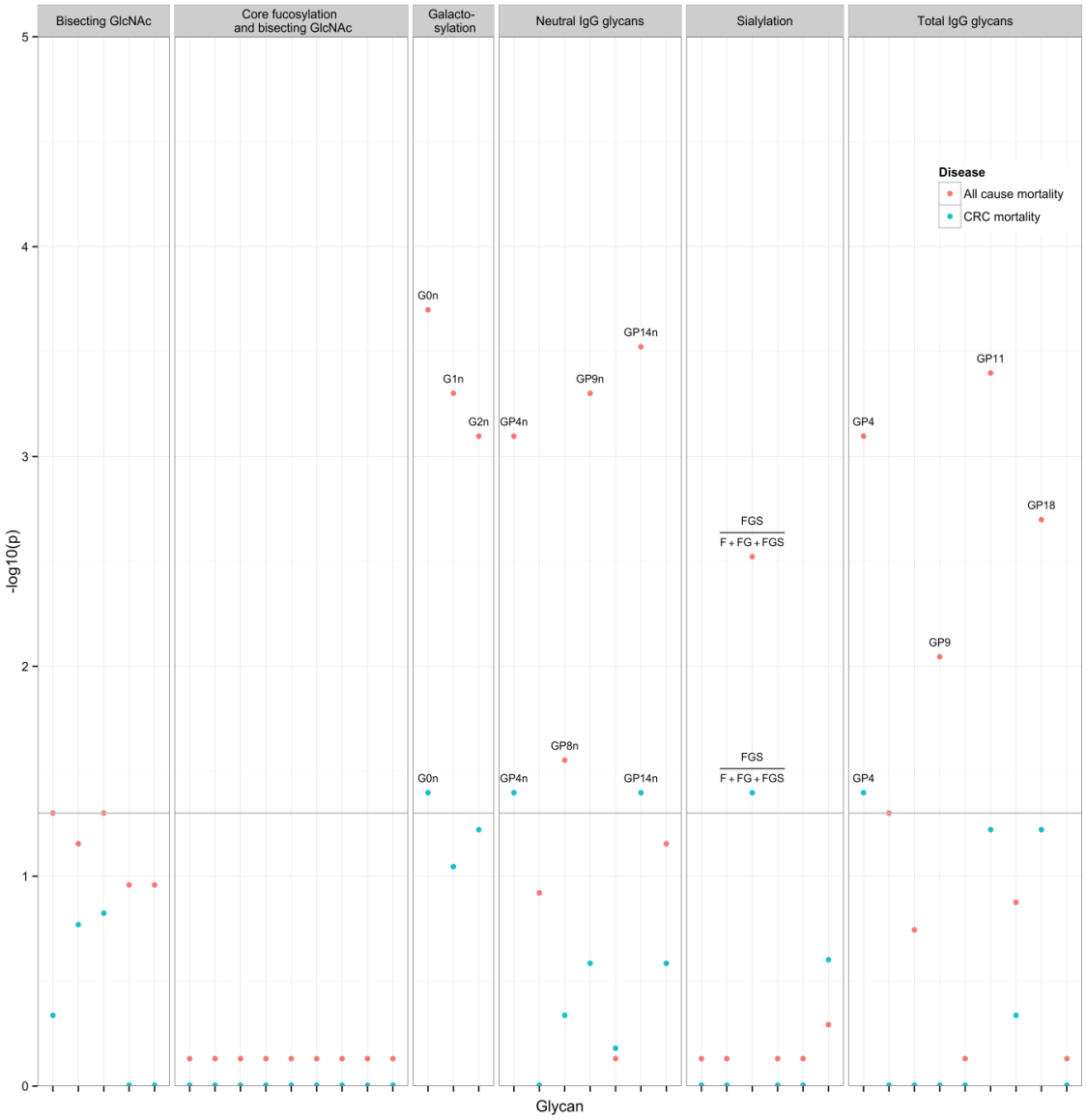


Figure 6. Minus logarithm of the FDR corrected p-values (q values) of all 39 glycan variables for all causes and CRC-specific mortality (Model III). Q- value threshold of significance <0.05.

Stratified analysis by stage for all-cause and CRC-specific mortality is presented in **(Tables 9 and 10).**

Table 9. All-cause mortality analysis by stage [Model II; rank transformed variables]

Code	AJCC stage 1 (n=210)		AJCC stage 2 (n=327)		AJCC stage 3 (n=313)		AJCC stage 4 (n=102)	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured								
IGP3	1.32 (0.96, 1.81)	0.09	1.93 (0.99, 3.78)	0.05	1.27 (1.05, 1.55)	0.02	1.15 (0.93, 1.42)	0.21
IGP5	1.29 (0.92, 1.82)	0.14	1.46 (1.16, 1.84)	0.001	1.20 (0.99, 1.47)	0.07	1.01 (0.83, 1.24)	0.90
IGP7	0.65 (0.47, 0.90)	0.009	0.76 (0.61, 0.96)	0.02	0.95 (0.79, 1.13)	0.55	1.00 (0.81, 1.22)	0.97
IGP8	0.77 (0.56, 1.06)	0.11	0.57 (0.46, 0.72)	$9.3x10^{-7}$	0.84 (0.71, 0.99)	0.04	1.25 (1.00, 1.57)	0.05
IGP9	0.95 (0.70, 1.29)	0.73	1.03 (0.83, 1.28)	0.77	0.98 (0.82, 1.18)	0.85	0.89 (0.70, 1.12)	0.32
IGP13	0.76 (0.55, 1.05)	0.09	0.72 (0.57, 0.93)	0.01	0.77 (0.62, 0.95)	0.02	0.91 (0.74, 1.12)	0.38
IGP14	0.89 (0.64, 1.23)	0.47	0.90 (0.72, 1.13)	0.38	0.85 (0.71, 1.03)	0.09	0.93 (0.73, 1.18)	0.55
IGP17	0.87 (0.63, 1.19)	0.38	0.79 (0.62, 1.01)	0.06	0.81 (0.67, 0.99)	0.04	0.79 (0.65, 0.97)	0.02
IGP18	0.91 (0.67, 1.23)	0.54	0.86 (0.70, 1.07)	0.17	1.02 (0.88, 1.19)	0.77	1.09 (0.85, 1.40)	0.49
Sialylation								
IGP24	1.19 (0.87, 1.64)	0.28	1.08 (0.86, 1.35)	0.51	0.92 (0.77, 1.09)	0.32	0.75 (0.63, 0.91)	0.003
IGP25	1.04 (0.76, 1.42)	0.80	0.97 (0.78, 1.20)	0.76	1.03 (0.88, 1.21)	0.70	1.00 (0.78, 1.28)	0.98
IGP26	0.90 (0.66, 1.22)	0.49	0.86 (0.68, 1.08)	0.20	0.80 (0.67, 0.97)	0.02	0.78 (0.63, 0.95)	0.01
IGP27	0.91 (0.66, 1.25)	0.57	0.84 (0.68, 1.03)	0.10	0.96 (0.82, 1.13)	0.66	0.98 (0.78, 1.23)	0.88
IGP29	1.43 (1.05, 1.96)	0.03	1.19 (0.95, 1.50)	0.13	1.11 (0.93, 1.33)	0.24	0.73 (0.60, 0.89)	0.002
IGP31	0.97 (0.71, 1.33)	0.84	0.87 (0.70, 1.07)	0.19	1.17 (0.99, 1.38)	0.06	1.39 (1.09, 1.76)	0.008
Bisecting GlcNAc								
IGP36	1.12 (0.81, 1.55)	0.48	1.18 (0.93, 1.49)	0.18	1.18 (0.98, 1.41)	0.08	1.24 (0.98, 1.57)	0.07
IGP37	1.04 (0.76, 1.42)	0.81	1.05 (0.85, 1.33)	0.70	1.19 (1.00, 1.43)	0.05	1.36 (1.09, 1.71)	0.008
IGP38	1.06 (0.77, 1.46)	0.72	1.06 (0.83, 1.35)	0.64	1.19 (1.00, 1.43)	0.05	1.36 (1.08, 1.71)	0.008
IGP39	1.17 (0.86, 1.59)	0.33	1.26 (1.00, 1.59)	0.05	1.23 (1.02, 1.47)	0.03	1.05 (0.86, 1.30)	0.63
IGP40	1.16 (0.85, 1.59)	0.34	1.25 (0.99, 1.58)	0.06	1.22 (1.02, 1.47)	0.03	1.05 (0.85, 1.29)	0.65
Neutral IgG glycans								
IGP43	1.35 (0.98, 1.85)	0.07	1.25 (0.98, 1.60)	0.07	1.27 (1.04, 1.54)	0.02	1.12 (0.91, 1.39)	0.29
IGP45	1.29 (0.92, 1.82)	0.14	1.49 (1.18, 1.88)	0.001	1.17 (0.96, 1.43)	0.11	0.96 (0.78, 1.19)	0.74
IGP47	0.67 (0.48, 0.92)	0.02	0.75 (0.60, 0.95)	0.02	0.89 (0.74, 1.07)	0.20	0.90 (0.74, 1.09)	0.29
IGP48	0.78 (0.57, 1.06)	0.12	0.57 (0.46, 0.72)	$8.1x10^{-7}$	0.79 (0.66, 0.95)	0.01	1.16 (0.92, 1.47)	0.20
IGP49	0.94 (0.68, 1.28)	0.68	1.01 (0.82, 1.25)	0.92	0.95 (0.80, 1.14)	0.58	0.83 (0.66, 1.05)	0.12
IGP53	0.77 (0.56, 1.06)	0.11	0.73 (0.57, 0.93)	0.01	0.77 (0.62, 0.94)	0.01	0.89 (0.72, 1.09)	0.25

IGP54	0.88 (0.64, 1.23)	0.46	0.90 (0.71, 1.12)	0.34	0.85 (0.70, 1.01)	0.07	0.90 (0.71, 1.12)	0.34
Galactosylation								
IGP55	1.39 (1.00, 1.92)	0.05	1.38 (1.08, 1.76)	0.009	1.31 (1.07, 1.60)	0.01	1.10 (0.89, 1.35)	0.38
IGP56	0.71 (0.51, 0.98)	0.04	0.69 (0.54, 0.87)	0.002	0.80 (0.66, 0.98)	0.03	0.95 (0.77, 1.17)	0.64
IGP57	0.78 (0.56, 1.07)	0.12	0.78 (0.61, 0.99)	0.05	0.76 (0.66, 0.93)	0.008	0.91 (0.73, 1.12)	0.35
Core fucosylation and bisecting GlcNAc								
IGP62	0.94 (0.68, 1.31)	0.73	0.75 (0.60, 0.94)	0.01	1.00 (0.84, 1.19)	0.99	1.03 (0.80, 1.32)	0.82
IGP63	1.03 (0.76, 1.41)	0.85	0.84 (0.64, 1.01)	0.06	1.03 (0.87, 1.23)	0.74	1.14 (0.91, 1.44)	0.25
IGP64	0.83 (0.60, 1.15)	0.26	0.77 (0.62, 0.96)	0.02	0.97 (0.81, 1.15)	0.69	1.09 (0.85, 1.41)	0.50
IGP66	1.13 (0.82, 1.55)	0.47	1.25 (1.00, 1.55)	0.05	1.03 (0.86, 1.23)	0.73	0.87 (0.70, 1.10)	0.24
IGP67	0.99 (0.73, 1.35)	0.97	1.15 (0.95, 1.49)	0.13	0.96 (0.80, 1.14)	0.64	0.84 (0.68, 1.05)	0.12
IGP68	1.20 (0.87, 1.66)	0.26	1.24 (1.00, 1.53)	0.05	1.05 (0.88, 1.25)	0.58	0.86 (0.68, 1.10)	0.23
IGP70	1.13 (0.81, 1.57)	0.47	1.28 (1.02, 1.59)	0.03	1.02 (0.86, 1.22)	0.79	0.88 (0.69, 1.11)	0.28
IGP71	1.11 (0.81, 1.53)	0.52	1.27 (1.02, 1.58)	0.03	1.02 (0.86, 1.22)	0.79	0.89 (0.71, 1.12)	0.32
IGP72	0.91 (0.66, 1.26)	0.59	0.79 (0.63, 0.98)	0.03	0.99 (0.83, 1.18)	0.92	1.11 (0.88, 1.40)	0.38

Table 10. CRC-specific mortality analysis by stage [Model II; rank transformed variables]

Code	AJCC stage 1 (n=210)		AJCC stage 2 (n=327)		AJCC stage 3 (n=313)		AJCC stage 4 (n=102)	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Total IgG glycans (neutral and charged); Measured								
IGP3	1.14 (0.67, 1.95)	0.63	1.21 (0.90, 1.62)	0.21	1.26 (1.02, 1.55)	0.04	1.15 (0.92, 1.43)	0.21
IGP5	0.94 (0.55, 1.61)	0.82	1.35 (1.01, 1.80)	0.04	1.10 (0.88, 1.36)	0.40	1.01 (0.82, 1.24)	0.94
IGP7	0.81 (0.48, 1.35)	0.42	0.74 (0.56, 0.98)	0.04	0.96 (0.79, 1.17)	0.71	1.00 (0.81, 1.24)	0.97
IGP8	0.72 (0.42, 1.24)	0.24	0.67 (0.51, 0.88)	0.004	0.85 (0.71, 1.02)	0.09	1.25 (0.99, 1.57)	0.06
IGP9	0.86 (0.52, 1.42)	0.56	1.00 (0.77, 1.31)	0.98	0.92 (0.76, 1.12)	0.42	0.88 (0.69, 1.12)	0.30
IGP13	0.98 (0.57, 1.69)	0.95	0.76 (0.56, 1.03)	0.08	0.78 (0.62, 0.98)	0.04	0.91 (0.71, 1.13)	0.39
IGP14	0.85 (0.50, 1.46)	0.56	0.91 (0.69, 1.20)	0.51	0.86 (0.70, 1.05)	0.13	0.92 (0.72, 1.18)	0.53
IGP17	1.12 (0.65, 1.94)	0.68	0.82 (0.61, 1.11)	0.20	0.85 (0.68, 1.05)	0.13	0.79 (0.64, 0.97)	0.02
IGP18	1.03 (0.62, 1.70)	0.92	0.84 (0.64, 1.08)	0.18	1.04 (0.88, 1.23)	0.64	1.07 (0.83, 1.38)	0.61
Sialylation								
IGP24	1.29 (0.77, 2.18)	0.34	1.05 (0.79, 1.39)	0.74	0.96 (0.79, 1.16)	0.64	0.75 (0.62, 0.90)	0.003
IGP25	1.23 (0.72, 2.11)	0.44	0.93 (0.71, 1.22)	0.61	1.08 (0.91, 1.29)	0.39	1.00 (0.78, 1.28)	0.99
IGP26	1.08 (0.64, 1.82)	0.77	0.87 (0.66, 1.16)	0.35	0.84 (0.68, 1.03)	0.09	0.77 (0.62, 0.95)	0.01
IGP27	1.19 (0.69, 2.05)	0.53	0.84 (0.64, 1.08)	0.18	1.02 (0.86, 1.22)	0.79	0.98 (0.77, 1.23)	0.85
IGP29	1.40 (0.85, 2.31)	0.19	1.17 (0.88, 1.55)	0.28	1.14 (0.94, 1.38)	0.20	0.73 (0.60, 0.90)	0.002
IGP31	1.05 (0.63, 1.73)	0.86	0.89 (0.69, 1.16)	0.41	1.18 (0.98, 1.42)	0.07	1.38 (1.08, 1.76)	0.01
Bisecting GlcNAc								
IGP36	1.02 (0.61, 1.72)	0.94	1.10 (0.82, 1.46)	0.53	1.15 (0.95, 1.40)	0.15	1.23 (0.97, 1.57)	0.09
IGP37	0.98 (0.59, 1.63)	0.94	0.99 (0.74, 1.33)	0.96	1.18 (0.97, 1.44)	0.10	1.36 (1.07, 1.71)	0.01
IGP38	0.99 (0.59, 1.65)	0.96	1.01 (0.75, 1.36)	0.93	1.18 (0.98, 1.44)	0.09	1.35 (1.07, 1.71)	0.01
IGP39	0.91 (0.55, 1.49)	0.71	1.22 (0.91, 1.63)	0.18	1.11 (0.91, 1.36)	0.29	1.05 (0.85, 1.30)	0.64
IGP40	0.91 (0.55, 1.50)	0.71	1.21 (0.90, 1.61)	0.21	1.11 (0.91, 1.36)	0.29	1.05 (0.85, 1.30)	0.66
Neutral IgG glycans								
IGP43	1.20 (0.70, 2.05)	0.51	1.21 (0.90, 1.63)	0.21	1.22 (1.02, 1.56)	0.03	1.12 (0.90, 1.40)	0.30
IGP45	0.95 (0.55, 1.63)	0.85	1.37 (1.02, 1.83)	0.03	1.08 (0.87, 1.33)	0.51	0.96 (0.77, 1.19)	0.69
IGP47	0.87 (0.51, 1.47)	0.60	0.74 (0.56, 0.99)	0.04	0.92 (0.75, 1.12)	0.41	0.90 (0.74, 1.10)	0.32
IGP48	0.78 (0.46, 1.31)	0.34	0.67 (0.51, 0.88)	0.003	0.82 (0.67, 0.99)	0.04	1.16 (0.91, 1.47)	0.23

