Aus dem Institut für Laboratoriumsmedizin, Klinische Chemie und Pathobiochemie, Campus Virchow Klinikum

der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Determination of *N*-Glycome signatures for Early Diagnosis of Epithelial Ovarian Cancer and Monitoring Response to Chemotherapy in an African Cohort

(Analyse von N-Glykom-Strukturen für die Frühdiagnose des epithelialen Ovarialkarzinoms und für das Ansprechen auf Chemotherapien in einer afrikanischen Kohorte)

zur Erlangung des akademischen Grades

Doctor rerum medicinalium (Dr. rer. medic)

vorgelegt der medizinischen Fakultät

Charité - Universitätsmedizin Berlin

Von

FRANCIS MUGENI WANYAMA

Aus Sio Port

(Kenia)

Datum der Promotion: 23.03.2024

Preface

Glycomics, a sub-discipline of 'omics' studies, has attracted an exponential growth of research activities in the last two decades, and especially in the area of biomarker discovery, owing to the development of new broad-spectrum high-throughput technologies that exhibit high sensitivities and specificities. This has led to a plethora of discoveries, and especially of biomarkers for various diseases and malignancies including ovarian cancer (OC), particularly in Caucasian and Asian ethnicities.

In my published review article that is a subject of this dissertation, *"Glycomic-Based Biomarkers for Ovarian Cancer: Advances and Challenges Diagnostics [Basel], 2021, 11(4)"*, I exhaustively documented the milestones that have been achieved so far in glycomics regarding discovery of OC biomarkers. Ovarian cancer is a lethal gynecological malignancy and a common cause of death among women. This is primarily because its diagnosis often happens late due to unavailable early diagnostic biomarkers. Other factors that have been attributed to reduced life expectancy of OC patients include resistance to chemotherapy agents as well as undiagnosed underlying viral infections such as Hepatitis B (HBV) and Human Immunodeficiency Virus (HIV) in patients scheduled for chemotherapy. Identification of HBV and HIV seropostive OC patients allows the clinicians to carefully select the optimal synergistic treatment strategies that curtail the viral growth as well as eliminating the tumor cells. In a second publication - also a subject of this dissertation, *"The Burden of Hepatitis B, Hepatitis C, and Human Immunodeficiency Viruses in Ovarian Cancer patients in Nairobi, Kenya Infect Dis Rep 2022, 14(3):433-445"*, high seroprevalence rates of HBV and HIV in Kenyan OC patients were reported. This indicated the need to incorporate screening of HBV and HIV in Kenyan ovarian OC patients to facilitate comprehensive management.

However, the substratum of this dissertation involved finding glycome-based biomarkers for the first time for early diagnosis of OC using a cohort of African ethnicity, as well as to find the differences or similarities in the expression of IgG_1 and IgG_2 among African and Caucasian ethnicities. The current findings therefore provide the information to the scientific community that glycome-based biomarkers for OC can be applied in a multi-ethnic setting. Additionally, the current findings provide a foundation on which future glycomic studies involving African and Caucasian ethnicities can be carried out.

<div align="right">

Francis Mugeni Wanyama

Doctoral Candidate (Mollecular Medicine)

Charité Universitätsmedizin Berlin, Germany

Also a Clinical Chemist at the University of Nairobi, Kenya

</div>

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Asn	asparagines
AUC	area under the curve
BOD	benign ovarian disease
BRCA1/BRCA2	breast cancer genes 1 and 2
CA 72-4	cancer antigen 72-4
CA125	cancer antigen 125
CDG	congenital disorder of glycosylation
DTE	dithioerythritol
Dol-P	dolichol phosphate
DMSO	dimethylsulfoxide
sDHB	super 2,5-dihydroxybenzoic acid
EOC	epithelial ovarian cancer
ER	endoplasmic reticulum
Fab	antigen binding fragment
Fuc	fucose
Fc	fragment crystallizable
FIGO	International Federation of Gynaecology and Obstetrics
Gal	galactose
GalNAc	*N*-Acetylgalactosamine
Glc	glucose
GlcNAc	*N*-acetylglucosamine
GlcNAc-T	*N*-acetylglucosaminyl-transferase
GLYCOV	proposed glycan serum marker for ovarian cancer (developed by Prof. Blanchard group)
HE4	human epididymis 4
HILIC	hydrophilic interaction liquid chromatography
HCC	Hepatocellular carcinoma
IAA	Iodoacetamide
IATA	International air transport association
IgG	immunoglobulin gamma
kDa	kilo dalton
KU/L	kilo units per litre
m/z	mass to charge ratio
Man	mannose

MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
MGAT	gene encoding the GlcNAc-T
ml	milliliter
MS	mass spectrometry
MPI-CDG	mannose phosphate isomerase-CDG
OC	ovarian cancer
OST	oligosaccharyltransferase
PNGase	peptide-*N*4-(*N*-acetyl-beta-glucosaminyl) asparagine amidase
PMM2-CDG	phosphomannomutase 2-CDG
ROC	receiver operating curve
SELDI-TOF	surface-enhanced laser desorption/ionization time-of-flight
Ser	serine
sLe	sialyl lewis epitope
SPSS	Statistical Package for the Social Sciences
Thr	threonine

Abstrakt

Die Glykosylierung ist der häufigste, bei Proteinen auftretende, posttranslationale Modifikationsprozess, der zu deren struktureller und funktioneller Veränderung führt. Die Glykanbiosynthese ist bei Krankheiten und bösartigen Tumoren, einschließlich des epithelialem Ovarialkarzinoms (EOC), der dritthäufigsten Todesursache aller gynäkologischen Malignome in Kenia, verändert. Eine späte Diagnose, aufgrund fehlender frühzeitiger diagnostischer Biomarker, Chemotherapieresistenz und nicht diagnostizierte virale Erkrankungen sind die Hauptgründe für die hohe Sterblichkeitsrate. Sowohl als Routinetest für den Nachweis von Krebserkrankungen als auch für die Überwachung der Wirksamkeit von Krebsmedikamenten dient das Krebsantigen (CA125). Dessen Bedeutung wird jedoch durch seine geringe Spezifität und Sensitivität beeinträchtigt.

Ziel der Arbeit war es, Lösungen für die Ursachen der hohen Sterblichkeitsrate bei EOC Patientinnen zu finden, zu denen neben den zugrundeliegenden Virusinfektionen auch der Mangel an wirksamen Biomarkern für eine frühe Diagnose und die Überwachung des Ansprechens der Patientinnen auf die Chemotherapie gehört. Zu diesem Zweck wurden erstmals Serumproben einer afrikanischen Kohorte analysiert und N-Glykan-Signaturen identifiziert, die im Vergleich zu CA125 eine höhere Genauigkeit bei der EOC-Diagnose und der Überwachung des Ansprechens auf die Chemotherapie aufweisen. Darüber hinaus wurden Kohorten afrikanischer und europäischer Ethnien analysiert, um gemeinsame IgG1/IgG2-Glykopeptid-Serumsignaturen zur Verbesserung der EOC-Diagnose zu identifizieren. Abschließend wurden die Seroprävalenzen des Humanen Immundefizienz-Virus (HIV), der Hepatitis-B- und -C-Viren (HBV und HCV) bei kenianischen EOC-Patienten ermittelt. Die Erzeugung von Glykopeptiden und die Freisetzung der N-Glykane erfolgte enzymatisch. Danach wurden diese gereinigt und mittels MALDI-TOF-MS gemessen.

Fünfundzwanzig N-Glykan-Signaturen korrelierten mit einem Ansprechen auf die Chemotherapie. Darunter waren 17 Strukturen mit m/z-Werten <3600, die auf ein frühes Ansprechen auf die Chemotherapie hinwiesen, und weitere acht Strukturen mit m/z-Werten >3600, die auf ein spätes Ansprechen hinwiesen. Die Analyse der IgG1/IgG2-Glykopeptide ergab spezifische altersabhängige, ethnienübergreifende Glykopeptid-Signaturen für EOC-Patientinnen. Für Patientinnen im Alter von 18-48 Jahren wurden drei IgG1- und fünf IgG2-Glykopeptide identifiziert, wohingegen für Patientinnen im Alter von ≥ 49 Jahren sieben IgG1- und vier IgG2-Glykopeptide festgestellt wurden. Bei den OC-Patientinnen wurden Seroprävalenzen von 29,1 % (HBV), 26,7 % (HIV) und 1,2 % (HCV) bestimmt. Die N-Glykan- und IgG1/IgG2-Glykopeptid-Signaturen zeigten eine frühzeitigere Feststellung des EOCs und eine bessere Überwachung des Ansprechens auf die Chemotherapie als CA125 allein. Die hohe HBV- und HIV-Positivität bei kenianischen OC-Patientinnen machte eine Routineuntersuchung für eine optimale Behandlung erforderlich.

Abstract

Glycosylation is the most frequent post-translational modification (PTM) process that occurs in proteins resulting in their structural and functional modifications. Glycan biosynthesis is altered in diseases and malignancies, leading to features that have shown promise in the search for new biomarkers of various malignancies including ovarian cancer (OC). OC is a deadly gynecologic cancer, and in Kenya it is the third most frequent cause of death of all gynecological malignancies. The high mortalities are attributed to the late-stage diagnosis due to a lack of early diagnostic biomarkers, resistance to chemotherapy and underlying undiagnosed viral diseases. Cancer antigen (CA125) is the routine test for OC detection as well as for monitoring the efficacy of anti-cancer agents; however, its significance is undermined by its low specificity and sensitivity.

This thesis sought to find solutions for the triggers of high mortalities in Epithelial OC (EOC) patients, which include a lack of effective biomarkers for timely diagnosis and monitoring patients' response to chemotherapy, in addition to the underlying viral infections. To this end, serum samples of an African cohort were analyzed for the first time to identify N-glycan signatures with superior accuracy for EOC diagnosis and for monitoring chemotherapy response compared to CA125. Additionally, cohorts of African and European ethnicities were analyzed to identify common serum IgG_1/IgG_2 glycopeptide signatures to improve EOC diagnosis. Finally, seroprevalence rates of Human immunodeficiency virus (HIV), Hepatitis B and C viruses (HBV and HCV) among Kenyan EOC patients were determined. The total N-Glycans were enzymatically released, purified, permethylated, and measured by MALDI-TOF-MS. Similarly, the tryptic serum IgG_1 and IgG_2 glycopeptides were measured by MALDI-TOF-MS. Finally, measurements of serum HBV, HCV, and HIV status were done on the Cobas e 801 immunoassay system.