IGP49	0.87 (0.53, 1.45)	0.60	0.99 (0.76, 1.29)	0.93	0.90 (0.74, 1.10)	0.30	0.82 (0.65, 1.04)	0.11
IGP53	1.00 (0.58, 1.72)	0.99	0.77 (0.57, 1.04)	0.09	0.79 (0.63, 0.98)	0.04	0.88 (0.72, 1.09)	0.26
IGP54	0.88 (0.51, 1.50)	0.63	0.91 (0.69, 1.20)	0.49	0.85 (0.70, 1.04)	0.12	0.89 (0.70, 1.12)	0.32
Galactosylation								
IGP55	1.14 (0.66, 1.97)	0.64	1.30 (0.96, 1.75)	0.09	1.27 (1.02, 1.59)	0.03	1.10 (0.89, 1.36)	0.40
IGP56	0.76 (0.45, 1.29)	0.32	0.73 (0.55, 0.98)	0.04	0.82 (0.66, 1.02)	0.07	0.95 (0.77, 1.18)	0.65
IGP57	0.96 (0.56, 1.65)	0.88	0.82 (0.61, 1.10)	0.18	0.78 (0.62, 0.97)	0.03	0.90 (0.73, 1.12)	0.36
Core fucosylation and bisecting GlcNAc								
IGP62	1.19 (0.69, 2.05)	0.53	0.80 (0.61, 1.05)	0.11	1.06 (0.88, 1.29)	0.52	1.04 (0.81, 1.33)	0.78
IGP63	1.23 (0.73, 2.06)	0.44	0.87 (0.65, 1.15)	0.33	1.10 (0.91, 1.33)	0.32	1.15 (0.91, 1.45)	0.25
IGP64	0.99 (0.58, 1.68)	0.96	0.80 (0.61, 1.06)	0.12	1.03 (0.85, 1.24)	0.77	1.11 (0.85, 1.43)	0.45
IGP66	0.88 (0.52, 1.47)	0.61	1.18 (0.90, 1.55)	0.22	0.96 (0.79, 1.17)	0.68	0.86 (0.69, 1.09)	0.21
IGP67	0.84 (0.50, 1.38)	0.49	1.14 (0.86, 1.50)	0.36	0.90 (0.74, 1.09)	0.29	0.84 (0.67, 1.04)	0.12
IGP68	0.98 (0.58, 1.64)	0.93	1.19 (0.91, 1.55)	0.19	0.98 (0.81, 1.19)	0.84	0.85 (0.67, 1.09)	0.20
IGP70	0.87 (0.51, 1.47)	0.61	1.22 (0.92, 1.60)	0.16	0.95 (0.78, 1.15)	0.61	0.87 (0.68, 1.10)	0.25
IGP71	0.87 (0.52, 1.46)	0.59	1.20 (0.92, 1.58)	0.18	0.95 (0.79, 1.16)	0.64	0.88 (0.70, 1.11)	0.29
IGP72	1.17 (0.70, 1.98)	0.55	0.83 (0.63, 1.09)	0.17	1.06 (0.88, 1.29)	0.54	1.12 (0.89, 1.42)	0.33

An increase in the percentage of agalactosylated structures ($G0^n$) and a decrease in mono- and di-galactosylated structures ($G1^n$, $G2^n$) was associated with poorer all-cause and CRC-specific mortality in stages 1, 2 and 3 (p-values from all-cause mortality models for $G0^n$: stage 1: 0.05, stage 2: 0.009 and stage 3: 0.01) but not in stage 4 (p-value for $G0^n$: 0.38).

In contrast, decrease in sialylation and increase in incidence of bisecting GlcNAc were statistically significantly associated with all-cause and CRC-specific mortality only in stage 4 (p-values from all-cause mortality models for stage 4 for FGS/(FG+FGS): 0.03; FGS/(F+FG+FGS): 0.01; FG2S1/(FG2+FG2S1+FG2S2): FBG2S1/(FBG2+FBG2S1+FBG2S2): 0.008; FBS1/FS1: 0.008).

Finally, only in stage 2 disease IgG glycans linked to core fucosylation were associated with all-cause and CRC-specific mortality **(Tables 9 and 10).**

Multivariate Cox regression clinical algorithms (including all the covariates of model III) showed good prediction of subsequent all cause (Harrell's C=0.73, AUC= 0.75, IDI=0.02 [as compared to model II that included AJCC stage, age and sex]) and CRC-mortality (Harrell's C=0.77, AUC= 0.79, IDI=0.04 [as compared to model II that included AJCC stage, age and sex]). Using glycans in addition to the clinical factors (that were selected by generalised boosted regression) did not lead to any statistically significant improvements for the whole sample analysis **(Table 11)** or after stage stratification **(Tables 12 and 13)**. This was reconfirmed by using Cox regression with L1 (LASSO) penalties on model parameters [191], as there were no significant differences in the validation deviances of models with and without glycans both for the whole sample and stage-stratified designs.

Table 11. Multivariate Cox regression of the a) clinical parameters and b) clinical and glycan parameters.

Clinical algorithm	All-cause mortality		Clinical algorithm	CRC mortality	
	HR (95% CI)	p-value		HR (95% CI)	p-value
Age	1.03 (1.01-1.04)	$9.2x10^{-6}$	Age	1.02 (1.00-1.03)	0.02
Sex	0.91 (0.74-1.12)	0.36	Sex	1.03 (0.82-1.29)	0.82
AJCC stage 2 vs 1	1.30 (0.92-1.83)	0.14	AJCC stage 2 vs 1	2.22 (1.33-3.72)	0.002
AJCC stage 3 vs 1	2.44 (1.77-3.38)	$7.3x10^{-8}$	AJCC stage 3 vs 1	5.02 (3.09-8.15)	$6.7x10^{-11}$
AJCC stage 4 vs 1	15.92 (11.16-22.70)	$<2.0x10^{-16}$	AJCC stage 4 vs 1	34.05 (20.60-56.29)	$<2.0x10^{-16}$
CRP	1.96 (1.47-2.62)	$4.9 x10^{-6}$	CRP	2.08 (1.52-2.86)	$4.8 x10^{-6}$
BMI	1.04 (1.01-1.06)	0.002	BMI	1.05 (1.02-1.08)	0.0001
Harrell's C	*0.73*		*Harrell's C*	*0.77*	
*IDI**	*0.02*		*IDI**	*0.04*	
AUC	0.74		*AUC*	0.79	
Clinical/glycans algorithm	**All-cause mortality**		**Clinical/glycans algorithm**	**CRC mortality**	
	HR (95% CI)	**p-value**		**HR (95% CI)**	**p-value**
Age	1.02 (1.01-1.03)	0.003	Age	1.01 (0.99-1.02)	0.28
Sex	0.91 (0.74-1.12)	0.36	Sex	1.04 (0.82-1.30)	0.77
AJCC stage 2 vs 1	1.33 (0.94-1.88)	0.11	AJCC stage 2 vs 1	2.24 (1.34-3.74)	0.002
AJCC stage 3 vs 1	2.49 (1.80-3.45)	$4.1x10^{-8}$	AJCC stage 3 vs 1	5.01 (3.09-8.14)	$7.1x10^{-11}$
AJCC stage 4 vs 1	15.72 (10.99-22.49)	$<2.0x10^{-16}$	AJCC stage 4 vs 1	33.63 (20.18-56.04)	$<2.0x10^{-16}$
CRP	1.66 (1.23-2.25)	0.001	CRP	1.84 (1.31-2.59)	0.0004
BMI	1.03 (1.01-1.06)	0.006	BMI	1.05 (1.02-1.07)	0.0007
IGP48	1.69 (0.64-4.47)	0.29	IGP29	0.92 (0.81-1.05)	0.21

IGP26	0.72 (0.55-0.95)	0.02	IGP13	0.85 (0.74-0.97)	0.01
IGP8	0.53 (0.21-1.30)	0.16	IGP8	0.91 (0.81-1.03)	0.14
Harrell's C	*0.73*		*Harrell's C*	*0.77*	
*IDI***	*0.05*		*IDI***	*0.03*	
AUC	0.75		*AUC*	0.79	

* The IDI was calculated based on the comparison of model II (adjusted for stage, sex and age) and the full clinical model III (adjusted for stage, age, sex, bmi and CRP – presented here).
** The IDI was calculated based on the comparison of the full clinical model III (adjusted for stage, age, sex, bmi and CRP) and the full clinical model III with the three top selected glycans.

Table 12. Multivariate Cox regression and estimate of the Harrell's concordance coefficient of the a) clinical parameters and b) clinical and glycan parameters by AJCC stage for all-cause mortality.

	Stage 1		Stage 2		Stage 3		Stage 4	
Clinical algorithm	**All-cause mortality**		**All-cause mortality**		**All-cause mortality**		**All-cause mortality**	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
Age	1.07 (1.03, 1.11)	0.0006	1.04 (1.02, 1.07)	0.0009	1.02 (1.00, 1.04)	0.03	1.01 (0.99, 1.03)	0.44
Sex	0.73 (0.40, 1.34)	0.31	0.96 (0.63, 1.46)	0.85	0.95 (0.68, 1.33)	0.77	0.94 (0.63, 1.40)	0.76
CRP	2.84 (1.20, 6.78)	0.02	1.96 (1.06, 3.63)	0.18	1.70 (1.02, 2.83)	0.04	1.95 (1.18, 3.22)	0.01
BMI	1.02 (0.95, 1.10)	0.61	1.09 (1.04, 1.14)	0.0005	1.01 (0.97, 1.05)	0.49	1.04 (0.99, 1.08)	0.10
Harrell's C	*0.68*		*0.65*		*0.55*		*0.61*	
IDI	*n/a**		*0.09*[§]		*0.08*[§]		*0.13*[§]	
AUC	*n/a**		*0.66*		*0.58*		*0.63*	
Clinical & glycans algorithm	**All-cause mortality***		**All-cause mortality**		**All-cause mortality**		**All-cause mortality**	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
Age			1.04 (1.01, 1.06)	0.006	1.01 (0.99, 1.03)	0.23	1.00 (0.98, 1.02)	0.45
Sex			0.98 (0.64, 1.50)	0.93	0.94 (0.67, 1.32)	0.73	1.00 (0.66, 1.52)	0.67
CRP			2.33 (1.24, 4.39)	0.01	1.47 (0.88, 2.47)	0.14	1.53 (0.86, 2.73)	0.03
BMI			1.07 (1.02, 1.12)	0.003	1.01 (0.97, 1.05)	0.65	1.03 (0.98, 1.08)	0.05
Top Glycan 1[**]			0.59 (0.46, 0.76)	$2.6x10^{-5}$	1.06 (0.85,1.32)	0.61	1.11 (0.81, 1.51)	0.53
Top Glycan 2[**]			0.95 (0.77, 1.17)	0.63	1.39 (1.03, 1.89)	0.03	0.98 (0.72, 1.33)	0.90
Top Glycan 3[**]			1.06 (0.83, 1.35)	0.62	0.75 (0.55, 1.01)	0.06	0.77 (0.58, 1.02)	0.07
Harrell's C			*0.67*		*0.53*		*0.61*	
IDI			*0.15*[§§]		*0.12*[§§]		*0.41*[§§]	
AUC			0.72		*0.60*		*0.69*	

* Due to the low number of observations, cross-validation was not possible and therefore we could not calculate the IDI and AUC values. Harrell's C coefficient were calculated based on the fitting all dataset

§ The IDI was calculated based on the comparison of the model II (adjusted for AJCC, age and sex) and the full clinical model III (adjusted for stage, age, sex, BMI and CRP – presented here).

**** Top Glycans for: Stage 2 IGP48, IGP18, IGP43; Stage 3 IGP43, IGP29, IGP24; Stage 4 IGP27, IGP49, IGP17**

§§ The IDI was calculated based on the comparison of full clinical model III (adjusted for stage, age, sex, bmi and CRP) and the model with these clinical factors plus the three top selected glycans.

Table 13. Multivariate Cox regression and estimate of the Harrell's concordance coefficient of the a) clinical parameters and b) clinical and glycan parameters by AJCC stage for CRC mortality.

	Stage 1		Stage 2		Stage 3		Stage 4	
Clinical algorithm	**CRC mortality**		**CRC mortality**		**CRC mortality**		**CRC mortality**	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
Age	1.06 (1.00, 1.12)	0.05	1.03 (1.00, 1.06)	0.05	1.01 (0.99-1.03)	0.23	1.01 (0.99, 1.03)	0.48
Sex	0.59 (0.21, 1.64)	0.31	1.35 (0.82, 2.22)	0.24	1.02 (0.71-1.46)	0.92	0.97 (0.65, 1.44)	0.86
CRP	1.04 (0.93, 1.16)	0.49	2.09 (1.02, 4.3)	0.04	1.92 (1.13-3.25)	0.02	1.04 (0.99, 1.08)	0.09
BMI	2.52 (0.57, 11.19)	0.22	1.13 (1.07, 1.19)	$7.9x10^{-6}$	1.02 (0.98-1.06)	0.34	1.04 (0.99, 1.08)	0.07
Harrell's C	*0.68*		*0.63*		*0.55*		*0.56*	
IDI	*n/a**		*0.11*[§]		*0.06*[§]		*0.12*[§]	
AUC	*n/a**		*0.68*		*0.56*		*0.63*	
Clinical & glycans algorithm	**CRC mortality***		**CRC mortality**		**CRC mortality**		**CRC mortality**	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
Age			1.02 (0.99,1.05)	0.17	1.01 (0.99, 1.03)	0.22	1.01 (0.99, 1.03)	0.42
Sex			1.38 (0.83,2.27)	0.21	1.03 (0.71, 1.49)	0.87	1.09 (0.71, 1.66)	0.70
CRP			2.32 (1.10,4.87)	0.03	1.80 (1.06, 3.07)	0.03	2.22 (1.21, 4.06)	0.01
BMI			1.11 (1.05,1.17)	8.34E-05	1.02 (0.98, 1.06)	0.39	1.04 (0.99, 1.09)	0.12

Top Glycan 1**			0.81 (0.24,2.68)	0.73	0.91 (0.62, 1.34)	0.64	1.17 (0.87, 1.59)	0.29
Top Glycan 2**			0.87 (0.59,1.29)	0.50	1.15 (0.95, 1.40)	0.16	0.70 (0.55,0.90)	0.005
Top Glycan 3**			0.88 (0.29,2.63)	0.82	0.97 (0.66, 1.41)	0.87	0.87 (0.67, 1.13)	0.30
Harrell's C			*0.61*		*0.51*		*0.61*	
IDI			*0.14§§*		*0.08§§*		*0.43§§*	
AUC			*0.71*		*0.59*		*0.71*	

* Due to the low number of observations, cross-validation was not possible and therefore we could not calculate the IDI and AUC values. Harrell's C coefficient were calculated based on the fitting all dataset

§ The IDI was calculated based on the comparison of the model II (adjusted for AJCC, age and sex) and the full clinical model III (adjusted for stage, age, sex, BMI and CRP – presented here).

**** Top Glycans for: Stage 2 IGP48, IGP56, IGP8; Stage 3 IGP67, IGP29, IGP49. IGP49, IGP63 and IGP9 were prioritised equally by generalised boosted approach, but results for IGP49 is only presented; Stage 4 IGP27, IGP24, IGP49**

§§ The IDI was calculated based on the comparison of full clinical model III (adjusted for stage, age, sex, bmi and CRP) and the model with these clinical factors plus the three top selected glycans

Similarly, predictions of the 5-year risk of death using the clinical factors stage, age, sex, BMI and CRP (e.g. AUC=0.80, Positive Predictive Value [PPV or precision] =0.80, using the Naïve Bayes classifier with a kernel density estimator for the marginal distributions) were not significantly improved by the addition of glycans data to the clinical factors **Tables 14 and 15**.

Table 14. Predictions of 5 year risk of CRC death for models with clinical factors and clinical and glycan factors using k-nearest neighbours, LASSO, Naïve Bayes, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Clinical model with age, sex and stage (n=950)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7379	-	0	1
k Nearest Neighbours	0.6619	0.7284	0.4833	0.5221	0.8017
LASSO	0.7786	0.8189	0.8738	0.3614	0.9815
Naive Bayes normal	0.7627	0.8179	0.8654	0.3614	0.9800
Naive Bayes kernel	0.7587	0.8179	0.8654	0.3614	0.9800
PAM	0.7626	0.7379	-	0	1
SVM linear	0.6831	0.7253	0.4805	0.5944	0.7718
SVM quadratic	0.6854	0.7000	0.4503	0.6546	0.7161
SVM cubic	0.7047	0.6884	0.4434	0.7390	0.6705
SVM RBF	0.7014	0.7389	0.5016	0.6125	0.7803
Decision Trees	0.6430	0.8189	0.8738	0.3614	0.9815
Boosted stumps	0.7691	0.8200	0.8750	0.3655	0.9815
Clinical model with age, sex, stage, BMI, CRP (n=950)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7379	-	0	1
k Nearest Neighbours	0.6310	0.7211	0.4661	0.4418	0.8203
LASSO	0.8052	0.8105	0.7315	0.4378	0.9429
Naive Bayes normal	0.8076	0.8116	0.7108	0.4739	0.9315
Naive Bayes kernel	0.7962	0.8095	0.7931	0.3695	0.9658
PAM	0.8039	0.7379	-	0	1
SVM linear	0.7237	0.7337	0.4944	0.7028	0.7447
SVM quadratic	0.7096	0.7568	0.5315	0.6104	0.8088
SVM cubic	0.6884	0.7274	0.4840	0.6064	0.7703
SVM RBF	0.6965	0.7221	0.4776	0.6426	0.7504
Decision Trees	0.6476	0.8189	0.8738	0.3614	0.9815
Boosted stumps	0.7978	0.8189	0.7770	0.4337	0.9558

Table 15. Predictions of 5 year risk of CRC death for models with the extended set of clinical factors with and without glycans using k-nearest neighbours, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Clinical model with age, sex, BMI, CRP, type of operation, time between operation and blood collection, and stage of cancer (n=949)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7379	-	0	1
k Nearest Neighbours	0.6174	0.7081	0.4398	0.4274	0.8074
LASSO	0.8042	0.8061	0.7192	0.4234	0.9415
PAM	0.8036	0.7492	0.9167	0.0444	0.9986
SVM linear	0.7229	0.7292	0.4875	0.7097	0.7361
SVM cubic	0.7047	0.7408	0.5032	0.6290	0.7803
SVM RBF	0.6764	0.7144	0.4639	0.5968	0.7561
Decision Trees	0.6331	0.8188	0.8725	0.3589	0.9815
Boosted stumps	0.8086	0.8124	0.7431	0.4315	0.9472
Clinical model with age, sex, BMI, CRP, type of operation, time between operation and blood collection, stage of cancer, and log-transformed glycans (n=949)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7387	-	0	1
k Nearest Neighbours	0.5713	0.6881	0.3857	0.3266	0.8160
LASSO	0.7980	0.8093	0.7557	0.3992	0.9544
PAM	0.6918	0.7576	0.6667	0.1452	0.9743
SVM linear	0.7068	0.7208	0.4759	0.6774	0.7361
SVM cubic	0.6449	0.7218	0.4688	0.4839	0.8060
SVM RBF	0.5	0.7387	-	0	1
Decision Trees	0.6447	0.8188	0.8725	0.3589	0.9815
Boosted stumps	0.7849	0.8072	0.7305	0.4153	0.9458

When patients were stratified by stage, adding glycans to the clinical variables improved the prediction results, did not change them, or made them worse, depending on the stage of cancer and on the chosen models. We tested whether independently of the choice of a model class, adding glycans to clinical covariates would improve predictions of a model of the same class estimated on independent test data using cross-validation. We performed two instances of the paired Wilcoxon sign-rank test comparing models with and without glycans, including all the considered models (W), or including only the models of disparate classes (Wd) as discussed in Methods. We showed that there was no significant improvement in the prediction of the rapid progressors using glycans (in addition to the clinical factors) for stage 2 ($p_W \sim 0.99$, $p_{Wd} \sim 0.98$) as measured by cumulative (merged) AUC on the validation data **(Table 16).** Similarly, for stage 3 the impact of the glycans was not consistent across the models, varied depending on the modelling assumptions, and was not significant overall ($p_W \sim 0.75$, $p_{Wd} \sim 0.58$; **(Table 17).**

Table 16. Predictions of rapid progressors in stage 2 for models with extending set of clinical factors with and without glycan using k-nearest neighbours, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Clinical model for stage 2 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, and stage of cancer (n=326)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.9356	-	0	1
k Nearest Neighbours	0.5863	0.8896	0.2000	0.2381	0.9344
LASSO	0.7820	0.9356	0.5000	0.0476	0.9967
PAM	0.7066	0.9356	-	0	1
SVM linear	0.7440	0.7699	0.1786	0.7143	0.7738
SVM quadratic	0.6939	0.8006	0.1765	0.5714	0.8164
SVM cubic	0.6200	0.8282	0.1569	0.3810	0.8590
SVM RBF	0.6331	0.8528	0.1860	0.3810	0.8852
Decision Trees	0.4746	0.9356	-	0	1
Boosted stumps	0.6911	0.9325	0.4286	0.1429	0.9869
Clinical model for stage 2 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, stage of cancer, and log-transformed glycans (n=326)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.9356	-	0	1
k Nearest Neighbours	0.5912	0.8988	0.2273	0.2381	0.9443
LASSO	0.7369	0.9356	0.5000	0.0476	0.9967
PAM	0.6623	0.9356	-	0	1
SVM linear	0.6537	0.7669	0.1429	0.5238	0.7836
SVM quadratic	0.5674	0.8957	0.1905	0.1905	0.9443
SVM cubic	0.4976	0.8896	0.0588	0.0476	0.9475
SVM RBF	0.5	0.9356	-	0	1
Decision Trees	0.4746	0.9356	-	0	1
Boosted stumps	0.5864	0.9294	0.2500	0.0476	0.9902

Table 17. Predictions of rapid progressors in stage 3 for models with the extended set of clinical factors with and without glycans using k-nearest neighbour, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Clinical model for stage 3 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, and stage of cancer (n=312)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.8782	-	0	1
k Nearest Neighbours	0.4438	0.7596	0.0256	0.0263	0.8613
LASSO	0.6259	0.8782	-	0	1
PAM	0.4802	0.8782	-	0	1
SVM linear	0.5864	0.7115	0.1905	0.4211	0.7518
SVM quadratic	0.5412	0.6122	0.1453	0.4474	0.6350
SVM cubic	0.5133	0.6827	0.1325	0.2895	0.7372
SVM RBF	0.5360	0.6827	0.1494	0.3421	0.7299
Decision Trees	0.4901	0.8782	-	0	1
Boosted stumps	0.5965	0.8782	-	0	1
Clinical model for stage 3 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, stage of cancer, and log-transformed glycans (n=312)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.8782	-	0	1
k Nearest Neighbours	0.4972	0.7949	0.1176	0.1053	0.8905
LASSO	0.5953	0.8782	-	0	1
PAM	0.5645	0.8814	1	0.0263	1
SVM linear	0.5941	0.7051	0.1932	0.4474	0.7409
SVM quadratic	0.5169	0.7885	0.1500	0.1579	0.8759
SVM cubic	0.5107	0.8173	0.1481	0.1053	0.9161
SVM RBF	0.5000	0.8782	-	0	1
Decision Trees	0.4745	0.8782	-	0	1
Boosted stumps	0.4483	0.8494	0.2353	0.1053	0.9526

On the other hand, there was a significant improvement in the prediction of the rapid progressors using glycans for stage 4 **(Table 18),** with $p_W \sim 0.01$, $p_{Wd} \sim 0.04$, leading to the median gain in the test AUC of 0.08. Importantly, the inclusion of glycans in the models consistently resulted in the improved quality of predictions across the range of the considered models, and independently of whether the restricted or extended sets of clinical variables were used in the adjustments. The results were qualitatively similar for multiple repetitions of 10-fold cross-validation with random partitions into non-overlapping test folds, and independently of whether 10-fold or two-fold cross-validation was used to estimate the AUC on test data for the considered models. The best extended clinical model had the test AUC of 0.58, with the PPV of 0.35. The best model augmented with unfiltered log-transformed glycans had the test AUC of 0.66, with the PPV of 0.62.

Table 18. Predictions of rapid progressors in stage 4 for models with the extended set of clinical factors with and without glycans using k-nearest neighbour, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Clinical model for stage 4 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, and stage of cancer (n=102)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7059	-	0	1
k Nearest Neighbours	0.5	0.5686	0.2941	0.3333	0.6667
LASSO	0.4331	0.6961	0.3333	0.0333	0.9722
PAM	0.4535	0.7059	-	0	1
SVM linear	0.5514	0.5588	0.3404	0.5333	0.5694
SVM quadratic	0.5819	0.6569	0.4138	0.4000	0.7639
SVM cubic	0.5472	0.6078	0.3529	0.4000	0.6944
SVM RBF	0.5278	0.6078	0.3333	0.3333	0.7222
Decision Trees	0.4899	0.7059	-	0	1
Boosted stumps	0.5505	0.6471	0.3125	0.1667	0.8472
Clinical model for stage 4 with age, sex, BMI, CRP, type of operation, time between operation and blood collection, stage of cancer, and log-transformed glycans (n=102)					
	AUC	**Accuracy**	**PPV**	**Sensitivity**	**Specificity**
Maximum Prior	0.5	0.7059	-	0	1
k Nearest Neighbours	0.5819	0.6569	0.4138	0.4000	0.7639
LASSO	0.5815	0.6863	0	0	0.9722
PAM	0.5257	0.6765	0	0	0.9583
SVM linear	0.5972	0.6373	0.4054	0.5000	0.6944
SVM quadratic	0.5778	0.6373	0.3939	0.4333	0.7222
SVM cubic	0.6486	0.7647	0.6875	0.3667	0.9306
SVM RBF	0.5000	0.7059	-	0	1
Decision Trees	0.6569	0.7353	0.6154	0.2667	0.9306
Boosted stumps	0.5565	0.6863	0.4500	0.3000	0.8472

5.3. IgG glycome composition in CRC patients and matching controls.

To test whether alteration on IgG glycome composition between CRC patients and controls could potentially have relevance as a clinically useful biomarker of CRC risk, analyses of clinical characteristics among 760 patients and 538 matching controls were done.

Descriptive information on CRC patients and healthy controls are presented in **(Table 19).**
In addition to 24 directly measured glycan structures, 12 derived traits were calculated from the directly measured glycans. These derived traits average particular glycosylation features (galactosylation, fucosylation, sialylation) across different individual glycan structures and consequently they are more related to individual enzymatic activities and underlying genetic polymorphisms **(Table 20).**

Table 19. Descriptive information on CRC patient and healthy controls

	Cases (n = 760)	Control (n = 538)	p-value
Age (median[IQR])	52 (48-56)	53 (48-56)	0,274
Sex (Men/Women)	415 (54.6%) / 345 (45.4%)	289 (53.7%) / 248 (46.2%)	0,821
Smoking status (Current/Ex/ Non/unknown)	133/172/271/184	94/139/211/103	0,836
BMI (median[IQR])	26,3 (23,4-26,8)	27,8 (25,8-28,6)	2,69E-10
Family History (Low/Medium or High/unknown)	524/197/39	517/6/15	6,50E-44

Table 20. Derived glycan traits

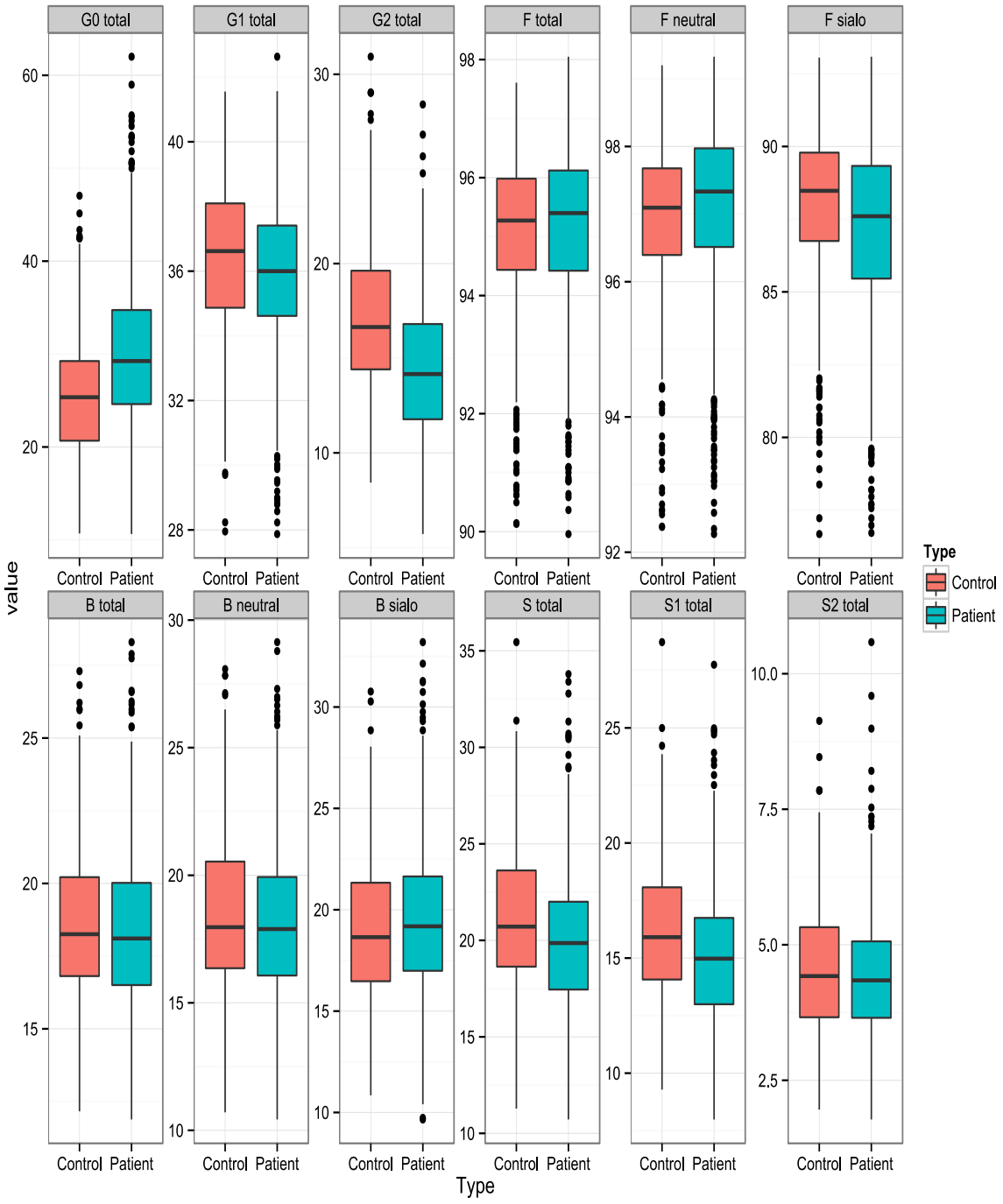
Derived glycan traits	Description	Formula
G0 total	Proportion of agalactosylated structures in total in total IgG glycans	(GP1+GP2+GP4+GP6)/GPt
G1 total	Proportion of monogalactosylated structures in total IgG glycans	(GP7+GP8+GP9+GP10+GP11)/GPt
G2 total	Proportion of digalactosylated structures in total IgG glycans	(GP12+GP13+GP14+GP15)/GPt
F total	Proportion of fucosylated structures in total IgG glycans	(GP1+GP4+GP6+GP8+GP9+GP10+GP11+ GP14+GP15+GP16+GP18+GP19+GP23+GP24)/GPt
F neutral	Proportion of fucosylated structures in total neutral glycans	(GP1+GP4+GP6+GP8+GP9+GP10+GP11+GP14+GP15)/GPn
F sialo	Proportion of fucosylated structures in total sialylated glycans	(GP16+GP18+GP19+GP23+GP24)/GPs
B total	Proportion of structures with bisecting GlcNAc in total IgG glycans	(GP3+GP6+GP10+GP11+GP13+GP15+GP19+GP22+GP24)/GPt
B neutral	Proportion of structures with bisecting GlcNAc in total neutral IgG glycans	(GP3+GP6+GP10+GP11+GP13+GP15)/GPn
B sialo	Proportion of structures with bisecting GlcNAc in total sialylated IgG glycans	(GP19+GP22+GP24)/GPs
S total	Proportion of sialylated structures in total IgG glycans	(GP16+GP17+GP18+GP19+GP20+GP21+GP22+GP23+GP24)/GPt
S1 total	Proportion of monosialylated structures in total IgG glycans	(GP16+GP17+GP18+GP19)/GPt
S2 total	Proportion of disialylated structures in total IgG glycans	(GP21+GP22+GP23+GP24)/GPt

Significant differences were observed in several features of the glycome **(Table 21, Figure 7),** primarily reflecting decreased galactosylation (OR=2.35, p=2.39E-22 for G0 and OR=0.36, p=6.59E-29 for G2) and sialylation (OR=0.72, p=2.73E-05 for S total), as well as increased fucosylation of neutral IgG glycans (OR=1.24, p=3.57E-03 for F total) and decreased fucosylation of sialylated glycans (OR=0.72, p=5.85E-05 for F sialo).

Table 21. IgG glycome composition in CRC patients and controls. Only the main derived traits describing glycome composition are shown.