Twenty-five N-glycan signatures correlated with the response to chemotherapy. They included 17 of m/z <3600, which indicated early response to chemotherapy, and a further eight of m/z >3600 that indicated late response. Furthermore, a total of 23 N-glycan signatures (12 N-glycans of m/z <3600 downregulated and 11 of m/z >3600 upregulated in EOC) differentiated primary EOC patients from the controls. Analysis of IgG_1/IgG_2 identified age-dependent cross-ethnic EOC diagnostic glycopeptide signatures. For patients of ages 18–48 years, three IgG_1 and five IgG_2 glycopeptides were identified, while those of age ≥49 years, had seven IgG_1 and four IgG_2 glycopeptides. Seroprevalence rates of 29.1% (HBV), 26.7% (HIV) and 1.2% (HCV) were reported in OC patients.

The N-glycan and IgG_1/IgG_2 glycopeptide signatures showed improved detection of EOC as well as chemotherapy response monitoring compared to CA125 when used alone, while the high HBV and HIV positivity in Kenyan OC patients necessitates routine testing for their optimal management.

1.0 Introduction

Ovarian cancer is the most lethal of all the gynecological malignancies [1], although its prevalence is lower than other female cancers such as breast and cervical cancers. The high mortalities associated with OC are often due to the delay in definitive diagnosis occasioned by the absence of effective early diagnostic biomarkers, unspecific symptoms as well as resistance to anti-cancer agents. Viral diseases including the Human immunodeficiency virus (HIV) and the Hepatitis B virus have also been shown to contribute to the early deaths of gynecological cancer patients, such as those with OC [2-4]. The most promising approach to reducing OC-related mortalities and morbidity lies in the discovery of effective biomarkers capable of detecting OC while still in the early stage (FIGO Stage I and II) when many treatment options are still available. Monitoring the response to anti-cancer treatment agents, resistance prediction as well as screening of HIV and HBV infections in OC patients scheduled for chemotherapy are factors that contribute to the longevity of patients but are not part of standard care worldwide [5, 6].

The current gold standard biomarker for OC is cancer antigen 125 (CA125), discovered in 1981 [7]. However, CA125 suffers from the limitations of low specificity and accuracy, as is the case with other biomarkers such as human epididymis 4 (HE4), and cancer antigen 19-9 that were discovered thereafter [8-12]. The suitability of CA125 in monitoring treatment response is also a matter of conjecture, as there is no established protocol for interpreting consecutive CA125 concentrations for a decision on the state of tumor progression to be made [13]. Based on these existing gaps in early diagnosis of OC and monitoring response to treatment, new effective biomarkers are urgently needed to mitigate the aggressiveness of the disease by reducing mortalities and improving the quality of patients' lives. Glycomics, a comprehensive study of the entire repertoire of glycans as expressed in biological systems under specified physiological and pathological conditions [14], has opened up new possibilities for finding biomarker signatures for diseases and malignancies including OC. Glycosylation is the most common post-translational modification (PTM) process that results in the formation of glycoconjugates (protein or lipid) through the linkage of one or more mono- or oligosaccharide units [15]. Perturbations of glycosylation happen in various malignancies, resulting in cancer-specific glycans, which form the basis upon which new effective biomarkers are hoped to be realized for diagnosis and monitoring the efficacy of anti-cancer agents in many cancers including OC.

Therefore, the overall long-term goal of the present study was to reduce the mortalities and improve the quality of life of EOC patients by finding effective, less invasive, and affordable biomarkers for early detection of EOC and monitoring of the response to anti-cancer agents, along with optimal patient management. This was addressed in the present doctoral thesis in three research questions where: For the first time, an investigation on malignant alterations of *N*-linked protein glycosylation

was carried out on the serum of an African cohort of EOC patients to determine potential *N*-glycan signatures that can be used for early diagnosis of EOC and in monitoring response to chemotherapy. Further evaluation of altered serum immunoglobulin G1 and G2 (IgG$_1$ and IgG$_2$) Fc glycosylation in EOC patients of African and European cohorts was carried out to determine common potential IgG$_1$/IgG$_2$ glycosylation signatures for the early diagnosis of EOC. Finally, this study also sought to determine for the first time, the burden of HIV, HBV, and HCV among EOC patients in Kenya.

1.1 Glycosylation

Glycosylation is a post-translational process that yields different types of glycoconjugates through the joining of saccharides to aglycones (proteins and lipids) by the formation of glycosidic linkages in a carefully enzyme-regulated process [16, 17]. The type of linkages formed and the aglycones involved describe the mechanisms of glycosylation. The major classes of glycoconjugates that result from various glycosylation processes include glycoproteins, glycosaminoglycans, glycolipids (glycosphingolipids), and glycosylphosphatidylinositol (GPI)-linked proteins.

Glycoproteins are formed by the attachment of one or more glycans to the polypeptide backbone via either a nitrogen (*N*-glycans) or an oxygen (O-glycans) linkage to an asparagine or serine/threonine, respectively. Other forms of glycoprotein glycosylation include phosphoglycosylation and C-mannosylation [18]. Glycosaminoglycans, on the other hand, are comprised of long polysaccharides made of repeating disaccharide units that are linked to the hydroxyl group of serine or threonine of a protein backbone in a process initiated by a conserved tetrasaccharide [19]. Glycosphingolipids are formed when one or more glycans are covalently attached via either a glucose or a galactose to the terminal primary hydroxyl group of the lipid [20]. Glycosphingolipids are one of the most abundant human glycolipids found on the outer leaflet of the cell plasma membrane [17]. Glycosylphosphatidylinositol (GPI) is comprised of ethanolaminophosphate, which forms an amide bond by linking its amino acid group with the protein, three mannoses and glucosamine, and then through the terminal inositolphospholipid. As a result, GPI anchors connect whole proteins to the membrane [21].

1.1.1   Glycan diversity

The highly regulated mammalian glycome is composed of glycan structures exhibiting enormous diversity that arises from branching and anomeric linkages [17, 19]. The great diversity exhibited by the mammalian glycans arises from an estimated 700 proteins, which include biosynthetic enzymes such as glycosyltransferases, glycosidases, kinases, and transporters as well as activated sugars and ten monosaccharides [19] [22, 23]. The ten monosaccharides include N-acetyl/glycolylneuraminic acid (Neu5Ac/ Neu5Gc), mannose (Man), glucose (Glc), galactose (Gal), fucose (Fuc), N-

acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), glucuronic acid (GlcA), iduronic acid (IdoA), and xylose (Xy) [18, 22]. Of the 700 proteins, approximately 200 are glycosyltransferase enzymes that are responsible for glycan biosynthesis, a biologically controlled process that is initiated in the Endoplasmic Reticulum (ER) and continues in the Golgi Apparatus [19, 24]. About half of all human proteins are glycosylated, and mostly by *N*-glycosylation [19, 22].   Of the ten monosaccharides, only seven take part in the synthesis of human *N*- and O-glycans. They include one acidic monosaccharide - Neu5Ac, and six neutral monosaccharides - Man, Glc, Gal, Fuc, GlcNAc, and GalNAc. Another type of acidic monosaccharide is Neu5Gc, which is mostly expressed in some malignancies (Figure 1) [25, 26].
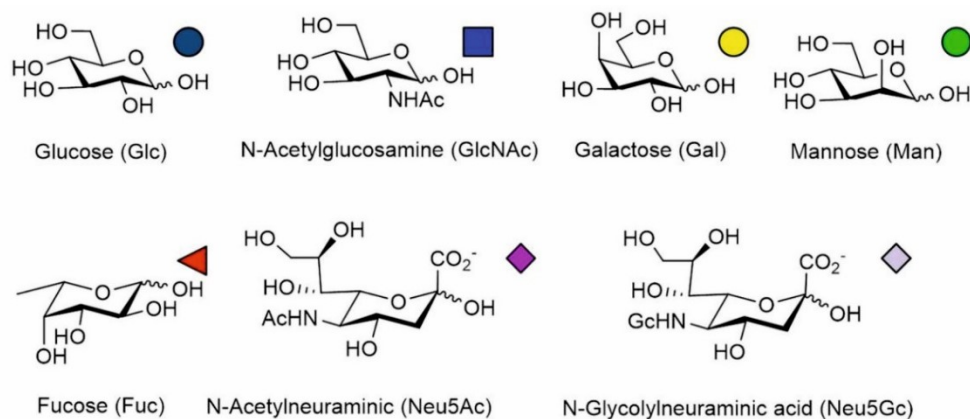
Figure 1. Structures and symbols of common monosaccharides found in *N*-glycans. Adapted from Chao et al., licensed under CC BY [27]

A range of highly branched glycan structures exhibiting high microheterogeneity are also generated from the different ways in which the monosaccharides are linked by the glycosidic bonds. The nature of microheterogeneity displayed by glycans is higher compared to either DNA or proteins, which originate from a conserved template that is linear [28]. *N*-Glycans are formed through linkage with the nitrogen atom of the asparagine (Asp) found in the consensus sequence Asp-X-Ser/Thr motif (where X denotes any amino acid other than proline) [29]. O-Glycans, on the other hand, are usually shorter than *N*-glycans and also exhibit branching. They are formed by the covalent attachment of monosaccharides to the oxygen atom of serine or threonine [30]. *N*-Glycans are the most abundant and complex due to their many branching possibilities, which generate more diversity. Further glycan diversity also arises from their stereoisomeric configurations of α and β-linkages. Glycosaminoglycans differ from the *N*- and O-glycans because of their linear configuration, but just like O-glycans, they are also linked to the serine and threonine of the protein backbone. Glycosaminoglycans are further post-synthetically modified by sulfation and/or epimerization. The diversity of glycosaminoglycans

produced through sulfation includes heparan sulfate, dermatan sulfate and chondroitin sulfate [17, 19]. Moreover, hyaluronan also adds to this type of diversity. Hyaluronan is a unique glycosaminoglycan found freely in the extracellular matrix. It is not attached to either protein or lipid, and is formed by alternate additions of GlcA and GlcNAc [16, 19, 31].

1.1.2 Structural properties and biological functions of *N*-glycans

Mammalian *N*-glycan structures are comprised of a core structure made of two *N*-acetylglucosamine and three mannose residues (Man$_3$GlcNAc$_2$) [32]. *N*-Glycans are of three types, i.e., high-mannose, hybrid- and complex-type structures, formed based on how the monosaccharides are added to the Man$_3$GlcNAc$_2$ core structure (Figure 2). High-mannose-N-glycans (or oligomannoses) are only constituted of mannose residues attached to the core structure. Hybrid-type *N*-glycans are biantennary structures, where one antenna is made up of high-mannoses only, while the second antenna is constituted of one or more different monosaccharides. Complex-type *N*-glycan structures may contain up to six antennae, each originating from GlcNAc residues attached to the Man of the core structure [33].
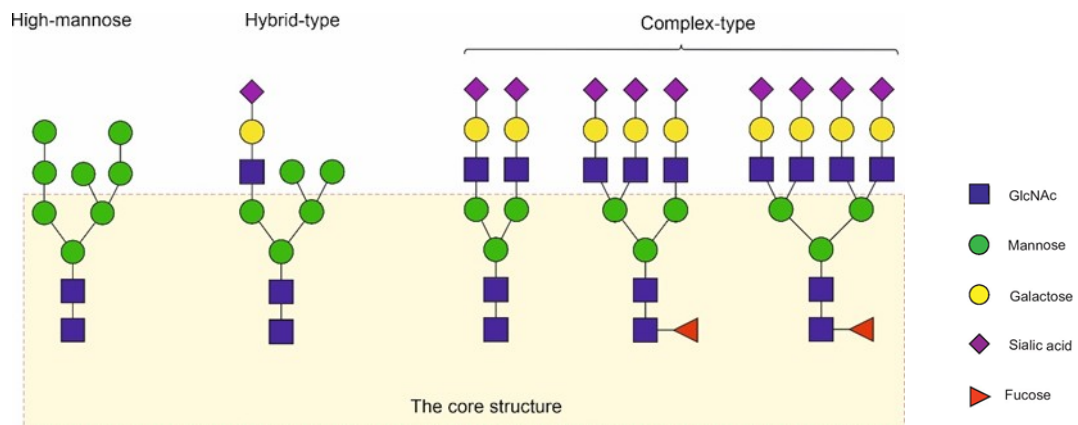
Figure 2. The three types of *N*-glycans, each containing the Man$_3$GlcNAc$_2$ core structure.

Glycosylation of proteins is a major contributor to the functional diversity manifested by proteins. Biological roles played by the *N*-linked glycoproteins include protein folding, provision of a physical barrier to pathogens using a thick glycan layer, and lubrication provided by the soluble and membrane-bound mucins [34]. Others include the solubility of macromolecules, modulation of membrane receptor signaling, anti-adhesive action by charge repulsion to inhibit cell-to-cell interactions, and of fundamentally importance, protection from proteases by steric hindrance and negative charge [34]. *N*-Glycans also play a key role in regulatory mechanisms that control

physiological and pathological processes in the human system [16, 34, 35]. Physiological functions regulated by glycans include aging, fertilization and embryogenesis. Pathological functions include cell-to-cell recognition, cell adhesion, signal transduction, and self/non-self recognition allowing antigen uptake and elimination [34]. Others include molecular trafficking and clearance, endocytosis, modulation of receptor activity and acting as a molecular switch for pro- and anti-inflammatory activities [16, 35].

1.1.3 *N*-Glycan biosynthesis

Synthesis of *N*-glycans is an ordered process that takes place in two organelles, namely the ER and the Golgi apparatus, and proceeds in two phases. It involves specific substrates (activated monosaccharides) and enzymes (glycosidases and glycosyltransferases), which are sequentially utilized to create highly complex but specific glycan structures on the proteins [16, 36]. The glycosidases perform a hydrolyzing role for specific glycan linkages, while the glycosyltransferases synthesize the glycan chains [16]. The two enzymes are differentially expressed according to the physiological and pathological states of the cell [32]. In summary, the first phase of *N*-glycan synthesis commences on the cytosolic side of the ER with the assembly of one or more oligosaccharides to a polypeptide backbone termed dolichol phosphate (Dol-P), via a nitrogen linkage in a highly conserved pathway (Figure 3) [29, 32]. In the second phase, the intermediate Man_5-$GlcNAc_2$-P-P-Dol is flipped into the lumen of ER, where the assembly and quality control of protein proceeds [37]. The well-folded glycoproteins then proceed to the Golgi apparatus to undergo further trimming, modification and maturation [16]. It is within the Golgi apparatus that *N*-glycans are further extended and branched as they transit through this latter portion of the secretory system before reaching the cell surface and extracellular compartments [16].

1.1.3.1 Synthesis of the dolichol precursor

The process of *N*-glycan synthesis is made possible initially by the synthesis of a five-carbon isoprene unit called the dolichol. Its purpose is to serve as a carrier of the oligosaccharide as well as the localization of the first phase of the N-glycan synthesis to the cytosolic membrane of the ER [38]. The enzyme cis-prenyltransferase is responsible for the synthesis of dolichol by the sequential addition of C5 isoprenoid units on the farnesyl-pyrophosphate [37]. On the cytosolic ER membrane, synthesis begins with the addition of GlcNAc from the activated UDP-GlcNAc to the Dol-P by the enzyme GlcNAc-1-phosphotransferase, to form the anhydride dolichol pyrophosphate N-acetylglucosamine (Dol-P-P-GlcNAc); Tunicamycin can inhibit this reaction, hence stopping the process of *N*-glycosylation in the cell [33]. A second GlcNAc residue is added to the Dol-P-P-GlcNAc by the GlcNAc-transferase, again using the substrate UDP-GlcNAc to form Dol-P-P-$GlcNAc_2$, which is also called the

"chitobiose core". The sequence is enlarged further by the addition of five mannose (Man) residues from GDP-Man by the mannosyltransferase to produce the $Man_5GlcNAc_2$-P-P-Dol. The $Man_5GlcNAc_2$-P-P-Dol precursor is then flipped from the cytosolic side to the luminal side of the ER by the $Man_5GlcNAc_2$-P-P-Dol flippase, marking the beginning of the second phase [32, 39]. Importantly, there is also another type of flippase, the Dol-P-P flippase, which plays the reverse role of returning the Dol-P back to the cytosol for reuse in the new biosynthetic cycle [39]. The reverse role of Dol-P-flippase is pertinent to the extent that together with the *de novo* synthesis, they maintain the levels of dolichol in the cytoplasmic side of the ER.

In the luminal side of the ER, four Man residues from the Dol-P-Man donor and three Glc residues from the Dol-P-Glc donor are added additionally onto the $Man_5GlcNAc_2$-P-P-Dol to form a 14-sugar *N*-glycan precursor structure, the $Glc_3Man_9GlcNAc_2$-P-P-Dol [32]. Importantly, the Dol-P-Man and Dol-P-Glc donors are formed on the cytoplasmic side of the ER membrane from the activated monosaccharides GDP-Man and UDP-Glc, which are also flipped into the ER lumen [33, 37]. At this point, the multimeric enzyme complex of four subunits, the oligosaccharyltransferase (OST), catalyzes the cleavage of the 14-sugar precursor structure from the Dol-P-P and its "*en bloc*" transfer to the consensus sequence of Asn–X-Ser/Thr motif of the newly synthesized proteins, which are transiting through the ER [17, 32, 38].

1.1.3.2 Quality control of protein folding

The newly formed nascent glycoprotein undergoes a quality control step that involves the sequential removal of the three Glc residues by the α-glucosidases I and II [32]. At first, α-glucosidases I and II remove the first two Glc to form the $GlcMan_9GlcNAc_2Asn$ structure. This acts as a ligand for the membrane-bound (calnexin) or soluble (calreticulin) lectin chaperones in a step that triggers protein folding [40]. In 1994, Ari Helenius described the calnexin and calreticulin cycle as being the chaperones in the ER that bind to the $GlcMan_9GlcNAc_2Asn$, while the ER contains the glycosyltransferase which re-transfers the glucose to the $Man_9GlcNAc_2Asn$ [41, 42]. Protein folding is prone to errors, hence the chaperone's role is quality control, whereby it recognizes and binds the misfolded proteins in order to refold them into the correct conformation [43]. However, misfolded proteins that cannot be rectified by chaperone intervention are allowed to enter the ERAD process (ER-assisted degradation), where they are degraded by the Man trimming enzymes based on the "mannose timer" model of ER quality control. In this model, the glycoproteins with slow folding speed, possibly due to mutations or incomplete oligomeric assembly, are eliminated [19, 40]. Overexpression of the ER mannosidase 1 leads to accelerated misfolded protein degradation by ERAD. However, a second α-mannosidase 1-like protein called EDEM (ER degradation-enhancing α-mannosidase like protein) accelerates the ERAD response to misfolded N-glycoproteins [43].

Misfolded glycoproteins showing slow cleavage of the terminal α1,2 Man by the EDEM are further recognized by OS9 and degraded in the ER before they exit to the Golgi [32]. As part of the housekeeping during protein folding, the lectins protect the nascent polypeptide from the possible effects of hydrophobic aggregation and non-qualitative disulfide bonding as well [44]. Glycoproteins that are properly folded undergo further trimming by the α-glucosidase II to remove the last Glc residue, resulting in the formation of $Man_9GlcNAc_2Asn$. Finally, before the glycoprotein exits the ER to the Golgi apparatus, the terminal α1,2 Man residue of the central arm is cleaved by α-mannosidase I to yield $Man_8GlcNAc_2Asn$ [19, 32, 45].

1.1.3.3 Processing, maturation, and formation of *N*-glycan diversity

In the cis-Golgi region, the nascent glycoprotein $Man_8GlcNAc_2$-Asn undergoes further trimming of three Man residues by a set of α1-2 mannosidases - 1A and 1B, to form the $Man_5GlcNAc_2$-Asn precursor. This molecule is responsible for mediating the formation of hybrid- and complex-type *N*-glycans [32]. $Man_5GlcNAc_2$-Asn is then transferred to the medial-Golgi for further processing and maturation. However, the structure may also remain unchanged, in which case the secreted glycoprotein or the mature membrane manifests $Man_{5-9}GlcNAc_2$-Asn [32, 46].

Maturation of *N*-glycans by the formation of hybrid- and complex-type structures takes place within the medial and trans-Golgi regions, with the $Man_5GlcNAc_2$-Asn as the precursor. The N-acetylglucosaminyltransferase-I (GlcNAcT-I or MGAT-1) transfers a GlcNAc residue to C-2 of the α1-3 Man in the core of the $Man_5GlcNAc_2$-Asn precursor to form $GlcNAc$-$Man_5GlcNAc_2$-Asn, and initiates synthesis of the hybrid- and complex-type *N*-glycan structures on nascent glycoproteins. Formation of the hybrid *N*-glycans involves retention of the five Man residues, while a Gal and a Neu5Ac may be added to the arm that received the GlcNAc. By contrast, complex *N*-glycans are formed by the removal of two Man residues from the $GlcNAcMan_5GlcNAc_2$-Asn by the action of α-mannosidase II to form $GlcNAcMan_3GlcNAc_2$-Asn. Subsequently, GlcNAcT-II (MGAT2) adds a second GlcNAc to the α1-6 linked core Man to form a biantennary complex-type *N*-glycan [32, 46].