Glycan	Control (median[IQR])		Patient (median[IQR])		OddsRatio	95%ConfInt	p.value	p.adjusted
G0 total	25,35	(20,67-29,24)	29,24	(24,62-34,74)	2,35	(1,95 - 2,82)	2,65E-23	2,39E-22
G1 total	36,62	(34,87-38,09)	35,94	(34,49-37,39)	0,64	(0,55 - 0,74)	8,09E-10	3,64E-09
G2 total	16,65	(14,41-19,62)	14,12	(11,71-16,78)	0,36	(0,30 - 0,44)	1,83E-30	6,59E-29
F total	95,27	(94,42-95,97)	95,39	(94,40-96,11)	1,11	(0,97 - 1,28)	1,27E-01	1,53E-01
F neutral	97,07	(96,38-97,66)	97,32	(96,50-97,97)	1,24	(1,08 - 1,43)	1,98E-03	3,57E-03
F sialo	88,47	(86,69-89,79)	87,58	(85,39-89,29)	0,72	(0,62 - 0,84)	2,44E-05	5,85E-05
B total	18,25	(16,81-20,22)	18,11	(16,49-20,02)	0,88	(0,76 - 1,01)	7,68E-02	9,87E-02
B neutral	17,97	(16,35-20,54)	17,89	(16,07-19,93)	0,84	(0,73 - 0,97)	1,86E-02	2,91E-02
B sialo	18,64	(16,48-21,33)	19,19	(16,99-21,71)	1,17	(1,01 - 1,35)	3,85E-02	5,13E-02
S total	20,71	(18,64-23,62)	19,85	(17,39-21,96)	0,72	(0,62 - 0,83)	1,06E-05	2,73E-05
S1 total	15,9	(14,07-18,07)	14,97	(12,94-16,74)	0,64	(0,54 - 0,75)	9,16E-09	3,66E-08
S2 total	4,41	(3,66-5,32)	4,34	(3,65-5,08)	0,96	(0,84 - 1,11)	6,03E-01	6,03E-01

Figure 7. IgG glycome composition in CRC patients and matching controls.



IgG glycome was analysed in 760 patients with CRC and 538 matching controls. Main features of the IgG glycome are presented as box plots showing median values and 25% (box) and 75% (line) percentiles for patients and controls.

Additional information is available in **(Table 21 and 22).**

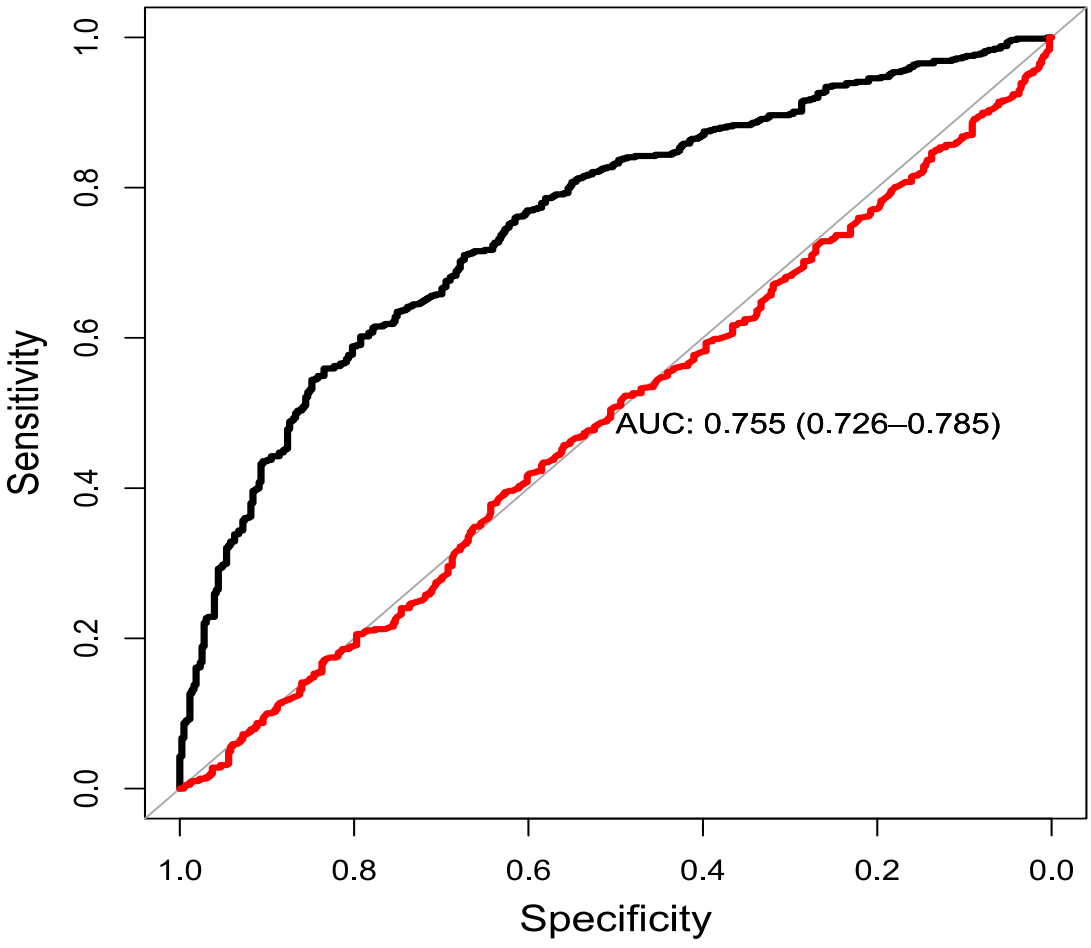
Table 22. IgG glycome composition in CRC patients and controls

Glycan peak	Control (median[IQR])		Patient (median[IQR])		OddsRatio	95%ConfInt	p.value	p.adjusted
GP1	0,12	(0,08-0,22)	0,15	(0,10-0,26)	1,28	(1,11 - 1,48)	5,20E-04	1,17E-03
GP2	0,49	(0,34-0,71)	0,54	(0,36-0,79)	1,18	(1,03 - 1,37)	2,10E-02	3,15E-02
GP3	0,37	(0,29-0,52)	0,46	(0,34-0,65)	1,5	(1,29 - 1,74)	4,79E-08	1,72E-07
GP4	18,68	(15,42-21,98)	22,34	(18,34-26,87)	2,33	(1,94 - 2,79)	1,83E-23	2,20E-22
GP5	0,29	(0,24-0,37)	0,32	(0,26-0,41)	1,26	(1,09 - 1,46)	1,30E-03	2,47E-03
GP6	4,97	(4,02-6,01)	5,39	(4,53-6,45)	1,44	(1,23 - 1,69)	4,13E-06	1,24E-05
GP7	0,57	(0,43-0,75)	0,54	(0,39-0,76)	0,88	(0,77 - 1,02)	8,42E-02	1,05E-01
GP8	18,9	(17,73-20,18)	18,82	(17,59-19,97)	0,85	(0,74 - 0,98)	2,43E-02	3,51E-02
GP9	9,87	(8,89-11,04)	9,66	(8,80-10,54)	0,78	(0,68 - 0,90)	5,52E-04	1,17E-03
GP10	5,62	(4,95-6,43)	5,37	(4,66-6,27)	0,72	(0,62 - 0,83)	4,65E-06	1,29E-05
GP11	0,82	(0,72-0,95)	0,8	(0,70-0,91)	0,78	(0,68 - 0,91)	8,05E-04	1,61E-03
GP12	0,68	(0,50-0,93)	0,55	(0,37-0,77)	0,6	(0,51 - 0,70)	9,93E-12	5,11E-11
GP13	0,45	(0,36-0,57)	0,4	(0,31-0,53)	0,68	(0,58 - 0,78)	1,09E-07	3,55E-07
GP14	13,59	(11,57-15,90)	11,45	(9,40-13,80)	0,37	(0,30 - 0,45)	4,21E-29	7,57E-28
GP15	1,73	(1,47-2,06)	1,47	(1,20-1,82)	0,49	(0,41 - 0,57)	6,94E-20	5,00E-19
GP16	3,27	(2,95-3,65)	3,37	(2,99-3,76)	1,17	(1,02 - 1,34)	2,65E-02	3,67E-02

GP17	0,94	(0,82-1,08)	0,92	(0,80-1,07)	0,92	(0,79 - 1,06)	2,39E-01	2,60E-01
GP18	9,52	(8,09-11,43)	8,47	(6,98-9,99)	0,56	(0,47 - 0,66)	8,05E-13	4,83E-12
GP19	1,88	(1,65-2,15)	1,81	(1,61-2,07)	0,82	(0,72 - 0,95)	5,63E-03	9,64E-03
GP20	0,36	(0,25-0,51)	0,38	(0,27-0,52)	1,12	(0,96 - 1,30)	1,41E-01	1,64E-01
GP21	0,89	(0,74-1,06)	0,92	(0,77-1,10)	1,11	(0,96 - 1,28)	1,71E-01	1,93E-01
GP22	0,16	(0,11-0,24)	0,15	(0,10-0,21)	0,84	(0,73 - 0,97)	1,62E-02	2,66E-02
GP23	1,45	(1,14-1,85)	1,4	(1,11-1,72)	0,96	(0,84 - 1,11)	5,96E-01	6,03E-01
GP24	1,78	(1,48-2,13)	1,74	(1,43-2,08)	0,95	(0,82 - 1,09)	4,44E-01	4,71E-01

As many glycan structures showed strong association with CRC, we attempted to build a predictive model using regularized logistic regression. Only the 24 directly measured glycan traits were used as predictors in the model. Evaluation of model performance was conducted using a 10-cross validation procedure. While a model based on age and sex did not show significant discriminative power (AUC = 0.499), the addition of glycan variables into the model considerably increased the discriminative power of the model (AUC = 0.755, P < 1x10E-16) **(Figure 8).**

Figure 8. ROC curve illustrating the performance of regularized logistic regression model in predicting disease status for CRC patients and healthy controls. While models based only on age and gender did not show predictive power (red line), addition of glycan traits increased predictive power of model (black line).



Glycome composition is known to change in acute inflammation [215] and to evaluate potential effects of surgery on the IgG glycome in CRC patients we analysed IgG glycome composition in 28 patients (i) before surgery, (ii) 24 after surgery), (iii) 48 h after surgery and (iv) 7 days after surgery. We did not observe any consistent and statistically significant changes in the IgG glycome that were caused by the surgery.

To determine whether the observed changes were present before the disease onset, we identified 39 individuals from the FINNRISK cohort that were sampled before the initial diagnosis. However, when compared to matching controls, no statistically significant changes were found.

6. DISCUSSION

This study represents the first comprehensive analysis of IgG glycosylation in CRC. By applying the recently developed method for high-throughput glyco-profiling of IgG [20] on a well-characterized large cohort of 1229 CRC patients and 538 matching controls an important gap in knowledge which may have relevance for understanding the immunosurveilance of CRC was filled [178].

The research presented here provides results regarding the relationship between the IgG glycome composition in plasma of CRC patients with survival outcomes and also it provides results about differences in IgG glycome composition between patients and controls.

Univariate and multivariate statistical models were applied to examine the associations between specific glycan changes and CRC-specific or all-cause mortality. IgG glycans linked to galactosylation, sialylation and bisecting GlcNAc were strongly associated with all-cause mortality and CRC mortality. Multivariate Cox regression clinical algorithms showed good prediction of outcome for all cause and CRC-mortality, but using glycans in addition to the clinical factors did not lead to any statistically significant improvements. However, when we investigated the prediction of rapid progressors within each AJCC stage, there was an improvement in the prediction of the rapid progressors using glycans for stages 3 and 4.

By analysing 760 CRC patients and 538 matching controls it is found that CRC is associated with three major alterations in the IgG glycome composition: **(i)** decrease in IgG galactosylation, **(ii)** decrease in IgG sialylation and **(iii)** increase in core-fucosylation of neutral glycans with concurrent decrease of core fucosylation of sialylated glycans.

6.1. Stage differences

It is well established that glycosylation changes are involved in the aetiology of cancer, and specifically mark tumor proliferation and metastasis [26]. IgG is produced and secreted by CRC cells and the expression levels of CRC-tumor derived IgG correlated with many clinical and pathological characteristics of the tumor (including stage) [192]. In particular it has been shown that expression of IgG was stronger in CRC tissues with TNM stage III–IV, than in those with TNM I–II. Similarly, in this study we observed different changes in IgG glycosylation status (levels of sialylation and incidence of bisecting GlcNAC) in late-stage disease and we saw an improvement in the prediction algorithms using glycans in addition to clinical factors in AJCC stage 3 and 4.

6.2. Galactosylation, sialylation, GlcNAc and core fucosylation residues present on IgG glycans in cancer

The immune system can identify and destroy new tumor cells through cancer immunosurveillance, which functions as an important defence against cancer. A recent review on the natural innate and adaptive immunity to cancer has presented evidence from mouse models that B cells (which create and release IgG) are important in the surveillance of the CRC [178]. Currently, it is known that inflammation is one of the features of cancer [193], but it is not known whether the inflammatory process mainly plays an important role in the development of cancer, or whether the cancer induces an inflammatory response, or both. IgG can demonstrate both pro and anti-inflammatory activity depending on its different glycan composition [81]. Alteration in IgG galactosylation, sialylation, bisecting GlcNAc and fucosylation have been previously reported in cancer studies **(Table 23).**. In particular, a decline in plasma IgG glycosylation has predominantly been observed with tumor progression and metastasis in gastric, lung, prostate and ovarian cancers [194, 164, 167, 162, 148, 161, 160, 195] as well as in chronic inflammatory diseases such as rheumatoid arthritis [143] and osteoarthritis [143], inflammatory bowel disease [196, 197], systemic lupus erithematosis [198], vasculitis [142] and some other disease [77]. All previous studies in cancer research had small sample sizes (<100 cancer cases) and this is the first time that similar changes in IgG galactosylation were observed in CRC prognosis in a study with more than 1000 CRC patients **(Table 23).**

Table 23. Studies on IgG glycosylation changes in cancer

Author	Year	Cancer	Method	Samples	Cases	Controls	Results
Galactosylation							
Kanoh Y[1]	2004	Prostate cancer	Fluorophore-assocd. carbohydrate electrophoresis (FACE)	serum	12	10	Fr 1 (monogalactosyl oligosaccharide) and Fr 2 (digalactosyl oligosaccharide) decreased significantly ($p<0.05$), while Fr 4 (agalactosyl IgG oligosaccharide) increased with PCa tumor progression. The Fr 4 / Fr 1 + 2 ratio in metastatic PCa patients was significantly higher than in healthy controls ($p<0.05$)
Aurer I[2]	2007	Multiple myeloma	Lectin blotting and densitometry	blood	16	16	IgG galactosylation was reduced in multiple myeloma
Kodar K[3]	2011	Gastric cancer	LC-ESI-MS	serum	80	51	Significant increase of agalactosylated (GnGnF, GnGn(bi)F), and decrease of galactosylated (AGn(bi), AGn(bi)F, AA(bi), AAF)
Bones J[4]	2011	Gastric cancer	Hydrophilic interaction liquid chromatography with fluorescence detection	serum	80	30	The data indicates that in the cancerous state there is a switch in IgG production toward the more pro-inflammatory IgG G0 glycoform (agalactosyl).
Gercel-Taylor C[5]	2001	Ovarian cancer	Concanavalin A affinity columns and sodium dodecyl sulfate-polyacrylamide gel electrophoresis	serum	62	50	This report demonstrated the presence of an aberrantly glycosylated IgG population in cancer patients.
Saldova R[6]	2007	Ovarian cancer	Quantitative NPHPLC and exoglycosidase digestion	serum	27	34	IgG containing agalactosylated structures (G0) (mostly represented by FA2) were doubled; monogalactosylated (G1) decreased; digalactosylated (G2) structures decreased
Alley WR[7]	2012	Ovarian cancer	MALDI-TOF Mass-spectrometric Analysis	serum	19	20	Increased levels of a-galactosylation structures were obsd. on N-linked glycans derived from IgG, which were independent of the presence of fucose residues.

Qian Y[8]	2013	Ovarian cancer	MALDI-TOF Mass-spectrometric Analysis	serum	32	26	G0/(G1 + G2·2) was found significantly higher in the malignant group than in the benign group (0.74 vs 0.34; $p < 0.0001$)
Sialylation							
Flemming SC[9]	1998	Multiple myeloma	High pressure anion exchange chromatography with pulsed electrochemical detection (HPAE-PED)	serum	47	14	Patients with myeloma showed an increase in the proportion of sialylated oligosaccharides in comparison with patients with MGUS
Saldova R[6]	2007	Ovarian cancer	Quantitative NPHPLC and exoglycosidase digestion	serum	27	34	The overall sialylation decreased
Kodar K[3]	2011	Gastric cancer	LC-ESI-MS	serum	80	51	Decrease of monosialylated IgG glycoforms (NaAF, NaA(bi)) in cancer patients.
Bisecting GlcNAC							
Kodar K[3]	2011	Gastric cancer	LC-ESI-MS	serum	80	51	A statistically significant decrease of bisecting GlcNAc was observed in tumor stage II and III

It has been reported that the environmental factors could modulate glycosylation of IgG Fc [199]. Since the composition of IgG Fc glycans is dependent on the regulation of glycosyltransferases and glycosidases, it is speculated that the expression and activity of these enzymes are differentially regulated in response to stimulation of environmental factors Furthermore, it has been hypothesised that decreased IgG galactosylation leads to a more pro-inflammatory antibody response [77, 119], which might influence cancer survival after diagnosis. Decreased of IgG galactosylation is caused by decreased Gal T-activity in plasma cells [200] or due to increased production of specific clones of plasma cells with low expression levels of galactosyltransferases [201]. It was shown that the decrease of Gal-T in B- lymphocyte caused the increase of IgG G0 glycoforms in rheumatoid arthritis and other rheumatic diseases [202, 203]. Thus, it was hypothesized that Gal-T activity in plasma cells is down-regulated during tumor progression and that this causes the significant increase of agalactosylated IgG, indicating that in cancerous state there could be a switch in IgG production toward the more pro-inflammatory IgG G0 glycoforms. Interestingly, in rheumatoid arthritis patients an increase in galactosylation of IgG during combined treatment of infliximab (anti-TNF-α neutralizing antibody) and methotrexate was shown, indicating that TNF-α may be involved in regulating IgG glycosylation patterns [204, 205]. Furthermore, experimental and clinical studies on the role of TNF- α have demonstrated that the TNF- α is a key player in progression of human CRC [206, 207]. Increased levels of TNF- α in patients with CRC were associated with poor prognosis [208]. Similar to galactosylation, decreased sialylation of IgG results in a proinflammatory IgG phenotype [119]. The key enzyme for the addition of α2,6-sialic acid to glycan termini is β-galactoside α2,6-sialyltransferase 1 (ST6Gal-1) Removal of the enzyme ST6Gal-1 in a mouse model, has been reported to results in more severe pulmonary inflammation [209] consistent with the importance of sialic acid in anti-inflammation [81]. Likewise, the same pattern of decreased IgG galactosylation and sialylation also occurs with aging in the general population [210], generating the hypothesis that decrease in IgG galactosylation and sialylation in CRC may be indicative of an inflammatory state.