From the biantennary *N*-glycan structures, about six more diverse branches can form. For instance, triantennary *N*-glycans are formed by the transfer of one β 1,4-linked GlcNAc to the tri-mannosyl core by the action of GlcNAcT-IV (MGAT4A). Additionally, the tetraantennary *N*-glycans result from the addition of the fourth β 1,6-linked GlcNAc to the trimannosyl core by the action of GlcNAcT-IV (MGAT4B) and GlcNAcT-V (MGAT5) [47]. Moreover, for both hybrid- and complex-type *N*-glycans, GlcNAc may also be added to the first Man of the core by the action of GlcNAc-T-III (MGAT3) to form bisecting *N*-glycan structures [32]. Some of the hybrid- and complex *N*-glycan structures extend their branches further by additions of Gal to the GlcNAc to yield the ubiquitous building block known as

the "LacNAc" sequence, which becomes poly-LacNAc when other LacNAcs are sequentially added. Where β-linked GalNAc is added to GlcNAc instead of Gal, GalNAcβ1-4GlcNAc (or LacdiNAc) is formed [32]. Following the reactions described during maturation, various types of *N*-glycans are formed, including the high-mannose, hybrid- and complex-type structures that differ in terms of antennarity, composition and length, yielding the great diversity that is the *N*-glycans repertoire. High-mannose *N*-glycans are formed in the cis-Golgi, while the hybrid- and complex-type structures are formed in the trans-Golgi.
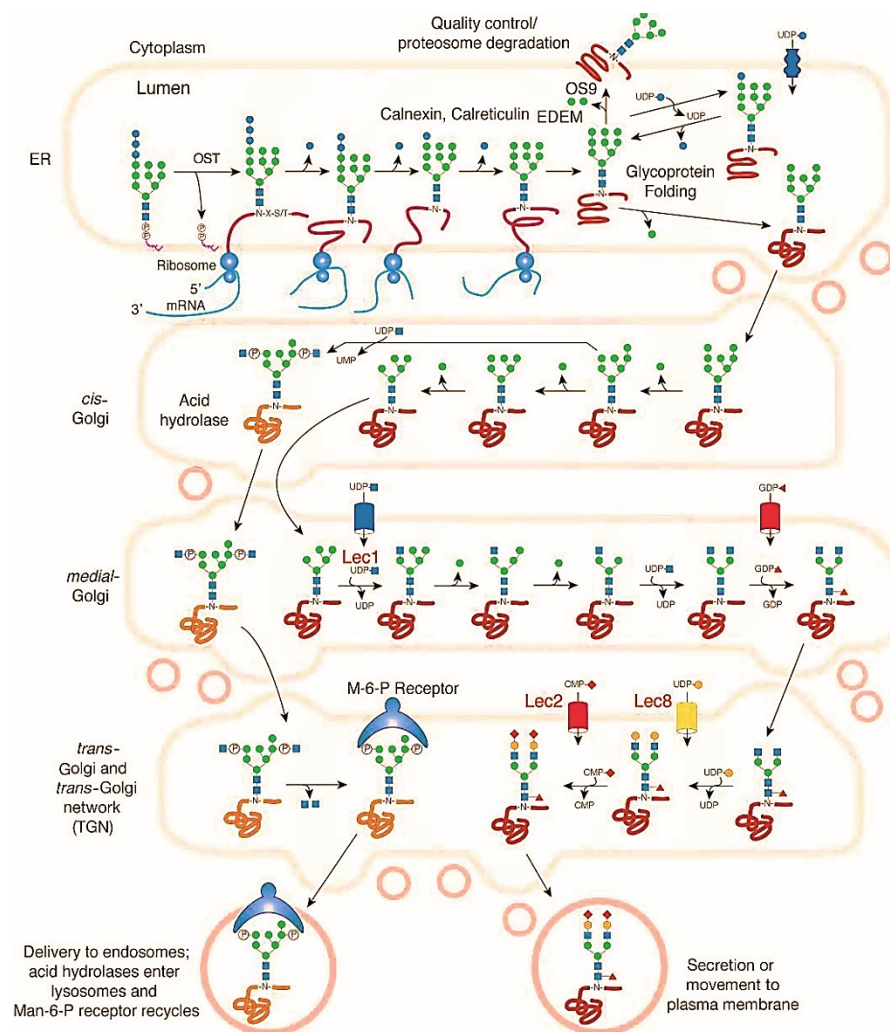
Figure 3. Biosynthesis of the N-glycans. The process begins with the attachment of GlcNAc to Dol-P to form GlcNAc-P-P-Dol, followed by the addition of a second GlcNAc and 5 Man to form $Man_5GlcNAc_2$-P-P-Dol, which is flipped into the lumen of ER. Then 4 Man and 3 Glc are added to form $Glc_3Man_9GlcNAc_2$-P-P-Dol, the 14 sugar N-glycan precursor, which is added to the nascent protein via Asn-X-Ser/Thr to form $Glc_3Man_9GlcNAc_2$-Asn. Then, 2 Glc are removed to form $GlcMan_9GlcNAc_2$-Asn, which initiates protein folding. For properly folded proteins, the remaining Glc is removed, followed by a terminal Man to form $Man_8GlcNAc_2$-Asn, which moves into cis-Golgi. 3 Man are then removed from the structure to form $Man_5GlcNAc_2$-Asn, and then transferred to the medial and later trans-Golgi for further processing as well as maturation, forming hybrid- and complex-type structures. Adapted from Essentials of Glycobiology 3rd edition, Stanley et al., under license CC BY-NC-ND 4.0 (additional color added) [32].

1.1.4   Inherited disorders of *N*-linked protein glycosylation

The defects of glycan biosynthesis lead to a group of diseases known as congenital disorders of glycosylation (CDG). CDGs are categorized into four classes depending on the mechanism of glycosylation involved. They include *N*-linked protein glycosylation, O-linked protein glycosylation, a combination of *N*- and O-glycosylations, and finally, GPI and lipid biosynthesis [48]. The majority of CDGs that result from *N*-linked glycosylation defects are severe, with clinical manifestations being felt in the muscular tone, brain development, neurological functions, growth and development [49]. Other notable symptoms of *N*-linked CDGs are hepatic fibrosis, nipple inversion, visual defects, and coagulation impairment, amongst others [17, 49]. *N*-Linked CDGs are classified into type 1 (CDG-I) and type II (CDG-II) [50]. For CDG-I, the defect occurs during *N*-glycan biosynthesis on the glycolipid precursor, before it is attached to the asparagine of the glycopeptide backbone. These defects are manifested in the synthesis and processing of the dolichol, in transport from the cytosolic face of the ER to the lumen, and also during the presentation or transport of the sugar residues [51]. By contrast, CDG-II, arises from the incomplete protein-bound glycans due to processing defects in the Golgi apparatus [17, 50]. The most common *N*-glycan-related CDGs include phosphomannomutase 2-CDG (PMM2-CDG), mannose phosphate isomerase-CDG (MPI-CDG), and ALG6-CDG.

PMM2-CDG is caused by the mutation in the phosphomannomutase 2 (PMM2) gene that encodes the corresponding PMM2 enzyme [17]. Its function is to catalyze the conversion of Man-6-Phosphate to Man-1-Phosphate, which is a precursor of GDP-Man and Dol-P-Man synthesis. These two are substrates for the mannosyltransferases that play a catalytic role in synthesizing the 14-carbon $Glc_3Man_9GlcNAc_2$-P-P-Dol [50, 51]. Currently, there are no therapeutic formulas available for patients with PMM2-CDG. It may present in two forms, the more severe form being called infantile-onset, where death occurs within the first 2 years of life, and the late-onset form, where the subjects have normal life expectancy but may present intellectual disability [51].

MPI-CDG, by contrast, is caused by the mutation of the MPI gene that is responsible for encoding the mannose phosphate isomerase, whose role is to catalyze the conversion of fructose-6-P to Man-6-P. However, Fructose-6-P can still be metabolized by the glycolytic pathway preventing its accumulation [48]. Patients suffering from this condition exhibit symptoms such as stunted growth, coagulopathy, hepatic fibrosis, severe diarrhea, and vomiting, but brain development is not impaired. However, they can be effectively treated with dietary supplements of mannose [50].

Finally, ALG6-CDG results from the mutation of the ALG6 gene that encodes glucosyltransferase 1. In *N*-glycan biosynthesis, three Glc residues are added sequentially to form the 14-carbon structure $Glc_3Man_9GlcNAc_2$-P-P-Dol [51, 52]. A deficiency of glucosyltransferase 1 interferes with the

attachment of the first Glc to the intermediate structure $Man_9GlcNAc_2$-P-P-Dol. Patients suffering from this condition present with seizures, hypotonia, and strabism [51].

1.1.5   Alterations of *N*-glycosylation in malignancy

Aberrant glycosylation of the total *N*-glycome is an established feature of many malignancies including ovarian cancer. The state of glycosylation is reflected by an individual's age, sex, and lifestyle [53-55]. However, in general, healthy individuals have been shown to exhibit remarkable stability in the glycosylation state [56]. Abnormal *N*-glycosylations have previously been demonstrated in various diseases, such as rheumatoid arthritis [57, 58], and systemic lupus erythematosus [59], including several malignancies [60-65]. *N*-Glycans are of particular interest in the biomarker research of diverse malignancies because structural alterations have previously been demonstrated in the tumor microenvironment, which are thought to promote invasion, progression and the spread of various tumors [31, 65-68]. Consequently, they have formed the basis upon which research for new biomarkers is being explored.

Previous reports have pointed to the fact that glycosylation changes could be a consequence of cancer invasion, though other research groups have also suggested that glycan alterations could be the cause of cancer [69]. Hakomori and Kannagi made two presuppositions on the key mechanisms thought to be responsible for the glycan alterations in malignancies - they included incomplete synthesis and neo-synthesis [70]. They indicated that incomplete synthesis occurs in the early cancer stages, and is supported by the transformation of normal glycosylation to the truncated glycosylation pattern that yields truncated structures such as sialyl Tn [17, 71]. Already, the expression of sialyl Tn structures has been shown in malignancies of the breast and gastrointestinal tract [72, 73], while for the neo-synthesis, they indicated that structures such as sialyl lewis^a and sialyl lewis^x (sLe^a and sLe^x), which are thought to promote cancer metastasis, are produced and have been associated with late-stage cancer [74, 75]. Some cancer types that have been shown to express sLe^a and sLe^x include prostate cancer, colorectal carcinoma, gastrointestinal cancer, and non-small cell lung cancer [60-63].

The reasons for glycan modifications in malignancies are multi-faceted and are aided by many factors. They include alteration in the expression of glycosyltransferases, the supply of acceptor substrates of glycosyltransferases, and increased availability of sugar nucleotide donors and cofactors [17, 31, 76]. Others involve alterations in the tertiary conformation of the peptide and the chain of the nascent glycan [71]. Decrease or increase of the glycosyltransferases expression as a function of oncogenic transformation affects both the core structure of the glycan and its function through increased antennarity. The specific transformation of the saccharides themselves to form unique terminal glycan motifs is also significant in this process [71]. The growth of tumors and their subsequent

migration to the metastatic sites relies on the ability of the malignant cells to escape the cellular division checkpoints, evade apoptotic signals and tumor immune surveillance mechanisms, and resist therapy [77]. The *N*-glycosylation changes during cancer are key enablers of the aforementioned factors and therefore offer much promise in cancer biomarker discoveries.

The most common and significant cancer-associated glycosylation changes that happen at the level of total *N*-glycome of blood glycoproteins are manifested in sialylation, fucosylation, bisection, and in *N*-linked glycan branching [64, 78-81]. Other *N*-glycosylation changes have also been reported at the immunoglobulin level and acute-phase proteins. At the immunoglobulin level, especially the IgG subclass, altered galactosylation and sialylation have been reported in various malignancies, such as colorectal, ovarian, and gastric cancers [82-85]. Previous studies have also reported increased sialylation as well as sLe^x structures, triantennary and tetrantennary structures, and core-fucosylated biantennary digalactosylation on the acute-phase proteins, such as haptoglobin, α1-acid glycoprotein, and α1- antichymotrypsin, in many cancers, including pancreatic, breast and ovarian cancers [86-89].

1.1.5.1 Sialylation

Sialic acids are nine-carbon acidic monosaccharide units usually found attached to the terminals of glycoproteins and glycolipids performing diverse roles [90]. Sialylation is essential for the stabilization of the glycoprotein structure [77]. Their characteristic positioning on the cell membrane, their electronegative charge, and their hydrophobic nature enable them to participate in functions such as cell signaling, cellular adhesion, ion transport, and recognition, including pathogen and toxin binding [31, 90, 91]. Other roles include conferring the negative charge on the erythrocytes to ward off unwarranted cellular interactions through repulsion, modulation of protein half-life in circulation, and influencing fertilization by facilitating sperm-egg interaction [90, 92].

The "sialome", which describes the diversity of sialic acid, is estimated to be made up of about 50 sialic acids. All known sialic acids are derived from four core molecules that vary at the C-5 carbon linkage, contributing immensely to the sialome. The core sialic acid molecules are *N*-acetylneuraminic acid (Neu5Ac), which is the most common in humans, N-glycolylneuraminic acid (Neu5Gc), 2-keto-deoxynonulosonic acid (Kdn) and Neuraminic acid (Neu) (Figure 4) [93]. The main mammalian sialic acids are Neu5Ac and Neu5Gc, where each differs from the other by a single oxygen atom [94]. The oxygen atom is added to the sialic acid structure, i.e., precursor CMP-Neu5Ac, by cytidine monophosphate N-acetylneuraminic acid hydroxylase (CMAH) to form CMP-Neu5Gc, but human beings lack this enzyme due to mutation of the CMAH gene [95]. Further diversity of sialic acid is due to the formation of α-linkages between carbon-2 and other monosaccharides as well as the anomeric

configurations. These are α2,3- or α2,6-linked to the Gal, α2,6 linked to the GalNAc or GlcNAc and α2,8 or α2,9 linked to another sialic acid. Anomeric configuration of sialic acid is also a fundamental contributor to the sialome. Sialic acids exist in two anomeric forms, namely α-anomer (the bound form on the glycans) and the β-anomer form, which forms over 90% of the sialic acid in solution [96].
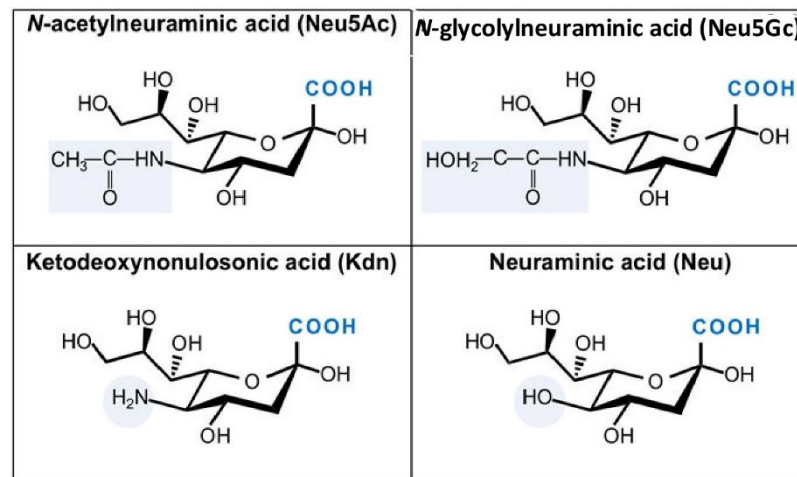
Figure 4. Primary core sialic acid structures, by Paul and Padler-Karavani, licensed under CC-BY NC [97].

Sialylation is a cellular glycosylation modification that involves the addition of sialic acid residues on the glycoconjugates [31], in a process that takes place in the Golgi apparatus. Sialyltransferases are responsible for regulating sialylation and are classified into four families based on the type of carbohydrate linkage they form as well as the structure to which they are transferred. For instance, the ST3Gal family (comprised of six ST3Gal1-6) form the α2,3-linkage with underlying Gal in a glycan, and the ST6Gal family (comprised of two members) form α2,6-linkage with underlying Gal residue. Others include the ST6GalNAc family (six members), which form α2,6-linkage to an underlying GalNAc, and the ST8Sia family (comprised of six members), which form α2,8-linkage with a terminal sialic acid [36, 98].

Malignant transformation of cellular surface sialylation following marked upregulation of the sialyltransferase is a major indicator of pathology. However, the levels of expression by the individual sialyltransferases vary quite significantly between different malignancies, and within tumors as well [99]. It is of fundamental importance that these cancer-mediated modifications, which are reflected in the expression of sialyltransferases and the resultant increased sialylation, have become plausible sources of new cutting-edge cancer biomarkers. Furthermore, the changes in the N-glycome also offer the opportunity for the discovery of new target-driven treatment strategies to minimize the emergence of resistance to chemotherapy agents. Indeed, the findings of various studies reported

perturbations of sialic acid expression in the body fluids and histological sections of patients with various diseases and malignancies such as ovarian carcinoma. Increased sialylation in cancer includes upregulation of sialylated derivatives of Lewis antigens (sialyl-lewis x, sLe[x] and sialyl-lewis a, sLe[a]), which are known to promote metastasis as they are the ligands of selectins [100, 101]. Selectins are vascular cell adhesive molecules expressed on the surface of endothelial cells and leukocytes, which are known to play immunological roles, such as in inflammation, wound repair, immune responses and cancer progression, through their interaction with sLe[x] [31, 102]. Increased sialylation has also been correlated with increased tumor migration, invasion, and cell adhesion, and hence poor patient prognosis [103-105]. Several studies reported increased α2,6- and α2,3-linked sialylation in various malignancies including ovarian cancer (OC) [28, 80, 106, 107]. Wang et. al observed that increased α2,3-linked sialylation in OC patients correlated with the increase in expression of ST3Gal1 [108].

The cancer cell surfaces of OC patients are characterized by increased expression of the sialylated lewis (sLe[x]) epitopes on terminal glycans [109]. This is a result of altered regulation of fucosyltransferases in hepatocytes, whereby the expression of α1,3 and α1,4 fucosyltransferases are increased, while α1,2 fucosyltransferases are suppressed. The α1,2 fucosyltransferases are suppressed as a result of being outcompeted off the substrate, which they share with the α2,3 sialyltransferase [109, 110]. Moreover, α1,3- and α1,4 fucosyltransferases facilitate fucosylation of the already sialylated core structure, resulting in increased sLe[x] structures as seen in OC [110]. The α2,6 sialyltransferase attaches the α2,6-linked sialic acid to the biantennary glycans, while the α2,3-linked sialic acids are majorly attached to the tri- and tetraantennary glycans [47]. The consequence of increased sLe[x] in malignancy is enhanced cancer invasiveness and metastasis through the adhesion of tumor cells to the endothelium [36, 109]. It is also important to note that sLe[x] are the ligands for selectins, whose interaction leads to the formation of aggregates of tumor cells and platelets, which contribute to tumor cell migration and disease progression [61]. Other mechanisms through which the increase in sialylation promotes tumor progression are the upregulation of integrin activity, suppression of galectin-3-induced apoptosis, and induction of resistance to some chemotherapy agents and immune evasion (Figure 5) [111].
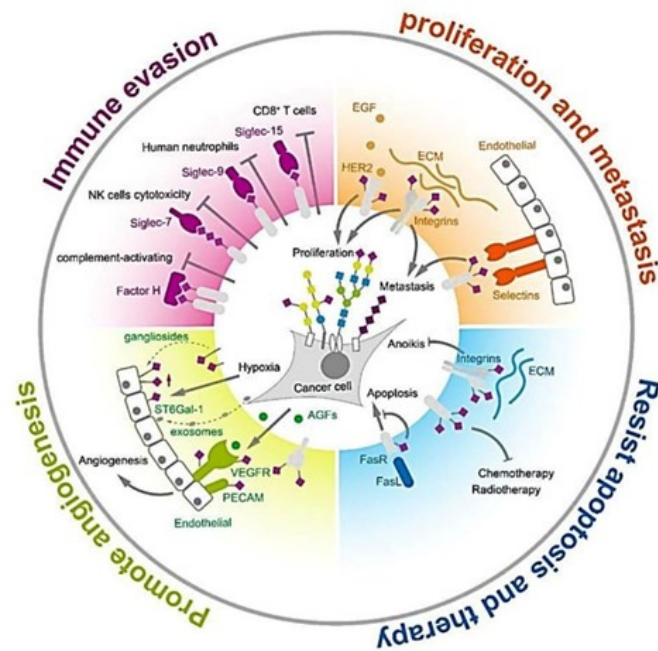
Figure 5. Functions of sialic acid in tumor biology. By Zou et al., licensed under CC-BY 4.0 [111].