Currently, there is a gap in knowledge whether the disease affects the glycosylation of all or just antigen-specific IgG, which then are responsible for the overall change observed in total IgG. There have been reports that agalactosylated and asialylated antigen-specific IgG are involved in the pathogenesis of some diseases, such as anti-citrullinated protein antibodies in RA and anti-proteinase 3 antibodies in Wegener's granulomatosis [211, 212]. A study further showed that there is a variability between glycosylation pattern of antigen-specific IgG and

total IgG. In our study we found that in the same manner as decreased galactosylation is associated with poorer prognosis, decreased sialylation was also linked to poorer prognosis, which replicated findings of two small studies on ovarian [195] and gastric cancer [162], but the opposite was found, in a study of multiple myeloma [213] where increased sialylation of IgG was linked to higher risk of multiple myeloma. Therefore, through both decreased galactosylation and decreased sialylation, IgG in CRC patients with poorer prognosis had significantly greater pro-inflammatory properties (decreased galactosylation and sialylation) than CRC patients with better prognosis. Furthermore elevated occurrence of bisecting GlcNAc and lack of core fucose results in increased ADCC activity. In our study we found higher occurrence of bisecting GlcNAc in CRC patients of poorer prognosis, but IgG core fucosylation changes were associated with all-cause or CRC-specific mortality only in stage 2 CRC patients.

Likewise, when comparing CRC patients with matching controls, we have observed consistent decreases in all structures with two terminal galactoses (A2BG2 (GP13), FA2G2 (GP14) and FA2BG2 (GP15)) and an increase in structures without galactoses (A2 (GP2), A2B (GP3), FA2 (GP4) and FA2B (GP6)) **(Table 22).** The decrease in galactosylation was also evident in the derived traits G0 and G2 that average galactosylation of several individual glycans (G0: OR=2,35; p =2,39E-22; G2: OR=0,36; p=6,59E-29). Several directly measured IgG glycans containing sialic acid, as well as the derived trait "S1 total" (measuring all monosialylated IgG glycans) were also decreased in CRC patients **(Table 21).**

Since decreases in galactosylation and sialylation have also been observed in a number of other diseases [77], this pattern of glycan changes (which are consistent with a decrease in the immunosuppressive potential of IgG) is not specific for CRC.

Taking into consideration that little is known about the mechanisms of regulation of IgG glycosylation [214], and since IgG N-glycosylation is controlled through a complex interplay between loci affecting an overlapping spectrum of glycome measurements, and through interection of genes directly involved in glycosylation and those that presumably have 'higher-level''regulatory function [189], it is very difficult to speculate about potential mechanisms and causes of differences in IgG glycosylation in CRC patients. Another feature of human IgG glycome is its possibility to change particularly in the stituation of disturbed homeostasis [215], despite the fact that the heritability of IgG glycosylation is relatively high (up to 80%) [20]. The heritability of galactosylation is estimated to be between 40% and 70% and of sialylation between 30% and 60% [20, 210]. Besides this, galactosylation of IgG can change quite rapidly in acute inflammation [215]. Recently, it was reported that in rheumatic arthritis,

decreased galactosylation has been demonstrated to predate the onset of disease [197], indicating that they may be a part of the disease pathophysiology. In an analogous attempt to address causality in this study we were able to identify 39 individuals that were sampled before the initial diagnosis of CRC. However, when compared to matching controls (individuals of the same age that did not develop CRC in the same period after recruitment) we did not identify any statistically significant differences. Since the sample size was very small it was hard to derive any conclusions from this exploratory part of the study, beside the fact that we were not able to show any differences in IgG glycome composition before the onset of CRC.

It is unquestionable that N-linked IgG Fc glycans are essential in determining distinct effector functions. Recently, Barb, A.W.A et al. [216] in their study using Nuclear Magnetic Rezonance (NMR) revealed that glycans attached to Fc part of IgG are more exposed structurally than previously shown, providing detailed description of the interaction between glycans on IgG Fc and the respective receptor, and give further insight into the immunomodulatory role of IgG Fc glycans. Furthermore,as ADCC is one of the crucial mechanisms in killing tumor cells, activation of ADCC is believed to be an important mechanism of therapeutic monoclonal antibodies, as indicated by the fact that a common single-nucleotide polymorphism in FcγRIIIa (V158F) is correlated with clinical responses to cetuximab [217] and rituximab [218]. Core-fucose is added to glycans by the fucosyltransferase 8 enzyme (encoded by *FUT8* gene), which has recently been reported to be functionally relevant in some cancers [219, 220]. The levels of IgG molecules without core-fucose vary between 1.3% and 19% and we postulated that this may have a significant impact on antibody-dependent cellular cytotoxicity (ADCC) and thus capacity to eliminate cancer cells [20]. The recently developed therapeutic IgG antibodies for cancer therapy are glyco-engineered to remove core-fucose on the Fc glycans in order to increase clinical efficacy of monoclonal antibodies due to enhancement of their therapeutic effect through ADCC mediated killing [221, 222]. Recently it has been shown that properly glyco-engineered antibodies [221] can efficiently elicit ADCC even in immunocompromised CRC patients [223]. This observation indicates that efficient immunosurveilance of tumor cells depends on antibody/Fc receptor affinity. Thus, increased levels of core-fucose on neutral IgG glycans in CRC patients may influence disease risk and course by decreasing the ability of IgG to activate ADCC.

Interestingly, core-fucose was decreased in sialylated glycans. The increase in core fucose on neutral glycans, with concurrent decrease of core-fucose on sialylated glycans has not been previously reported. This is the first report of different direction of changes in fucosylation in sialylated and neutral glycans. Recent studies clearly demonstrated that some antigen-specific

antibodies can have significantly different levels of fucose [224] but it is hard to evaluate the importance in the different fucosylation of sialylated and neutral IgGs, since the relevance of sialylation on the impact of fucose on IgG function is currently not known [225]. This is further supported by the observation that polymorphisms in the *FUT8* gene seem to be associated with increased risk for CRC [226]. Additional evidence supporting a possible functional importance of FUT8 in colorectal cancer is the recent observation that micro RNA MiR-198 represses tumor growth and metastasis in colorectal cancer by targeting FUT8 [227]. The finding that similar pattern of changes in *FUT8* expression can be observed or inferred in both the tumor tissue and the antibody-producing B lymphocytes [226] points to the importance of general mechanisms controlling fucosylation in CRC and implies that the same features of genetic makeup influence glyco-gene expression (and thereby glycosylation profile) in both the tumor tissue and in B cells. This view is further strengthened by the observation of differential expression of the B-cell-specific transcription factor Ikaros IKZF1- which has been reported to be associated with the risk of various cancers [228] in tumor tissue of CRC patients.

7. CONCLUSION

Alteration in IgG glycosylation has a significant impact in the pathogenesis of numerous diseases providing insights into disease state and progression. Consequently, there is an indispensable need to search for new IgG glycan biomarkers that could serve as more sensitive diagnostic and prognostic tools which could be used to distinguish between different forms/stages of disease and to monitor the efficacy of various new treatment options. Inter-individual differences in IgG glycome are very important in this aspect, by virtue of their role in understanding the host defence reaction responses to the presence of the disease.

1. The plasma IgG glycan differences which we observed at the time of CRC diagnosis are consistent with significantly increased IgG pro-inflammatory activity being associated with poorer CRC prognosis, especially in late stage (stages 3 and 4) CRC. In the absence of validated biomarkers to improve upon prognostic information from existing clinicopathological factors the potential of these novel IgG glycan biomarkers merits further investigation.

In particular, the improved predictive power in models including glycan factors in stage 4 patients is interesting. Currently, there are various strategies that are employed when using chemotherapy [229] in stage 4 disease. Therefore having a novel biomarker or prediction model that could help selected patients that may have a better prognosis and a more indolent disease course would be useful as these patients could perhaps be offered sequential single agent chemotherapy with lower toxicity when compared to a more aggressive combination strategy. Furtheremore, there is a great interest in novel immunotherapies in cancer and therefore it would be useful to identify a more 'immunogenic' tumor based on identified IgG glyco-markers with a particular response to immunotherapy. Certainly to date the most encouraging results for immunotherapies (incl. PD-1 inhibitors) have been in tumors such as melanoma that are thought to be highly immunogenic and there is interest in investigating mismatch repair deficient colon cancer which are often associated histologically with a heavy immune infiltrate. Recent studies [215] demonstrated that IgG glycosylation changes are very dynamic and variable between individuals, thus longitudinal studies are needed to fully investigate the prognostic potential of IgG glycosylation changes in CRC.

2. Significant plasma IgG glycome composition differences are demonstrated between CRC patients and controls. We were not able to detect these differences in historical samples (taken

before CRC had developed in these patients). This could indicate that the changes are due to reverse causality (due to the disease process or treatment effects). However, it may also be due to inadequate study power in this small sub-study and so additional studies are required to investigate this further.

Considering the functional relevance of IgG glycosylation for both tumor immunosurveilance and clinical efficacy of therapy with monoclonal antibodies, individual variation in IgG glycosylation may turn out to be important for prediction of disease course or the choice of therapy, thus warranting further, more detailed studies of IgG glycosylation in CRC.

8. SAŽETAK / ABSTRACT IN CROATIAN

UVOD: Rak debelog crijeva je zloćudna novotvorina debelog crijeva i rektuma, te je još uvijek povezan s lošom prognozom, niskom stopom preživljavanja i vrlo često relativno kasnom dijagnozom.

MATERIJALI I METODE: Analizirali smo glikozilaciju imunoglobulina G (IgG) u populaciji od 1229 ljudi oboljenih od raka debelog crijeva te 538 odgovarajućih kontrolnih uzoraka. Utjecaj operacije na glikozilaciju procjenjen je kod 28 bolesnika uzorkovanih prije i tri puta nakon operacije. Nadalje, glikozilacija IgG-a analizirana je i u 39 uzoraka krvne plazme izuzete prije prve dijagnoze.

REZULTATI: Klinički algoritmi pokazali su dobro predviđanje ukupne smrtnosti kao i smrtnosti uzrokovane rakom debelog crijeva. Uključivanje podataka o IgG glikanima u regresijske modele nije dovelo do statistički značajnog poboljšanja u ukupnoj prognozi preživljavanja (Harrellov C indeks: 0.73, 0.77; AUC: 0.75, 0.79; IDI: 0.02, 0.04). Međutim, u odnosu na modele temeljene isključivo na kliničkim podacima uključivanjem podataka o IgG glikanima značajno je poboljšano predviđanje "rapidnih progresora" kod bolesnika u AJCC četvrtoj fazi (AUC 0.53 vs 0.75, IDI 0.21). Analizom kliničkih podataka 760 pacijenata i 538 odgovarajućih kontrola utvrđena je povezanost raka debelog crijeva s padom stupnja galaktozilacije i sijalinizacije IgG glikana, s porastom fukozilacije neutralnih IgG glikana te padom fukozilacije sijaliniziranih IgG glikana.

ZAKLJUČAK: Karakteristike glikozilacije IgG-a kod pacijenata oboljelih od raka debelog crijeva u skladu su sa značajno povećanom proupalnom aktivnošću IgG-a povezanom s lošijom prognozom, pogotovo u kasnijem stadiju. S obzirom na funkcionalnu važnost glikozilacije IgG-a u imunološkom nadzoru i učinkovitosti terapije monoklonskim antitijelima, individualne varijacije u glikozilaciji IgG-a mogle bi imati važnu ulogu u predviđanju tijeka bolesti ili izboru terapije.

9. ABSTRACT IN ENGLISH

IMMUNOGLOBULIN G GLYCOSYLATION IN PATIENTS WITH COLORECTAL CANCER

PhD candidate: Kujtim Thaçi
Year: 2017

INTRODUCTION: Colorectal cancer (CRC) is a malignant neoplasm of the colon and the rectum. CRC is still associated with poor prognosis, low survival rate and usually relatively late diagnosis.

MATERIALS AND METHODS: We analysed IgG glycome composition in 1229 patients with CRC and 538 matching controls. Effects of surgery were evaluated in 28 patients sampled before and three times after surgery. Furthermore, IgG glycome composition was analysed in 39 plasma samples collected before initial diagnosis of CRC.

RESULTS: Clinical algorithms showed good prediction of all cause and CRC mortality. The inclusion of IgG glycan data in regression models did not lead to any statistically significant improvements in overall prediction of survival (Harrell's C: 0.73, 0.77; AUC: 0.75, 0.79, IDI: 0.02, 0.04 respectively). However, the inclusion of IgG glycan data substantially improved the prediction of rapid progressors over clinical models in AJCC stage 4 patients (AUC 0.53 vs. 0.75, IDI 0.21). When analysing clinical characteristics among 760 patients and 538 matching controls it was found that CRC associates with decrease in IgG galactosylation, IgG sialylation and increase in core-fucosylation of neutral glycans with concurrent decrease of core fucosylation of sialylated glycans.

COCLUSION: The glycan differences among CRC patients are consistent with significantly increased IgG pro-inflammatory activity being associated with poorer CRC prognosis, especially in late stage CRC**.** Considering the functional relevance of IgG glycosylation for both tumor immunosurveilance and clinical efficacy of therapy with monoclonal antibodies, individual variation in IgG glycosylation may turn out to be important for prediction of disease course or the choice of therapy.

10. REFERENCES

1. Lebrilla CB, An HJ. The prospects of glycan biomarkers for the diagnosis of diseases. Mol Biosyst. 2009;5(1):17.

2. Werz DB, Ranzinger R, Herget S, Adibekian A, von der Lieth C-W, Seeberger PH. Exploring the structural diversity of mammalian carbohydrates ("glycospace") by statistical databank analysis. ACS Chem Biol. 2007 Oct 19;2(10):685–91.

3. Ohtsubo K, Marth JD. Glycosylation in Cellular Mechanisms of Health and Disease. Cell. 2006 Sep;126(5):855–67.

4. Dwek RA. Glycobiology: Toward Understanding the Function of Sugars. Chem Rev. 1996 Jan;96(2):683–720.

5. Cummings RD. The repertoire of glycan determinants in the human glycome. Mol Biosyst. 2009;5(10):1087.

6. Skropeta D. The effect of individual N-glycans on enzyme activity. Bioorg Med Chem. 2009 Apr;17(7):2645–53.

7. Lee RT, Lauc G, Lee YC. Glycoproteomics: protein modifications for versatile functions. EMBO Rep. 2005 May;6(11):1018–22.

8. Marth JD, Grewal PK. Mammalian glycosylation in immunity. Nat Rev Immunol. 2008 Nov;8(11):874–87.

9. Freeze HH. Genetic defects in the human glycome. Nat Rev Genet. 2006 Jul;7(7):537–51.

10. Abu-Qarn M, Eichler J, Sharon N. Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. Curr Opin Struct Biol. 2008 Oct;18(5):544–50.

11. Helenius A. Intracellular Functions of N-Linked Glycans. Science. 2001 Mar 23;291(5512):2364–9.

12. Zielinska DF, Gnad F, Wiśniewski JR, Mann M. Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. Cell. 2010 May;141(5):897–907.

13. Marek KW, Vijay IK, Marth JD. A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. Glycobiology. 1999 Nov;9(11):1263–71.

14. Lauc G, Rudan I, Campbell H, Rudd PM. Complex genetic regulation of protein glycosylation. Mol Biosyst. 2010;6(2):329.

15. Abbott KL, Nairn AV, Hall EM, Horton MB, McDonald JF, Moremen KW, et al. Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. PROTEOMICS. 2008 Aug;8(16):3210–20.

16. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol. 2012 Jun 22;13(7):448–62.

17. Knezevic A, Gornik O, Polasek O, Pucic M, Redzic I, Novokmet M, et al. Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. Glycobiology. 2010 Aug 1;20(8):959–69.

18. Lauc G, Zoldoš V. Protein glycosylation—an evolutionary crossroad between genes and environment. Mol Biosyst. 2010;6(12):2373.

19. Knežević A, Polašek O, Gornik O, Rudan I, Campbell H, Hayward C, et al. Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. J Proteome Res. 2009 Feb 6;8(2):694–701.

20. Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High Throughput Isolation and Glycosylation Analysis of IgG-Variability and Heritability of the IgG Glycome in Three Isolated Human Populations. Mol Cell Proteomics. 2011 Oct 1;10(10):M111.010090–M111.010090.

21. Huffman JE, Knezevic A, Vitart V, Kattla J, Adamczyk B, Novokmet M, et al. Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. Hum Mol Genet. 2011 Dec 15;20(24):5000–11.

22. Varki A. Biological roles of oligosaccharides: all of the theories are correct. Glycobiology. 1993 Apr;3(2):97–130.

23. Dhiman N, Haralambieva IH, Kennedy RB, Vierkant RA, O'Byrne MM, Ovsyannikova IG, et al. SNP/haplotype associations in cytokine and cytokine receptor genes and immunity to rubella vaccine. Immunogenetics. 2010 Apr;62(4):197–210.

24. National Research Council (U.S.), National Research Council (U.S.), National Research Council (U.S.). Transforming glycoscience: a roadmap for the future. Washington, D.C: National Academies Press; 2012. 191 p.

25. Dube DH, Bertozzi CR. Glycans in cancer and inflammation — potential for therapeutics and diagnostics. Nat Rev Drug Discov. 2005 Jun;4(6):477–88.

26. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer. 2005 Jul;5(7):526–42.

27. An HJ, Peavy TR, Hedrick JL, Lebrilla CB. Determination of N-glycosylation sites and site heterogeneity in glycoproteins. Anal Chem. 2003 Oct 15;75(20):5628–37.

28. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta. 1999 Dec 6;1473(1):4–8.

29. Dennis JW, Granovsky M, Warren CE. Glycoprotein glycosylation and cancer progression. Biochim Biophys Acta. 1999 Dec 6;1473(1):21–34.

30. Meany DL, Chan DW. Aberrant glycosylation associated with enzymes as cancer biomarkers. Clin Proteomics. 2011;8(1):7.

31. Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. Nat Rev Immunol. 2007 Apr;7(4):255–66.

32. Potapenko IO, Haakensen VD, Lüders T, Helland Å, Bukholm I, Sørlie T, et al. Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. Mol Oncol. 2010 Apr;4(2):98–118.

33. Reis CA, Osorio H, Silva L, Gomes C, David L. Alterations in glycosylation as biomarkers for cancer detection. J Clin Pathol. 2010 Apr 1;63(4):322–9.

34. Packer NH, von der Lieth C-W, Aoki-Kinoshita KF, Lebrilla CB, Paulson JC, Raman R, et al. Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11-13, 2006). Proteomics. 2008 Jan;8(1):8–20.

35. Vanderschaeghe D, Szekrényes Á, Wenz C, Gassmann M, Naik N, Bynum M, et al. High-Throughput Profiling of the Serum N-Glycome on Capillary Electrophoresis Microfluidics Systems: Toward Clinical Implementation of GlycoHepatoTest. Anal Chem. 2010 Sep;82(17):7408–15.

36. Wuhrer M. Glycosylation profiling in clinical proteomics: heading for glycan biomarkers. Expert Rev Proteomics. 2007 Apr;4(2):135–6.

37. Peracaula R, Tabarés G, Royle L, Harvey DJ, Dwek RA, Rudd PM, et al. Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. Glycobiology. 2003 Jun;13(6):457–70.

38. Sarrats A, Saldova R, Comet J, O'Donoghue N, de Llorens R, Rudd PM, et al. Glycan Characterization of PSA 2-DE Subforms from Serum and Seminal Plasma. OMICS J Integr Biol. 2010 Aug;14(4):465–74.

39. Arnold JN, Saldova R, Hamid UMA, Rudd PM. Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. PROTEOMICS. 2008 Aug;8(16):3284–93.

40. Pierce A, Saldova R, Abd Hamid UM, Abrahams JL, McDermott EW, Evoy D, et al. Levels of specific glycans significantly distinguish lymph node-positive from lymph node-negative breast cancer patients. Glycobiology. 2010 Oct 1;20(10):1283–8.

41. Saldova R, Wormald MR, Dwek RA, Rudd PM. Glycosylation changes on serum glycoproteins in ovarian cancer may contribute to disease pathogenesis. Dis Markers. 2008;25(4-5):219–32.

42. Taniguchi N, Miyoshi E, Ko JH, Ikeda Y, Ihara Y. Implication of N-acetylglucosaminyltransferases III and V in cancer: gene regulation and signaling mechanism. Biochim Biophys Acta. 1999 Oct 8;1455(2-3):287–300.

43. Murata K, Miyoshi E, Kameyama M, Ishikawa O, Kabuto T, Sasaki Y, et al. Expression of N-acetylglucosaminyltransferase V in colorectal cancer correlates with metastasis and poor prognosis. Clin Cancer Res Off J Am Assoc Cancer Res. 2000 May;6(5):1772–7.

44. Yamamoto E, Ino K, Miyoshi E, Shibata K, Takahashi N, Kajiyama H, et al. Expression of N-acetylglucosaminyltransferase V in endometrial cancer correlates with poor prognosis. Br J Cancer. 2007 Dec 3;97(11):1538–44.

45. Li D, Li Y, Wu X, Li Q, Yu J, Gen J, et al. Knockdown of Mgat5 inhibits breast cancer cell growth with activation of CD4+ T cells and macrophages. J Immunol Baltim Md 1950. 2008 Mar 1;180(5):3158–65.

46. Kim Y-S, Hwang SY, Kang H-Y, Sohn H, Oh S, Kim J-Y, et al. Functional proteomics study reveals that N-Acetylglucosaminyltransferase V reinforces the invasive/metastatic potential of colon cancer through aberrant glycosylation on tissue inhibitor of metalloproteinase-1. Mol Cell Proteomics MCP. 2008 Jan;7(1):1–14.

47. Barthel SR, Wiese GK, Cho J, Opperman MJ, Hays DL, Siddiqui J, et al. Alpha 1,3 fucosyltransferases are master regulators of prostate cancer cell trafficking. Proc Natl Acad Sci. 2009 Nov 17;106(46):19491–6.

48. Bos PD, Zhang XH-F, Nadal C, Shu W, Gomis RR, Nguyen DX, et al. Genes that mediate breast cancer metastasis to the brain. Nature. 2009 Jun 18;459(7249):1005–9.

49. Sellers TA, Huang Y, Cunningham J, Goode EL, Sutphen R, Vierkant RA, et al. Association of Single Nucleotide Polymorphisms in Glycosylation Genes with Risk of Epithelial Ovarian Cancer. Cancer Epidemiol Biomarkers Prev. 2008 Feb 4;17(2):397–404.

50. Dennis JW, Granovsky M, Warren CE. Protein glycosylation in development and disease. BioEssays News Rev Mol Cell Dev Biol. 1999 May;21(5):412–21.

51. Zhao Y, Sato Y, Isaji T, Fukuda T, Matsumoto A, Miyoshi E, et al. Branched N-glycans regulate the biological functions of integrins and cadherins: Biological functions of branched N-glycans. FEBS J. 2008 May;275(9):1939–48.

52. Miyoshi E, Ito Y, Miyoshi Y. Involvement of Aberrant Glycosylation in Thyroid Cancer. J Oncol. 2010;2010:1–7.

53. Miyoshi E, Moriwaki K, Nakagawa T. Biological Function of Fucosylation in Cancer Biology. J Biochem (Tokyo). 2007 Nov 16;143(6):725–9.

54. Miyoshi E, Nakano M. Fucosylated haptoglobin is a novel marker for pancreatic cancer: Detailed analyses of oligosaccharide structures. PROTEOMICS. 2008 Aug;8(16):3257–62.

55. Li D, Mallory T, Satomura S. AFP-L3: a new generation of tumor marker for hepatocellular carcinoma. Clin Chim Acta Int J Clin Chem. 2001 Nov;313(1-2):15–9.

56. Barrabés S, Pagès-Pons L, Radcliffe CM, Tabarés G, Fort E, Royle L, et al. Glycosylation of serum ribonuclease 1 indicates a major endothelial origin and reveals an increase in core fucosylation in pancreatic cancer. Glycobiology. 2007 Apr;17(4):388–400.

57. Thanabalasingham G, Huffman JE, Kattla JJ, Novokmet M, Rudan I, Gloyn AL, et al. Mutations in HNF1A Result in Marked Alterations of Plasma Glycan Profile. Diabetes. 2013 Apr 1;62(4):1329–37.

58. Orntoft TF, Vestergaard EM. Clinical aspects of altered glycosylation of glycoproteins in cancer. Electrophoresis. 1999 Feb;20(2):362–71.

59. Gornik O, Gornik I, Gašparović V, Lauc G. Change in transferrin sialylation is a potential prognostic marker for severity of acute pancreatitis. Clin Biochem. 2008 May;41(7-8):504–10.

60. Saldova R, Fan Y, Fitzpatrick JM, Watson RWG, Rudd PM. Core fucosylation and 2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. Glycobiology. 2011 Feb 1;21(2):195–205.

61. Vanhooren V, Liu X-E, Franceschi C, Gao C-F, Libert C, Contreras R, et al. N-glycan profiles as tools in diagnosis of hepatocellular carcinoma and prediction of healthy human ageing. Mech Ageing Dev. 2009 Jan;130(1-2):92–7.

62. Gerber-Lemaire S, Juillerat-Jeanneret L. Glycosylation pathways as drug targets for cancer: glycosidase inhibitors. Mini Rev Med Chem. 2006 Sep;6(9):1043–52.

63. Kim YJ, Varki A. Perspectives on the significance of altered glycosylation of glycoproteins in cancer. Glycoconj J. 1997 Aug;14(5):569–76.

64. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA. The impact of glycosylation on the biological function and structure of human immunoglobulins. Annu Rev Immunol. 2007;25:21–50.

65. Schur PH. IgG subclasses. A historical perspective. Monogr Allergy. 1988;23:1–11.

66. Schroeder HW, Cavacini L. Structure and function of immunoglobulins. J Allergy Clin Immunol. 2010 Feb;125(2):S41–52.

67. Nezlin R, Ghetie V. Interactions of immunoglobulins outside the antigen-combining site. Adv Immunol. 2004;82:155–215.

68. Masuda K, Yamaguchi Y, Kato K, Takahashi N, Shimada I, Arata Y. Pairing of oligosaccharides in the Fc region of immunoglobulin G. FEBS Lett. 2000 May 19;473(3):349–57.

69. Huber R, Deisenhofer J, Colman PM, Matsushima M, Palm W. Crystallographic structure studies of an IgG molecule and an Fc fragment. Nature. 1976 Dec 2;264(5585):415–20.

70. Wright A, Morrison SL. Effect of glycosylation on antibody function: implications for genetic engineering. Trends Biotechnol. 1997 Jan;15(1):26–32.

71. Bruhns P, Iannascoli B, England P, Mancardi DA, Fernandez N, Jorieux S, et al. Specificity and affinity of human Fc receptors and their polymorphic variants for human IgG subclasses. Blood. 2009 Apr 16;113(16):3716–25.

72. Stapleton NM, Andersen JT, Stemerding AM, Bjarnarson SP, Verheul RC, Gerritsen J, et al. Competition for FcRn-mediated transport gives rise to short half-life of human IgG3 and offers therapeutic potential. Nat Commun. 2011 Dec 20;2:599.

73. Lund J, Takahashi N, Pound JD, Goodall M, Jefferis R. Multiple interactions of IgG with its core oligosaccharide can modulate recognition by complement and human Fc gamma receptor I and influence the synthesis of its oligosaccharide chains. J Immunol Baltim Md 1950. 1996 Dec 1;157(11):4963–9.

74. Lux A, Nimmerjahn F. Impact of Differential Glycosylation on IgG Activity. In: Pulendran B, Katsikis PD, Schoenberger SP, editors. Crossroads between Innate and Adaptive Immunity III [Internet]. New York, NY: Springer New York; 2011 [cited 2015 Feb 9]. p. 113–24. Available from: http://link.springer.com/10.1007/978-1-4419-5632-3_10

75. Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. PROTEOMICS. 2008 Jul;8(14):2858–71.

76. Wuhrer M, Stam JC, van de Geijn FE, Koeleman CAM, Verrips CT, Dolhain RJEM, et al. Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. Proteomics. 2007 Nov;7(22):4070–81.

77. Gornik O, Lauc G. Glycosylation of serum proteins in inflammatory diseases. Dis Markers. 2008;25(4-5):267–78.

78. Mizuochi T, Taniguchi T, Shimizu A, Kobata A. Structural and numerical variations of the carbohydrate moiety of immunoglobulin G. J Immunol Baltim Md 1950. 1982 Nov;129(5):2016–20.

79. Shikata K, Yasuda T, Takeuchi F, Konishi T, Nakata M, Mizuochi T. Structural changes in the oligosaccharide moiety of human IgG with aging. Glycoconj J. 1998 Jul;15(7):683–9.

80. Raju TS. Terminal sugars of Fc glycans influence antibody effector functions of IgGs. Curr Opin Immunol. 2008 Aug;20(4):471–8.

81. Kaneko Y, Nimmerjahn F, Ravetch JV. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. Science. 2006 Aug 4;313(5787):670–3.

82. Pilar Rauter A, Křen V, Royal Society of Chemistry (Great Britain). Carbohydrate chemistry. chemical and biological approaches Volume 37 Volume 37 [Internet]. Cambridge: Royal Society of Chemistry; 2011 [cited 2015 Feb 9]. Available from: http://dx.doi.org/10.1039/9781849732765

83. Horvat T, Zoldoš V, Lauc G. Evolutional and clinical implications of the epigenetic regulation of protein glycosylation. Clin Epigenetics. 2011 Aug;2(2):425–32.

84. Kobata A. The N-linked sugar chains of human immunoglobulin G: their unique pattern, and their functional roles. Biochim Biophys Acta. 2008 Mar;1780(3):472–8.

85. Jefferis R. Glycosylation of recombinant antibody therapeutics. Biotechnol Prog. 2005 Feb;21(1):11–6.

86. Zhu D, Ottensmeier CH, Du M-Q, McCarthy H, Stevenson FK. Incidence of potential glycosylation sites in immunoglobulin variable regions distinguishes between subsets of Burkitt's lymphoma and mucosa-associated lymphoid tissue lymphoma. Br J Haematol. 2003 Jan;120(2):217–22.

87. Rademacher TW, Homans SW, Parekh RB, Dwek RA. Immunoglobulin G as a glycoprotein. Biochem Soc Symp. 1986;51:131–48.

88. Del Val IJ, Kontoravdi C, Nagy JM. Towards the implementation of quality by design to the production of therapeutic monoclonal antibodies with desired glycosylation patterns. Biotechnol Prog. 2010 Nov;26(6):1505–27.

89. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. Glycosylation and the immune system. Science. 2001 Mar 23;291(5512):2370–6.

90. Nimmerjahn F, Ravetch JV. Fcγ receptors as regulators of immune responses. Nat Rev Immunol. 2008 Jan;8(1):34–47.

91. Fanger MW, Graziano RF, Shen L, Guyre PM. Fc gamma R in cytotoxicity exerted by mononuclear cells. Chem Immunol. 1989;47:214–53.

92. Anderson CL, Shen L, Eicher DM, Wewers MD, Gill JK. Phagocytosis mediated by three distinct Fc gamma receptor classes on human leukocytes. J Exp Med. 1990 Apr 1;171(4):1333–45.

93. Anderson CL, Guyre PM, Whitin JC, Ryan DH, Looney RJ, Fanger MW. Monoclonal antibodies to Fc receptors for IgG on human mononuclear phagocytes. Antibody characterization and induction of superoxide production in a monocyte cell line. J Biol Chem. 1986 Sep 25;261(27):12856–64.

94. Anegón I, Cuturi MC, Trinchieri G, Perussia B. Interaction of Fc receptor (CD16) ligands induces transcription of interleukin 2 receptor (CD25) and lymphokine genes and expression of their products in human natural killer cells. J Exp Med. 1988 Feb 1;167(2):452–72.

95. Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, et al. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. J Biol Chem. 2002 Jul 26;277(30):26733–40.

96. Barbin K, Stieglmaier J, Saul D, Stieglmaier K, Stockmeyer B, Pfeiffer M, et al. Influence of variable N-glycosylation on the cytolytic potential of chimeric CD19 antibodies. J Immunother Hagerstown Md 1997. 2006 Apr;29(2):122–33.

97. Roopenian DC, Akilesh S. FcRn: the neonatal Fc receptor comes of age. Nat Rev Immunol. 2007 Sep;7(9):715–25.

98. Simister NE, Story CM, Chen HL, Hunt JS. An IgG-transporting Fc receptor expressed in the syncytiotrophoblast of human placenta. Eur J Immunol. 1996 Jul;26(7):1527–31.

99. Nimmerjahn F, Ravetch JV. Fcgamma receptors as regulators of immune responses. Nat Rev Immunol. 2008 Jan;8(1):34–47.

100. Jefferis R. Glycosylation as a strategy to improve antibody-based therapeutics. Nat Rev Drug Discov. 2009 Mar;8(3):226–34.

101. Karsten CM, Pandey MK, Figge J, Kilchenstein R, Taylor PR, Rosas M, et al. Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. Nat Med. 2012 Sep;18(9):1401–6.