1.1.5.2 Bisecting structures

Decreased expression of the bisecting *N*-glycan structures is a common occurrence in tumor malignancies including OC. Bisecting structures are formed through the enzymatic addition of bisecting β1-4-linked GlcNAc residue to the core mannose of the *N*-glycans. The enzyme responsible for this catalytic reaction is GlcNAcT-III [36]. Bisecting *N*-glycan structures promote tumor suppression by inhibiting the addition of the branched complex *N*-linked glycans that aid the spread of malignant cells [8]. The decrease of bisecting structures in OC, therefore, is a predictor of poor patient prognosis as it promotes tumor progression and metastasis by default. Previously, Zahradnikova et al. reported a correlation between the bisecting GlcNAcylated biantennary structures and the resistance to primary chemotherapy [112]. In view of the above, both bisecting and tri-/tetra-antennary structures present opportunities for potential new biomarkers for predicting prognosis as well as resistance to anti-cancer treatment.

1.1.5.3 Branching structures

Increased *N*-glycan branching is one of the most conspicuous and well-characterized malignant transformations apart from sialylation. Increased branching leads to the formation of many triantennary and tetraantennary structures, which in turn provides more sites for the addition of terminal sialic acid [103]. Furthermore, an association was established between increased tri- and tetraantennary structures with cancer progression, due to their negative correlation with bisecting structures that are inhibitory to cancer progression [8]. Increased triantennary and tetraantennary

structures have been reported in malignancies such as prostate, breast, lung and ovarian cancers [28, 79, 106, 113, 114]. Increased *N*-glycan branching expression in cancer is mediated by the upregulation of the GlcNAcT-IV and GlcNAcT-V, which catalyzes the addition of β1-6 GlcNAc to the *N*-glycan core. The inverse relationship between branching and bisecting structures is informed by enzymic competition of the GlcNAcT-III (responsible for the formation of the bisecting structures) against GlcNAcT-IV and GlcNAcT-V (responsible for the formation of branches) [69]. GlcNAcT-III competes with GlcNAcT-IV and GlcNAcT-V, whereby once a bisecting GlcNAc has been added to the core mannose of the biantennary *N*-glycan, GlcNAcT-IV and GlcNAcT-V can no longer form more branches [31, 115]. Similarly, tri- and tetraantennary structures inhibit GlcNAcT-III, preventing the formation of the bisecting structures which are critical in suppressing the tumor; hence, they promote cancer growth and metastasis by default [79].

1.1.5.4 High-mannosylation

An increase or decrease of high mannosylated glycans due to altered expression of the mannosyltransferases has been reported in different types of malignancies. Decreased high-mannosylation has been reported in malignancies such as ovarian cancer and gastric cancers [68, 107, 116, 117]. Conversely, an increase in high-mannose has been reported in breast cancer [114]. Although aberration of high-mannose has been reported in various cancers, the role played by the high-mannose in tumor growth remains to be elucidated.

1.1.5.5 Fucosylation

Fucosylation is expressed either as a core or terminal in a process regulated enzymatically via fucosyltransferases (Fut). Eleven fucosyltransferases are involved in the synthesis of fucosylated glycans, and are functionally classified into four groups, including: Fut1 and Fut2 (synthesis of α1,2 fucose); Fut3, Fut4, Fut5, Fut6, Fut7, and Fut9 (synthesis of α1,3/α1,4 fucose); Fut8 (synthesis of α1-6 fucose); and Fut10 and Fut11, whose function is yet to be elucidated [118].

Core-fucosylation is formed by the enzymatic addition of α1-6 fucose to the innermost GlcNAc residue by α1,6-fucosyltransferase (Fut8). Terminal fucosylated glycans (Lewis epitopes) are synthesized by Fut3-9, whereby the fucose residue is attached to the α2,3 and or 4 linkages at the terminus of the *N*-glycan structure, producing specific lewis antigens such as Lewis x/Lewis y (Le[x/y]) and Lewis a/Lewis b (Le[a/b]) [118]. Lewis antigens are categorized into two types, type 1 and type 2. Type 1 includes H1, Lewis[a] (Le[a]), Lewis[b] (Le[b]), and sialyl lewis[a], which share a common structure consisting of β(1,3)-linked Gal-GlcNAc disaccharide, while Type 2 is comprised of H2, lewis[x] (Le[x]), Lewis[y] (Le[y]) and sialyl lewis[x] (sLe[x]) containing the β(1,4)-linked Gal-GlcNAc disaccharide [119, 120]. The biosynthetic pathway of the sialyl Lewis (sLe) antigens comprises of fucosylation of α1,3 or α1,4 of the previously α2,3-sialylated type 1 or type 2 (sLe[x]) chains. Generally, expression of type 1 lewis antigens in malignancies has been shown to decrease, while type II increases [121, 122].

The upregulation of fucosyltransferases is responsible for the increased core- or terminal fucosylation reported in many cancers [31, 123]. Experimental studies have correlated higher metastatic potential and mortalities in breast cancer and non-small cell lung cancer patients due to overexpression of Fut8 and core-fucosylation [124, 125]. Several studies also reported upregulation of core-fucosylation in various cancers including OC [64, 116, 126], colon cancer, and lung cancer [127, 128]. However, decreased core-fucosylation was also reported in gastric and prostate cancers due to the downregulation of Fut8 [28, 129, 130]. Malignant increases of terminal fucosylation have also been reported, where an increase of sLe^x was associated with a poor prognosis of breast cancer [122]. Furthermore, an increase of Le^x and sLe^x epitopes was reported in colon cancer, whereby in cases where both were expressed, the patient had a shortened survival period compared to patients who did not express the two epitopes [131]. Of note, α-fetoprotein, which is core-fucosylated, is already an approved biomarker for the early detection of hepatocellular cancer [132]. Hence, the increase in core-fucosylation evident in ovarian cancer is a potential area in which discoveries of effective biomarkers for the diagnosis of OC as well as monitoring the response to treatment can be made. Additionally, the Lewis epitopes also demonstrate the potential for developing predictive biomarkers for the patient's prognosis [120, 133, 134].
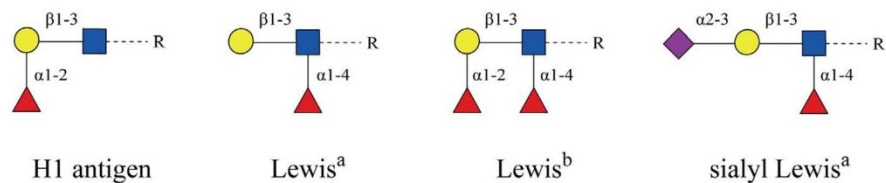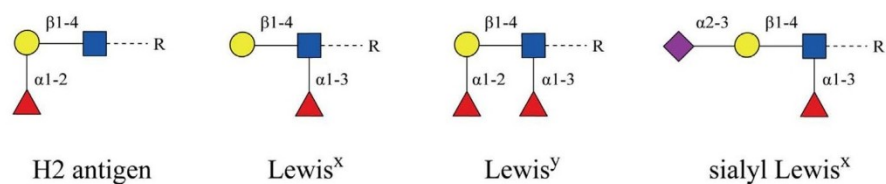
Figure 6. Terminal fucosylated glycans, lewis antigens, which are altered in malignancies. By Blanas et al., under license CC-BY [120].

1.2   Immunoglobulin G Fc *N*-linked glycosylation

Immunoglobulins (Ig) are the major class of serum glycoprotein, comprising five classes: IgA, IgM, IgG, IgD, and IgE. Structurally, the human immunoglobulin is made up of two identical heavy chain (HC) and two identical light chain (LC) subunits (Figure 7). The HC is further divided into four domains, comprising one variable (VH) and three constants (CH$_1$, CH$_2$, and CH$_3$), while the LC has two domains,

one variable (VL) and one constant (CL) [135]. The two light chains, with a part of the heavy chain, (VH and CH1), form the two fragment antigen-binding (Fab) sites, which are linked to the fragment crystallizable (Fc), made up of the CH2 and CH3 domains by a flexible hinge region [136]. Fab and the Fc regions are the functional sites of immunoglobulins and are differentially glycosylated at the variable domain [57]. Functionally, glycosylation controls the immune regulatory roles of the immunoglobulins [35]. The role of Fab in immunity is to bind the specific antigen for neutralization and elimination. By contrast, the Fc region executes the effector function by binding the gamma receptors (FcγRs) of various immune cells to trigger an immune response when an antigen binds [137]. The types of immune responses triggered include complement activation, complement-dependent cytotoxicity (CDC), antibody-dependent cell-mediated cytotoxicity (ADCC), antigen neutralization, phagocytosis, and opsonization [138]. The FcγRs are found on immune cells, such as lymphocytes, neutrophils, eosinophils, macrophages, and natural killer cells, which are tasked with implementing the immune responses against antigens.

The nature of glycosylation reflected by human immunoglobulin is a function of an individual's physiological (age, sex, pregnancy status), and pathological states [139, 140]. In a normal physiological state, an individual's immunoglobulin glycosylation is normally stable and reproducible. Several studies have shown perturbations of immunoglobulins' glycosylation, especially of either the IgA, IgG, or IgM classes, in various malignancies including ovarian cancer, colorectal cancer, breast cancer and gastric cancer, amongst others   [84, 141, 142]. In one of the ovarian cancer studies, for instance, it was reported that although the IgA, IgG, and IgM classes had transformed Fc glycosylation, IgG had the most pronounced effect compared to other immunoglobulin classes [83].

IgG is the most abundant class of immunoglobulin in human circulation, made up of four subclasses (IgG$_1$, IgG$_2$, IgG$_3$, and IgG$_4$) that are differentiated from each other based on the constant regions of their polypeptide chain (γ-chain sequences and patterns of the disulfide bridge) [57, 135]. All IgG molecules contain a single *N*-glycan attached to a highly conserved site, Asn-297, in each of the CH$_2$ domains of the Fc region (Figure 7). The Fc *N*-glycan is responsible for maintaining the quaternary structure and stability of the Fc fragment [137]. By contrast, the Fab region may contain one or more *N*-glycosylation sites restricted to the variable domains and have no conserved glycosylation sites [138]. The Fc-linked *N*-glycans are characterized by a high degree of core-fucosylation and complex-type biantennary structures with variable numbers of galactoses that result in many glycoforms. The number of galactoses in a glycoform varies between zero and two: G0 (zero or agalactose), G1 (one or monogalactose) and G2 (two or digalactose). Furthermore, a few of these glycans may be bisected or contain terminal sialic acids [143].
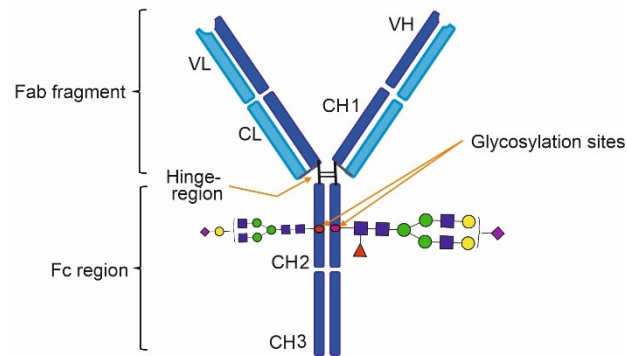
Figure 7. Schematic presentation of an immunoglobulin G structure
showing two *N*-glycans attached to Asn in the CH2 region.

1.2.1   IgG Fc *N*-glycosylation changes in cancer pathology

The composition of IgG glycoforms is usually stable and reproducible in a normal physiological state
[144]. Therefore, when quantified, deviations in the IgG Fc glycosylation that happen in various
pathological processes, such as in autoimmune diseases and cancer [59, 145-148], can help to
differentiate normal from diseased states. IgG Fc *N*-glycosylation is a key effector of immunity
functions such as pro-inflammatory, anti-inflammatory, and antibody-dependent cellular cytotoxicity
(ADCC). Accordingly, the aberrations of IgG Fc *N*-glycosylation in various pathologies not only offer
invaluable opportunities for new signature biomarkers, but also new insights that contribute to
further understanding the molecular basis of diseases and malignancies.

1.2.1.1 Agalactosylation

Expression of an individual's IgG Fc galactosylation (IgG-G1/G2) and agalactosylation (IgG-G0) exhibits
relative stability in normal health. Of the total IgG *N*-glycan pool, IgG-G1 and IgG-G0 account for about
35% each, while IgG-G2 makes up about 16% [138, 144]. The degree of IgG Fc galactosylation and
agalactosylation are influenced differently according to the prevailing physiological or pathological
stimuli.

Physiological variations in the levels of galactosylation are reflected in age, sex, and pregnancy.
Galactosylation was reported to progressively increase with age from birth to around 27 years and
then decrease as one advances in age to a maximal value of less than 50% throughout one's lifetime
[54, 138, 149]. Furthermore, variations in the degree of galactosylation according to sex have also
been implied. Females of the same age as males are reported to express higher galactosylation
compared to their male counterparts. Furthermore, in pregnancy, an increase in the expression of
IgG-G2 and IgG-G2S1 was found to be proportional to the term of pregnancy, reaching a peak at week
30, and then decreasing after delivery as well as in menopause [138, 150, 151]. This points to the
possible role played by the estrogen hormone, as Chen et al. previously reported [152], although the
role is so far yet to be confirmed.

In various pathologies, galactosylated Fc *N*-glycans have been associated with modulation of the inflammatory process by initiating an anti-inflammatory signaling cascade through binding to the inhibitory FcγRIIB [138, 153]. Evidence of the decrease in the degree of IgG galactosylation in the serum of patients with various autoimmune diseases and malignancies has been adduced by various studies. Earlier studies on autoimmune diseases, such as rheumatoid arthritis [57, 58], and systemic lupus erythematosus [59], showed decreases in the levels of IgG galactosylation. Furthermore, several cancer studies also reported a decrease in galactosylation, for instance, in multiple myeloma [154], prostate, gastric, lung, and breast cancers [146, 155-157].

Conversely, IgG agalactosylated glycoforms are reported to promote the pro-inflammatory response. The higher affinity of the agalactosylated IgG glycoforms to bind the FcγRIII, thereby inducing pro-inflammatory activities, is contributed to by their lack of sialic acid or galactose residues (Figure 8) [139, 158]. It is thought that the increase in agalactosylation could be due to a decrease in the expression of galactosyltransferases, and not as a consequence of elevated immunoglobulin synthesis by the B cells [109, 159]. This was confirmed by Parekh et al., who showed that the increase in IgG-G0 glycoforms did not correlate with changes in the serum IgG concentration [160]. However, the current state of conflicting views over the control of galactosyltransferases, namely as to whether it is at the transcriptional or the post-transcriptional level, calls for the undertaking of targeted research to help elucidate the cause of the decrease in galactosyltransferases. An increase in the serum levels of IgG-G0 has been reported in various malignancies, such as ovarian cancer, gastric cancer, prostate cancer and breast cancer [82, 146, 161, 162]. Besides malignancy, autoimmune diseases such as rheumatoid arthritis, systemic lupus syndrome, Crohn's disease, and Sjögren's syndrome have also presented with increased agalactosylation [148, 160, 163]. The chronic inflammation that characterizes ovarian cancer can be ascribed to the increase in agalactosylation.

1.2.1.2 Sialylation

IgG Fc sialylated (mono- and very rarely di-sialylated) structures in healthy individuals account for approximately 10 - 14% of the total glycan pool [138, 144, 164]. Sialic acids are thought to act as the molecular switch between pro- and anti-inflammatory responses, which may be triggered by the change in homeostasis [138]. The presence of terminal sialic acid(s) is a trigger for an anti-inflammatory response, whereas their absence promotes a pro-inflammatory response [138, 165] (Figure 8). For instance, terminal α2,6- linked sialic acid is anti-inflammatory and acts by decreasing the affinity of IgG Fc from binding the activating FcγRs, hence promoting DC-SIGN, which leads to increased expression of the inhibitory FcγRIIB [149]. Altered sialylation is mediated by changes in the expression of sialyltransferase and the sialidases [80, 166] arising from the homeostatic disturbance. Research by independent groups has associated sialylation with possibilities for various biomarkers roles. Kemna et al. reported decreased sialylation and galactosylation of IgG_1 as an important

antecedent to reactivation of granulomatosis with polyangiitis, which could hence serve as indicators for the necessitating of pre-emptive therapy [167]. Furthermore, patient response to the treatment of Kawasaki disease and Guillain-Barre syndrome was shown to correlate with restoration of the serum sialylation of the autoantibodies [168, 169]. These findings, therefore, are suggestive of the possibility of using sialylation as a biomarker for monitoring disease progression and efficacy of treatment as well as detection of relapse.

1.2.1.3 Bisecting GlcNAc

The bisecting GlcNAc containing IgG Fc glycans accounts for about 10 -15% of the total glycan pool [138, 144, 164]. The bisecting GlcNAc and the core-fucose roles are antagonistic to each other from the perspective of the effector functions of the Fc glycan. Just like the lack of core-fucose, higher levels of bisecting GlcNAc increase the affinity of IgG for the FcγRs, resulting in an enhanced ADCC [138, 170, 171] (Figure 8).

1.2.1.4 Core-fucosylation

Several studies have reported increases of either core- or terminal fucosylation in many cancers, including ovarian cancer, colorectal cancer, and breast cancer [80, 172, 173]. The increase of fucosylation in ovarian cancer is mediated by the upregulation of the fucosyltransferases, the enzymes encoded by FUT 1-11 genes [31]. Over 90% of the total IgG glycoforms contain fucose attached to the first N-acetylglucosamine of their structures [138, 174]. This is in contrast to other immunoglobulin classes, which are largely not core fucosylated [175]. Core-fucose is thought to be protective from harmful antibody-dependent cellular cytotoxicity (ADCC) activities by acting as the "safety switch" [176]. This means that the majority of the circulating IgG pool has safety switches. The presence of core-fucose reduces the ability of IgG Fc to bind FcγRIIIA and FcγRIIIB, hence the activities of the ADCC are countermanded. Lack of fucose, by contrast, increases the binding affinity for the FcγRIIIA, promoting ADCC activity by about 100-fold [165, 170, 177, 178] (Figure 8).
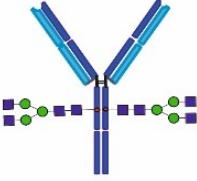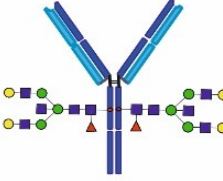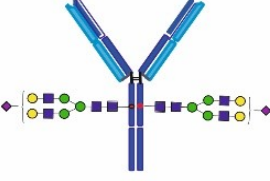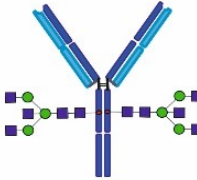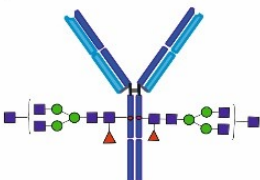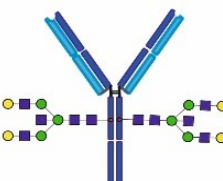
N-glycosylation change	IgG Fc N-glycosylated structure	Effector function
Agalactosylated		Pro-inflammatory
Galactosylated		Anti-inflammatory
Sialylated		Anti-inflammatory
Bisecting GlcNAc		Enhanced ADCC
Core-fucosylated		Decreased ADCC
Not core-fucosylated		Enhanced ADCC

Figure 8. Illustration of the effector functions of the modulated IgG FC glycosylation.

1.3   Ovarian cancer

Ovarian cancer is a group of diseases that forms in the ovary, fallopian tubes, and peritoneum, which manifests as fast-growing abnormal cells that can invade and spread to other parts of the body [8]. Ovarian cancers are classified into three histological types: epithelial tumors (the most dominant), sex cord-stromal tumors, and germ cell tumors [179, 180]. Currently, EOC is sub-classified further into seven histotypes: serous type, mucinous, clear cell, endometrioid, seromucinous, transitional cell (Brenner tumors), and mixed epithelial carcinomas [36, 181]. In the developed world, for instance in Germany, EOC is the most common, accounting for approximately 90% [179, 182]. A similar trend is also reflected in the developing parts of the world where, for example in Kenya, approximately 86%

of OC cases are of EOC type [183]. Consequently, owing to its numerical significance, the present research focuses on epithelial ovarian cancer.

Ovarian cancers are highly lethal and a common cause of death in all gynecological cancers [184], notwithstanding current advances in therapy. The Global Cancer Observatory (GLOBOCAN) reported in 2018 that of the total 8.8 million female cancers globally, an estimated 295,414 cases were of OC. Furthermore, of the total 4.1 million global female cancer deaths, 184,799 mortalities were of OCs [1, 8]. The factors attributed to the high mortalities are late diagnosis and misdiagnosis due to unspecific symptoms in the initial stages of cancer and a lack of effective early diagnostic biomarkers. Hence, the cancer is mostly diagnosed in late stages, when it has already metastasized to distant sites of the body, and the treatment options are greatly diminished. Other contributors to high mortalities or early deaths are resistance to chemotherapy agents as well as infectious agents such as HIV and Hepatitis B virus [3, 4, 6].