102. Nimmerjahn F, Ravetch JV. Fcgamma receptors: old friends and new family members. Immunity. 2006 Jan;24(1):19–28.

103. Nose M, Wigzell H. Biological significance of carbohydrate chains on monoclonal antibodies. Proc Natl Acad Sci U S A. 1983 Nov;80(21):6632–6.

104. Tao MH, Morrison SL. Studies of aglycosylated chimeric mouse-human IgG. Role of carbohydrate in the structure and effector functions mediated by the human IgG constant region. J Immunol Baltim Md 1950. 1989 Oct 15;143(8):2595–601.

105. Feige MJ, Nath S, Catharino SR, Weinfurtner D, Steinbacher S, Buchner J. Structure of the Murine Unglycosylated IgG1 Fc Fragment. J Mol Biol. 2009 Aug;391(3):599–608.

106. Miyoshi E, Noda K, Yamaguchi Y, Inoue S, Ikeda Y, Wang W, et al. The alpha1-6-fucosyltransferase gene and its biological significance. Biochim Biophys Acta. 1999 Dec 6;1473(1):9–20.

107. Mimura Y, Ashton PR, Takahashi N, Harvey DJ, Jefferis R. Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. J Immunol Methods. 2007 Sep 30;326(1-2):116–26.

108. Scanlan CN, Burton DR, Dwek RA. Making autoantibodies safe. Proc Natl Acad Sci. 2008 Mar 18;105(11):4081–2.

109. Shibata-Koyama M, Iida S, Okazaki A, Mori K, Kitajima-Miyama K, Saitou S, et al. The N-linked oligosaccharide at Fc RIIIa Asn-45: an inhibitory element for high Fc RIIIa binding affinity to IgG glycoforms lacking core fucosylation. Glycobiology. 2008 Oct 24;19(2):126–34.

110. Iida S, Misaka H, Inoue M, Shibata M, Nakano R, Yamane-Ohnuki N, et al. Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcgammaRIIIa. Clin Cancer Res Off J Am Assoc Cancer Res. 2006 May 1;12(9):2879–87.

111. Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, et al. The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. J Biol Chem. 2003 Jan 31;278(5):3466–73.

112. Nimmerjahn F, Ravetch JV. Antibodies, Fc receptors and cancer. Curr Opin Immunol. 2007 Apr;19(2):239–45.

113. Mimura Y, Lund J, Church S, Dong S, Li J, Goodall M, et al. Butyrate increases production of human chimeric IgG in CHO-K1 cells whilst maintaining function and glycoform profile. J Immunol Methods. 2001 Jan 1;247(1-2):205–16.

114. Imai-Nishiya H, Mori K, Inoue M, Wakitani M, Iida S, Shitara K, et al. Double knockdown of alpha1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC. BMC Biotechnol. 2007;7(1):84.

115. Yamane-Ohnuki N, Kinoshita S, Inoue-Urakubo M, Kusunoki M, Iida S, Nakano R, et al. Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. Biotechnol Bioeng. 2004 Sep 5;87(5):614–22.

116. Patnaik SK, Stanley P. Lectin-resistant CHO glycosylation mutants. Methods Enzymol. 2006;416:159–82.

117. Davies J, Jiang L, Pan LZ, LaBarre MJ, Anderson D, Reff M. Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FC gamma RIII. Biotechnol Bioeng. 2001 Aug 20;74(4):288–94.

118. Mori K, Iida S, Yamane-Ohnuki N, Kanda Y, Kuni-Kamochi R, Nakano R, et al. Non-fucosylated therapeutic antibodies: the next generation of therapeutic antibodies. Cytotechnology. 2007 Dec 3;55(2-3):109–14.

119. Gornik O, Pavić T, Lauc G. Alternative glycosylation modulates function of IgG and other proteins — Implications on evolution and disease. Biochim Biophys Acta BBA - Gen Subj. 2012 Sep;1820(9):1318–26.

120. Scallon BJ, Tam SH, McCarthy SG, Cai AN, Raju TS. Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. Mol Immunol. 2007 Mar;44(7):1524–34.

121. Anthony RM, Kobayashi T, Wermeling F, Ravetch JV. Intravenous gammaglobulin suppresses inflammation through a novel TH2 pathway. Nature. 2011 Jun 19;475(7354):110–3.

122. Samuelsson A, Towers TL, Ravetch JV. Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. Science. 2001 Jan 19;291(5503):484–6.

123. Kleinau S. The impact of Fc receptors on the development of autoimmune diseases. Curr Pharm Des. 2003;9(23):1861–70.

124. Ravetch, J.V. 2003. Fc receptors. In Fundamental Immunology. W.E. Paul, ed-itor. Lippincott-Raven, Philadelphia, PA. 685-700. In.

125. Nimmerjahn F, Ravetch JV. Anti-Inflammatory Actions of Intravenous Immunoglobulin. Annu Rev Immunol. 2008 Apr;26(1):513–33.

126. Negi V-S, Elluru S, Sibéril S, Graff-Dubois S, Mouthon L, Kazatchkine MD, et al. Intravenous immunoglobulin: an update on the clinical use and mechanisms of action. J Clin Immunol. 2007 May;27(3):233–45.

127. Anthony RM, Wermeling F, Karlsson MCI, Ravetch JV. Identification of a receptor required for the anti-inflammatory activity of IVIG. Proc Natl Acad Sci. 2008 Dec 16;105(50):19571–8.

128. Debré M, Bonnet MC, Fridman WH, Carosella E, Philippe N, Reinert P, et al. Infusion of Fc gamma fragments for treatment of children with acute immune thrombocytopenic purpura. Lancet. 1993 Oct 16;342(8877):945–9.

129. Anthony RM, Nimmerjahn F, Ashline DJ, Reinhold VN, Paulson JC, Ravetch JV. Recapitulation of IVIG Anti-Inflammatory Activity with a Recombinant IgG Fc. Science. 2008 Apr 18;320(5874):373–6.

130. Kaneko Y, Nimmerjahn F, Madaio MP, Ravetch JV. Pathology and protection in nephrotoxic nephritis is determined by selective engagement of specific Fc receptors. J Exp Med. 2006 Mar 20;203(3):789–97.

131. Schwab I, Biburger M, Krönke G, Schett G, Nimmerjahn F. IVIg-mediated amelioration of ITP in mice is dependent on sialic acid and SIGNR1: HIGHLIGHTS. Eur J Immunol. 2012 Apr;42(4):826–30.

132. Bruhns P, Samuelsson A, Pollard JW, Ravetch JV. Colony-stimulating factor-1-dependent macrophages are responsible for IVIG protection in antibody-induced autoimmune disease. Immunity. 2003 Apr;18(4):573–81.

133. Hodoniczky J, Zheng YZ, James DC. Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. Biotechnol Prog. 2005 Dec;21(6):1644–52.

134. Huhn C, Selman MHJ, Ruhaak LR, Deelder AM, Wuhrer M. IgG glycosylation analysis. PROTEOMICS. 2009 Feb;9(4):882–913.

135. Selman MHJ, Niks EH, Titulaer MJ, Verschuuren JJGM, Wuhrer M, Deelder AM. IgG Fc N - Glycosylation Changes in Lambert-Eaton Myasthenic Syndrome and Myasthenia Gravis. J Proteome Res. 2011 Jan 7;10(1):143–52.

136. Ackerman ME, Crispin M, Yu X, Baruah K, Boesch AW, Harvey DJ, et al. Natural variation in Fc glycosylation of HIV-specific antibodies impacts antiviral activity. J Clin Invest. 2013 May 1;123(5):2183–92.

137. Dubé R, Rook GA, Steele J, Brealey R, Dwek R, Rademacher T, et al. Agalactosyl IgG in inflammatory bowel disease: correlation with C-reactive protein. Gut. 1990 Apr;31(4):431–4.

138. Malhotra R, Wormald MR, Rudd PM, Fischer PB, Dwek RA, Sim RB. Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. Nat Med. 1995 Mar;1(3):237–43.

139. Arnold JN, Dwek RA, Rudd PM, Sim RB. Mannan binding lectin and its interaction with immunoglobulins in health and in disease. Immunol Lett. 2006 Aug 15;106(2):103–10.

140. Nimmerjahn F, Anthony RM, Ravetch JV. Agalactosylated IgG antibodies depend on cellular Fc receptors for in vivo activity. Proc Natl Acad Sci U S A. 2007 May 15;104(20):8433–7.

141. Lauc G, Pezer M, Rudan I, Campbell H. Mechanisms of disease: The human N-glycome. Biochim Biophys Acta. 2015 Oct 21;

142. Holland M, Takada K, Okumoto T, Takahashi N, Kato K, Adu D, et al. Hypogalactosylation of serum IgG in patients with ANCA-associated systemic vasculitis. Clin Exp Immunol. 2002 Jul;129(1):183–90.

143. Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, Stanworth D, et al. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. Nature. 1985 Aug 1;316(6027):452–7.

144. Parekh RB, Roitt IM, Isenberg DA, Dwek RA, Ansell BM, Rademacher TW. Galactosylation of IgG associated oligosaccharides: reduction in patients with adult and juvenile onset rheumatoid arthritis and relation to disease activity. Lancet. 1988 Apr 30;1(8592):966–9.

145. Mehta AS, Long RE, Comunale MA, Wang M, Rodemich L, Krakover J, et al. Increased Levels of Galactose-Deficient Anti-Gal Immunoglobulin G in the Sera of Hepatitis C Virus-Infected Individuals with Fibrosis and Cirrhosis. J Virol. 2008 Feb 1;82(3):1259–70.

146. Moore JS, Wu X, Kulhavy R, Tomana M, Novak J, Moldoveanu Z, et al. Increased levels of galactose-deficient IgG in sera of HIV-1-infected individuals: AIDS. 2005 Mar;19(4):381–9.

147. Bones J, Mittermayr S, O'Donoghue N, Guttman A, Rudd PM. Ultra Performance Liquid Chromatographic Profiling of Serum *N* -Glycans for Fast and Efficient Identification of Cancer Associated Alterations in Glycosylation. Anal Chem. 2010 Dec 15;82(24):10208–15.

148. Bones J, Byrne JC, O'Donoghue N, McManus C, Scaife C, Boissin H, et al. Glycomic and Glycoproteomic Analysis of Serum from Patients with Stomach Cancer Reveals Potential Markers Arising from Host Defense Response Mechanisms. J Proteome Res. 2011 Mar 4;10(3):1246–65.

149. Parekh R, Roitt I, Isenberg D, Dwek R, Rademacher T. Age-related galactosylation of the N-linked oligosaccharides of human serum IgG. J Exp Med. 1988 May 1;167(5):1731–6.

150. Yamada E, Tsukamoto Y, Sasaki R, Yagyu K, Takahashi N. Structural changes of immunoglobulin G oligosaccharides with age in healthy human serum. Glycoconj J. 1997 Apr;14(3):401–5.

151. Ruhaak LR, Uh H-W, Beekman M, Koeleman CAM, Hokke CH, Westendorp RGJ, et al. Decreased Levels of Bisecting GlcNAc Glycoforms of IgG Are Associated with Human Longevity. Lucia A, editor. PLoS ONE. 2010 Sep 7;5(9):e12566.

152. Kri ti J, Vu kovi F., Menni C, Klari L, Keser T, Beceheli I, et al. Glycans Are a Novel Biomarker of Chronological and Biological Ages. J Gerontol A Biol Sci Med Sci. 2014 Jul 1;69(7):779–89.

153. Einarsdottir HK, Selman MHJ, Kapur R, Scherjon S, Koeleman CAM, Deelder AM, et al. Comparison of the Fc glycosylation of fetal and maternal immunoglobulin G. Glycoconj J. 2013 Feb;30(2):147–57.

154. Williams PJ, Arkwright PD, Rudd P, Scragg IG, Edge CJ, Wormald MR, et al. Short communication: selective placental transport of maternal IgG to the fetus. Placenta. 1995 Dec;16(8):749–56.

155. Parekh R, Isenberg D, Rook G, Roitt I, Dwek R, Rademacher T. A comparative analysis of disease-associated changes in the galactosylation of serum IgG. J Autoimmun. 1989 Apr;2(2):101–14.

156. Young A, Sumar N, Bodman K, Goyal S, Sinclair H, Roitt I, et al. Agalactosyl IgG: an aid to differential diagnosis in early synovitis. Arthritis Rheum. 1991 Nov;34(11):1425–9.

157. Soltys AJ, Hay FC, Bond A, Axford JS, Jones MG, Randen I, et al. The binding of synovial tissue-derived human monoclonal immunoglobulin M rheumatoid factor to immunoglobulin G preparations of differing galactose content. Scand J Immunol. 1994 Aug;40(2):135–43.

158. Rademacher TW, Parekh RB, Dwek RA, Isenberg D, Rook G, Axford JS, et al. The role of IgG glycoforms in the pathogenesis of rheumatoid arthritis. Springer Semin Immunopathol. 1988;10(2-3):231–49.

159. Rook GA, Steele J, Brealey R, Whyte A, Isenberg D, Sumar N, et al. Changes in IgG glycoform levels are associated with remission of arthritis during pregnancy. J Autoimmun. 1991 Oct;4(5):779–94.

160. Gerçel-Taylor C, Bazzett LB, Taylor DD. Presence of aberrant tumor-reactive immunoglobulins in the circulation of patients with ovarian cancer. Gynecol Oncol. 2001 Apr;81(1):71–6.

161. Alley WR, Vasseur JA, Goetz JA, Svoboda M, Mann BF, Matei DE, et al. N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients. J Proteome Res. 2012 Apr 6;11(4):2282–300.

162. Kodar K, Stadlmann J, Klaamas K, Sergeyev B, Kurtenkov O. Immunoglobulin G Fc N-glycan profiling in patients with gastric cancer by LC-ESI-MS: relation to tumor progression and survival. Glycoconj J. 2012 Jan;29(1):57–66.

163. Chen G, Wang Y, Qin X, Li H, Guo Y, Wang Y, et al. Change in IgG $_1$ Fc *N* -linked glycosylation in human lung cancer: Age- and sex-related diagnostic potential: Proteomics and 2DE. ELECTROPHORESIS. 2013 Aug;34(16):2407–16.

164. Kanoh Y, Mashiko T, Danbara M, Takayama Y, Ohtani S, Egawa S, et al. Changes in serum IgG oligosaccharide chains with prostate cancer progression. Anticancer Res. 2004 Oct;24(5B):3135–9.

165. Ruhaak LR, Barkauskas DA, Torres J, Cooke CL, Wu LD, Stroble C, et al. The serum immunoglobulin G glycosylation signature of gastric cancer. EuPA Open Proteomics. 2015 Mar;6:1–9.

166. Chen G, Wang Y, Qiu L, Qin X, Liu H, Wang X, et al. Human IgG Fc-glycosylation profiling reveals associations with age, sex, female sex hormones and thyroid cancer. J Proteomics. 2012 Jun 6;75(10):2824–34.

167. Aurer I, Lauc G, Dumić J, Rendić D, Matisić D, Milos M, et al. Aberrant glycosylation of Igg heavy chain in multiple myeloma. Coll Antropol. 2007 Mar;31(1):247–51.

168. Kim H-J, Yu M-H, Kim H, Byun J, Lee C. Noninvasive molecular biomarkers for the detection of colorectal cancer. BMB Rep. 2008 Oct 31;41(10):685–92.

169. Yabroff KR, Lund J, Kepka D, Mariotto A. Economic burden of cancer in the United States: estimates, projections, and future research. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2011 Oct;20(10):2006–14.

170. Benson AB. Epidemiology, disease progression, and economic burden of colorectal cancer. J Manag Care Pharm JMCP. 2007 Aug;13(6 Suppl C):S5–18.

171. Hawk ET, Limburg PJ, Viner JL. Epidemiology and prevention of colorectal cancer. Surg Clin North Am. 2002 Oct;82(5):905–41.

172. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, et al. Early detection: The case for early detection. Nat Rev Cancer. 2003 Apr;3(4):243–52.

173. Koessler T, Azzato EM, Perkins B, Macinnis RJ, Greenberg D, Easton DF, et al. Common germline variation in mismatch repair genes and survival after a diagnosis of colorectal cancer. Int J Cancer J Int Cancer. 2009 Apr 15;124(8):1887–91.

174. Gockel I, Sgourakis G, Lyros O, Polotzek U, Schimanski CC, Lang H, et al. Risk of lymph node metastasis in submucosal esophageal cancer: a review of surgically resected patients. Expert Rev Gastroenterol Hepatol. 2011 Jun;5(3):371–84.

175. National Institute for Health and Clinical Excellence. Colorectal cancer. The diagnosis and management of colorectal cancer. NICE clinical guideline 131, 2011.

176. Garcea G, Sharma RA, Dennison A, Steward WP, Gescher A, Berry DP. Molecular biomarkers of colorectal carcinogenesis and their role in surveillance and early intervention. Eur J Cancer Oxf Engl 1990. 2003 May;39(8):1041–52.

177. Walt D, Aoki-Kinoshita KF, Bendiak B, Bertozzi CR, Boons GJ, Darvill A, et al. Transforming Glycoscience: A Roadmap for the Future. Washington: Nacional Academies Press; 2012.

178. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural Innate and Adaptive Immunity to Cancer. Annu Rev Immunol. 2011 Apr 23;29(1):235–71.

179. Burnet M. Cancer; a biological approach. I. The processes of control. Br Med J. 1957 Apr 6;1(5022):779–86.

180. Wolfert MA, Boons G-J. Adaptive immune activation: glycosylation does matter. Nat Chem Biol. 2013 Nov 14;9(12):776–84.

181. Weiner LM, Surana R, Wang S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. Nat Rev Immunol. 2010 May;10(5):317–27.

182. Mariño K, Bones J, Kattla JJ, Rudd PM. A systematic approach to protein glycosylation analysis: a path through the maze. Nat Chem Biol. 2010 Oct;6(10):713–23.

183. Theodoratou E, Kyle J, Cetnarskyj R, Farrington SM, Tenesa A, Barnetson R, et al. Dietary Flavonoids and the Risk of Colorectal Cancer. Cancer Epidemiol Biomarkers Prev. 2007 Apr 1;16(4):684–93.

184. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol. 2010 Apr 1;39(2):504–18.

185. Baković MP, Selman MHJ, Hoffmann M, Rudan I, Campbell H, Deelder AM, et al. High-Throughput IgG Fc N-Glycosylation Profiling by Mass Spectrometry of Glycopeptides. J Proteome Res. 2013 Feb;12(2):821–31.

186. Friedman J. Greedy function approximation: A gradient boosting machine. The Annals of Statistics 2001;29(5):1189-232. In.

187. Ridgeway G. The State of Boosting. Computing Science and Statistics 1999;31:172-81. In.

188. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B. 1995:289-300.

189. Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, Mužinić A, et al. Loci Associated with N-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. Gibson G, editor. PLoS Genet. 2013 Jan 31;9(1):e1003225.

190. Campbell MP, Royle L, Radcliffe CM, Dwek RA, Rudd PM. GlycoBase and autoGU: tools for HPLC-based glycan analysis. Bioinformatics. 2008 May 1;24(9):1214–6.

191. Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. Bioinforma Oxf Engl. 2005 Jul 1;21(13):3001–8.

192. Niu N, Zhang J, Huang T, Sun Y, Chen Z, Yi W, et al. IgG Expression in Human Colorectal Cancer and Its Relationship to Cancer Cell Behaviors. Hold GL, editor. PLoS ONE. 2012 Nov 1;7(11):e47362.

193. Coussens LM, Zitvogel L, Palucka AK. Neutralizing Tumor-Promoting Chronic Inflammation: A Magic Bullet? Science. 2013 Jan 18;339(6117):286–91.

194. Qian Y, Wang Y, Zhang X, Zhou L, Zhang Z, Xu J, et al. Quantitative Analysis of Serum IgG Galactosylation Assists Differential Diagnosis of Ovarian Cancer. J Proteome Res. 2013 Sep 6;12(9):4046–55.

195. Saldova R, Royle L, Radcliffe CM, Abd Hamid UM, Evans R, Arnold JN, et al. Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. Glycobiology. 2007 Dec;17(12):1344–56.

196. Trbojević Akmačić I, Ventham NT, Theodoratou E, Vučković F, Kennedy NA, Krištić J, et al. Inflammatory Bowel Disease Associates with Proinflammatory Potential of the Immunoglobulin G Glycome: Inflamm Bowel Dis. 2015 Apr;1.

197. Rombouts Y, Ewing E, van de Stadt LA, Selman MHJ, Trouw LA, Deelder AM, et al. Anti-citrullinated protein antibodies acquire a pro-inflammatory Fc glycosylation phenotype prior to the onset of rheumatoid arthritis. Ann Rheum Dis. 2015 Jan 1;74(1):234–41.

198. Vučković F, Krištić J, Gudelj I, Teruel M, Keser T, Pezer M, et al. Association of Systemic Lupus Erythematosus With Decreased Immunosuppressive Potential of the IgG Glycome: ASSOCIATION OF SLE WITH ALTERED IgG GLYCOSYLATION. Arthritis Rheumatol. 2015 Nov;67(11):2978–89.

199. Wang J, Balog CIA, Stavenhagen K, Koeleman CAM, Scherer HU, Selman MHJ, et al. Fc-Glycosylation of IgG1 is Modulated by B-cell Stimuli. Mol Cell Proteomics. 2011 May 1;10(5):M110.004655–M110.004655.

200. Axford JS, Sumar N, Alavi A, Isenberg DA, Young A, Bodman KB, et al. Changes in normal glycosylation mechanisms in autoimmune rheumatic disease. J Clin Invest. 1992 Mar;89(3):1021–31.

201. Omtvedt LA, Royle L, Husby G, Sletten K, Radcliffe CM, Harvey DJ, et al. Glycan analysis of monoclonal antibodies secreted in deposition disorders indicates that subsets of plasma cells differentially process IgG glycans. Arthritis Rheum. 2006 Nov;54(11):3433–40.

202. Axford JS. Decreased B-cell galactosyltransferase activity in rheumatoid arthritis. Br J Rheumatol. 1988;27 Suppl 2:170.

203. Alavi A, Axford JS, Pool AJ. Serum galactosyltransferase isoform changes in rheumatoid arthritis. J Rheumatol. 2004 Aug;31(8):1513–20.

204. Croce A, Firuzi O, Altieri F, Eufemi M, Agostino R, Priori R, et al. Effect of infliximab on the glycosylation of IgG of patients with rheumatoid arthritis. J Clin Lab Anal. 2007;21(5):303–14.

205. Pasek M, Duk M, Podbielska M, Sokolik R, Szechiński J, Lisowska E, et al. Galactosylation of IgG from rheumatoid arthritis (RA) patients--changes during therapy. Glycoconj J. 2006 Nov;23(7-8):463–71.

206. Bertazza L, Mocellin S. The dual role of tumor necrosis factor (TNF) in cancer biology. Curr Med Chem. 2010;17(29):3337–52.

207. Grimm M, Lazariotou M, Kircher S, Höfelmayr A, Germer CT, von Rahden BHA, et al. Tumor necrosis factor-α is associated with positive lymph node status in patients with recurrence of colorectal cancer—indications for anti-TNF-α agents in cancer treatment. Cell Oncol. 2011 Aug;34(4):315–26.

208. Balkwill F. Tumor necrosis factor or tumor promoting factor? Cytokine Growth Factor Rev. 2002 Apr;13(2):135–41.

209. Nasirikenari M, Chandrasekaran EV, Matta KL, Segal BH, Bogner PN, Lugade AA, et al. Altered eosinophil profile in mice with ST6Gal-1 deficiency: an additional role for

ST6Gal-1 generated by the P1 promoter in regulating allergic inflammation. J Leukoc Biol. 2010 Mar 1;87(3):457–66.

210. Menni C, Keser T, Mangino M, Bell JT, Erte I, Akmačić I, et al. Glycosylation of Immunoglobulin G: Role of Genetic and Epigenetic Influences. Kaufmann GF, editor. PLoS ONE. 2013 Dec 6;8(12):e82558.

211. Espy C, Morelle W, Kavian N, Grange P, Goulvestre C, Viallon V, et al. Sialylation levels of anti-proteinase 3 antibodies are associated with the activity of granulomatosis with polyangiitis (Wegener's). Arthritis Rheum. 2011 Jul;63(7):2105–15.

212. Scherer HU, van der Woude D, Ioan-Facsinay A, el Bannoudi H, Trouw LA, Wang J, et al. Glycan profiling of anti-citrullinated protein antibodies isolated from human serum and synovial fluid. Arthritis Rheum. 2010 Feb 22;62(6):1620–9.

213. Fleming SC, Smith S, Knowles D, Skillen A, Self CH. Increased sialylation of oligosaccharides on IgG paraproteins--a potential new tumour marker in multiple myeloma. J Clin Pathol. 1998 Nov;51(11):825–30.

214. KrištićJ, Zoldoš V, Lauc G. Complex Genetics of Protein N-Glycosylation. In: Endo T, Seeberger PH, Hart GW, Wong C-H, Taniguchi N, editors. Glycoscience: Biology and Medicine: Springer Japan; 2014. p. 1-7.

215. Novokmet M, Lukić E, Vučković F, –Durić Ž, Keser T, Rajšl K, et al. Changes in IgG and total plasma protein glycomes in acute systemic inflammation. Sci Rep [Internet]. 2014 Mar 11 [cited 2016 Jan 27];4. Available from: http://www.nature.com/articles/srep04347

216. Barb AW, Prestegard JH. NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. Nat Chem Biol. 2011 Mar;7(3):147–53.

217. Zhang W, Gordon M, Schultheis AM, Yang DY, Nagashima F, Azuma M, et al. FCGR2A and FCGR3A polymorphisms associated with clinical outcome of epidermal growth factor receptor expressing metastatic colorectal cancer patients treated with single-agent cetuximab. J Clin Oncol Off J Am Soc Clin Oncol. 2007 Aug 20;25(24):3712–8.

218. Cartron G, Dacheux L, Salles G, Solal-Celigny P, Bardos P, Colombat P, et al. Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcgammaRIIIa gene. Blood. 2002 Feb 1;99(3):754–8.

219. Chen C-Y, Jan Y-H, Juan Y-H, Yang C-J, Huang M-S, Yu C-J, et al. Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. Proc Natl Acad Sci. 2013 Jan 8;110(2):630–5.

220. Wang X, Chen J, Li QK, Peskoe SB, Zhang B, Choi C, et al. Overexpression of (1,6) fucosyltransferase associated with aggressive prostate cancer. Glycobiology. 2014 Oct 1;24(10):935–44.

221. Gerdes CA, Nicolini VG, Herter S, van Puijenbroek E, Lang S, Roemmele M, et al. GA201 (RG7160): A Novel, Humanized, Glycoengineered Anti-EGFR Antibody with Enhanced ADCC and Superior In Vivo Efficacy Compared with Cetuximab. Clin Cancer Res. 2013 Mar 1;19(5):1126–38.

222. Gasdaska JR, Sherwood S, Regan JT, Dickey LF. An afucosylated anti-CD20 monoclonal antibody with greater antibody-dependent cellular cytotoxicity and B-cell depletion and lower complement-dependent cytotoxicity than rituximab. Mol Immunol. 2012 Mar;50(3):134–41.

223. Oppenheim DE, Spreafico R, Etuk A, Malone D, Amofah E, Peña-Murillo C, et al. Glyco-engineered anti-EGFR mAb elicits ADCC by NK cells from colorectal cancer patients irrespective of chemotherapy. Br J Cancer. 2014 Mar 4;110(5):1221–7.

224. Kapur R, Kustiawan I, Vestrheim A, Koeleman CAM, Visser R, Einarsdottir HK, et al. A prominent lack of IgG1-Fc fucosylation of platelet alloantibodies in pregnancy. Blood. 2014 Jan 23;123(4):471–80.

225. Biburger M, Lux A, Nimmerjahn F. How Immunoglobulin G Antibodies Kill Target Cells. In: Advances in Immunology [Internet]. Elsevier; 2014 [cited 2016 Jan 28]. p. 67–94. Available from: http://linkinghub.elsevier.com/retrieve/pii/B9780128001479000030

226. Vuckovic F, Theodoratou E, Thaci K, Timofeeva M, Vojta A, tambuk J, et al. IgG glycome in colorectal cancer. Clin Cancer Res [Internet]. 2016 Feb 1 [cited 2016 Feb 4];

Available from: http://clincancerres.aacrjournals.org/cgi/doi/10.1158/1078-0432.CCR-15-1867

227. Wang M, Wang J, Kong X, Chen H, Wang Y, Qin M, et al. MiR-198 represses tumor growth and metastasis in colorectal cancer by targeting fucosyl transferase 8. Sci Rep. 2014 Sep 1;4:6145.

228. Yang L, Luo Y, Wei J. Integrative genomic analyses on Ikaros and its expression related to solid cancer prognosis. Oncol Rep. 2010 Aug;24(2):571–7.

229. Van Cutsem E, Nordlinger B, Cervantes A, ESMO Guidelines Working Group. Advanced colorectal cancer: ESMO Clinical Practice Guidelines for treatment. Ann Oncol Off J Eur Soc Med Oncol ESMO. 2010 May;21 Suppl 5:v93–7.

11. LIST OF FIGURES AND TABLES

Figures

Figure 1. The structure of an IgG1 molecule. (A) Fab and Fc portion (B).Complex biantennary glycan structure attached to Asn 297 in Fc portion of IgG molecule. Structural schemes are given in terms of blue square (N-acetylglucosamine), red triangle (fucose), green circle (mannose), yellow circle (galactose), and purple diamond (N-acetylneuraminic acid)

Figure 2. Structural variations in IgG glycans. Initial GlcNAc2Man3GlcNAc2 structure (red square) can be modified by the addition of bisecting GlcNAc (GnTIII), fucose (FUT8) or galactose (GalT). These resulting structures can further be modified by the activity of the same enzymes or by the addition of the sialic acid (SiaT)

Figure 3. Modulation of IgG function by alternative glycosylation.Structure of the glycan on IgG Fc part can significantly affect effector function of IgG

Figure 4. The scematic representation of the human IgG structure and functional implication of alternative glycosylation

Figure 5. UPLC analysis of immunoglobulin G (IgG) glycosylation. Each IgG contains one conserved N-glycosylation site on Asn197 of its heavy chain. Different glycans can be attached to this site and the process seems to be highly regulated. UPLC analysis can reveal composition of the glycome attached to a population of IgG molecules by separating total IgG N-glycome into 24 chromatographic glycan peaks (GP1–GP24), mostly corresponding to individual glycan structures

Figure 6. Minus logarithm of the FDR corrected p-values (q values) of all 39 glycan variables for all causes and CRC-specific mortality (Model III). Q- value threshold of significance <0.05

Figure 7. **I**gG glycome composition in CRC patients and matching controls

Figure 8. ROC curve illustrating the performance of regularized logistic regression model in predicting disease status for CRC patients and healthy controls. While models based only on age and gender did not show predictive power (red line), addition of glycan traits increased predictive power of model (black line).

TABLES

Table 1. Glycans annotation and experimental variation for each glycan variable

Table 2. Summary statistics and univariate Cox regression for factors influencing all-cause and CRC mortality

Table 3. All cause analysis for rank transformed glycans. Q value represents the adjusted p-values using the false discovery rate method (Benjamini–Hochberg procedure).

Table 4. All cause analysis

Table 5. CRC-specific analysis for rank transformed glycans. Q value represents the adjusted p-values using the false discovery rate method (Benjamini–Hochberg procedure).

Table 6. CRC-specific analysis

Table 7. All-cause mortality analysis for stages 1-3

Table 8. CRC-specific mortality analysis for stages 1-3

Table 9. All-cause mortality analysis by stage [Model II; rank transformed variables

Table 10. CRC-specific mortality analysis by stage [Model II; rank transformed variables

Table 11. Multivariate Cox regression of the a) clinical parameters and b) clinical and glycan parameters

Table 12. Multivariate Cox regression and estimate of the Harrell's concordance coefficient of the a) clinical parameters and b) clinical and glycan parameters by AJCC stage for all-cause mortality.

Table 13. Multivariate Cox regression and estimate of the Harrell's concordance coefficient of the a) clinical parameters and b) clinical and glycan parameters by AJCC stage for CRC mortality

Table 14. Predictions of 5 year risk of CRC death for models with clinical factors and clinical and glycan factors using k-nearest neighbours, LASSO, Naïve Bayes, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Table 15. Predictions of 5 year risk of CRC death for models with the extended set of clinical factors with and without glycans using k-nearest neighbours, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Table 16. Predictions of rapid progressors in stage 2 for models with extending set of clinical factors with and without glycan using k-nearest neighbours, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Table 17. Predictions of rapid progressors in stage 3 for models with the extended set of clinical factors with and without glycans using k-nearest neighbour, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Table 18. Predictions of rapid progressors in stage 4 for models with the extended set of clinical factors with and without glycans using k-nearest neighbour, LASSO, PAM, Support Vector Machines, Decision Trees, and Boosted Stump classifiers. The results are summarized over 10 cross-validation folds.

Table 19. Descriptive information on CRC patient and healthy controls

Table 20. Derived glycan traits

Table 21. IgG glycome composition in CRC patients and controls. Only the main derived traits describing glycome composition are shown.

Table 22. IgG glycome composition in CRC patients and controls

Table 23. Studies on IgG glycosylation changes in cancer

12. CURRICULUM VITAE

I was born on March 04, 1984, in Podujevo, Kosovo, where I completed primary school and gymnasium. I studied pharmacy at the University of Prishtina School of Medicine from 2002-2008. I joined the Clinical Biochemistry residency programme at Institute of Clinical Biochemistry in University Clinical Center of Kosovo (UCCK) from 2011-2015. I enrolled in PhD programme of Biomedicine and Health Sciences at School of Medicine, University of Zagreb in 2012. In the same year I spent one year in Genos Glycoscience Research Laboratory doing a basic research on a project which resulted in this PhD thesis.

Publications: 6 papers in peer reviewed journals.

Professional affiliation: College of Medical Sciences "REZONANCA" Prishtina.
Teaching assistant in Clinical Biohemistry (2016-continue).

Languages: Albanian, Croatian, English,

Personal data: married, father of two sons.