The diagnostic patterns of OC are similar in both the developed and the developing parts of the world, whereby the majority of women are diagnosed when cancer has already spread, a situation that portends greatly diminished chances of complete cure and longevity. In developed countries, more than 75% of OC are diagnosed in the advanced stage (FIGO stage III and IV) [185-187], a situation that is replicated in the developing world, where for instance in Kenya, over 73% of OC cases were diagnosed in advanced stage [183]. Ovarian cancer patients diagnosed in the advanced stage have greatly diminished five-year survival due to cancer spread compared to those diagnosed in early stage. For instance, the 2011 United Kingdom data on OC patients showed that only 20% of stage III survived after five years, while for stage IV only 6% survived [188]. In Kenya, 50% of OC patients survived within two years of diagnosis [183]. Unfortunately, less than 30% of OC diagnosed cases are made in stage 1, which accords better five-year survival of about 75% and a 90% cure rate [188-191].

1.3.1   Pathogenesis of epithelial ovarian cancer

The disease process of EOC is well elucidated by the International Federation of Gynecology and Obstetrics (FIGO) through their staging system. For EOC patients in FIGO Stage I, the tumor is limited to the ovary or fallopian tube, while in Stage II the tumor remains confined to the ovary or fallopian tube, but with an extension to the pelvic region. When the tumor spreads to the peritoneum in the abdomen and retroperitoneal lymph nodes, it is classified as Stage III, while in Stage IV, the tumor spread goes beyond the peritoneum cavity to distant parts of the body [192]. Although the disease process of EOC is well understood, its etiology is not yet clear. Several models have been put forward to elucidate the etiology of OC, which include genetic mutations, incessant ovulation, hormonal stimulation, and chronic inflammation.

The model of genetic mutation, which is the most important risk factor of EOC, is anchored on the inheritance of the mutant genes. Inheritance of the germline mutations in the Breast cancer 1 and 2 (BRCA1 and BRCA2) tumor suppressor genes accounts for the majority of hereditary OCs, and 10 - 15 % of all ovarian cancers [191]. The carriers of these genetic mutations have lifetime probabilities of developing ovarian cancer at 30-60% and 15-30% for BRCA1 and BRCA2, respectively [191, 193].

The model of incessant ovulation is elucidated as the risk index of developing EOC arising from the number of times that an individual gets the ovulatory cycles in their lifetime [194, 195]. Excessive ovulations subject the epithelium to constant damage, thereby increasing the chances of spontaneous mutations of DNA repair and inactivation of tumor-suppressor genes leading to EOC [195, 196]. This model is positioned by the biological or therapeutic risk factors that facilitate the increased number of lifetime ovulation cycles. They include early onset of menarche, delayed menopause, nulliparity, fertility medication, and hormone therapy in menopausal women [197].

The model of hormonal stimulation is based on the elevations of the gonadotropin levels acting in concert with estrogen. This model explains the key events that lead to the development of ovarian cancer. It involves entrapment of surface epithelium in inclusion cysts, leading to its stimulation by estrogen or estrogen precursors, especially in the presence of persistently high levels of the gonadotropins (luteinizing hormone and follicle-stimulating hormone) [196, 198, 199]. Finally, chronic inflammation has also been advanced as one way of OC pathogenesis, whereby the toxic oxidants produced during this process for the purpose of eliminating pathogens may also cause direct damage to DNA, leading to carcinogenesis [200]. Furthermore, it is also linked to increased cell division, which enhances formation of errors of replication that arise from DNA repair and consequently an increased risk of mutagenesis. Exposure of ovaries to pelvic contaminants, such as ectopic implantation of uterine lining and talc, may also result in inflammation leading to EOC [196].

1.3.2   Preventive factors for epithelial ovarian cancer

The factors thought to protect women from developing OC are in majority those that cause unovulation. These include pregnancy, parity number of live births, prolonged lactation, use of oral contraceptives, and tubal ligation or hysterectomy [194]. During pregnancy, there is high production of estrogen and progesterone hormones, which suppress the levels of LH and FSH to prevent ovulation. Similarly, during lactation, there is high production of prolactin, which has a negative feedback on the production of estrogen and LH, and consequently, ovulation is suppressed [182]. Similarly, the use of oral contraceptives stabilizes the levels of estrogen and progesterone, which inhibits the secretion of gonadotropins curtailing ovulation [199]. The fact that these factors are

protective from EOC is suggestive of a common pathway of ovarian carcinogenesis through ovulation or gonadotropin hormones.

1.3.3   Diagnosis of epithelial ovarian cancer

Early diagnosis of OC patients offers the most practical and promising approach to reducing mortalities and morbidity associated with the late-stage diagnosis of OC. The most widely validated routine serum biomarker currently used for the diagnosis of OC is the cancer antigen CA125; however, its analytical usefulness is undermined by its low specificity and sensitivity [107, 191]. CA125 levels are found to be normal in approximately 20% of OC cases, but are also found to be elevated in conditions such as benign ovarian diseases, liver diseases, menstruation and many other states [201]. Owing to these inadequacies, CA125 is used in combination with other markers such as clinical pelvic examination and transvaginal ultrasound or imaging to bolster its diagnostic performance [189, 202]. But even with the inclusion of CA125 and transvaginal ultrasound into the screening programs, there has not been any significant change in the trajectory of mortalities [202, 203]. The recently approved biomarker for monitoring OC relapse and progression, human epididymis 4 (HE4), has a better performance in detecting OC in premenopausal women compared to CA125 [64]. When the two tests HE4 and CA125 are used as a combination to detect EOC in the early stage, the problem of specificity has persisted, as shown by risk of ovarian malignancy algorithm (ROMA) score (specificity, 75 – 84%), which is based on the two biomarkers. A detailed discussion on this, and other approved clinical biomarkers for OC detection, is addressed in our recent review [8]. Therefore, the need for new effective biomarkers cannot be gainsaid.

1.3.4   Treatment of EOC and biomarkers for monitoring treatment efficacy

Early-stage diagnosis of OC, effective monitoring strategies and optimal treatment are important factors that contribute to the reduction in OC mortalities and increased patient survival. The gold-standard treatment of OC consists of surgical debulking followed by chemotherapy [204, 205]. The first-line chemotherapy strategy for OC is a combination of platinum and paclitaxel administered in six cycles, with each cycle given every 3 weeks after the previous one [187]. However, new transformative treatment options are emerging, such as targeted therapy and immunotherapy [206], though some are still within the trial phase. Surgical intervention usually aims for the complete removal of the tumor; however, the majority of patients still retain some small residual tumors in the peritoneum, which are difficult to examine physically or by imaging [13]. Approximately 30% of EOC patients may not respond to platinum and paclitaxel combination therapy, while of the 70% that respond, it is not possible to detect residual cancers in about one-half of them using the available imaging and biomarker techniques after five months of treatment [191].

Although CA125 is largely used in monitoring the stability or progression of OC, its suitability for monitoring treatment response has not been ascertained regarding the interpretation of consecutive CA125 concentrations in making the correct judgment on the state of tumor progression [13]. This is because the criteria that have been proposed so far have failed to eliminate the false positives that indicate tumor progression [207]. Hence, the need to not only discover effective biomarkers for early detection of OC but also for monitoring the response of the patients to the chemotherapy agents is of fundamental importance.

1.3.5   Potential glycomic-based biomarkers for ovarian cancer

Many of the biomarkers currently used in cancer diagnosis and monitoring are largely glycosylated proteins, but their measurement is mostly based on specific immunochemical detection of the protein moiety that is part of the glycoprotein as opposed to the glycan [208]. Some of the currently approved cancer biomarkers derived from the glycoproteins include CA125 for OC, CA15-3 for breast cancer, CA19-9 for pancreatic cancer, carcinoembryonic antigen (CEA) for colorectal cancer, and prostate-specific antigen (PSA) for prostate cancer. Various research groups have reported aberrant *N*-glycosylation at the levels of total glycome, immunoglobulin, and acute-phase proteins that shows great potential for future new biomarkers of various malignancies.

At the level of total *N*-glycome, several research groups have reported transformation of *N*-glycans that included downregulation of high-mannose, asialylation hybrid, and mono- and bi-antennary structures [64, 78, 80, 107, 116], while sialylated, fucosylated, and tri- and tetra-antennary *N*-glycan structures were reportedly upregulated [64, 79, 80, 107, 116, 126]. Cancer-specific aberrations of the total *N*-glycome include the increase of sialylated and fucosylated glycan structures in breast cancer [209], while in gastric cancer, a decrease of high-mannose, bigalactosylated biantennary with an increase of the agalactosylated biantennary glycans was reported [68, 129]. Saldova and colleagues also reported an increase in core-fucosylated biantennary and α2-3-linked sialic acids in prostate cancer patients [81]. Our research group previously described 11 *N*-glycan signatures that were altered in EOC, and were able to discriminate primary EOC patients from healthy and BOD women, with better specificity compared to CA125. They comprised four high-mannose (*m/z* 1579.8, 1783.9, 1988.0, and 2192.1) and seven multiantennary complex-type *N*-glycans characterized by high sialylation and fucosylation (*m/z* 3776.9, 3951.0, 4226.1, 4400.2, 4587.3, 4761.4, and 4935.5) [64, 107]. A score generated from the 11 *N*-glycans named "GLYCOV" was able to differentiate primary EOC patients from the healthy or BOD controls and was also able to assign the EOC patients in their respective tumor stages with higher accuracy compared to the CA125 [107].

A few studies have also reported on the potential use of altered glycans as biomarkers for predicting resistance and response to chemotherapy. A previous study identified six *N*-glycans in tissues of OC patients with the potential for use as markers of resistance to chemotherapy. The six *N*-glycans displayed features of tetraantennary, sialylation and/or fucosylation and bisecting structures [112]. Additionally, lewis type biantennary, triantennary tri-sialylated, and lewis type triantennary structures were also isolated from the serum of OC patients and were found to be useful in predicting response to chemotherapy in OC patients [210]. In another study involving breast cancer patients, alterations in the abundance of high-mannose, core-fucose, and galactosylation were reported following chemotherapy administration [211], suggesting their possible utility as markers of drug efficacy.

Investigation of malignant glycosylation changes at the level of immunoglobulin has seen more interest directed at the predominant antibody of the secondary immune response, the IgG. Various independent research groups have reported increased agalactosylation and fucosylation, with a decrease in galactosylation among OC patients [82, 87, 212]. Ruhaak et al. found IgG to be the most accurate (91%) in discriminating EOC patients from the healthy controls, followed by IgA (85%), and the least accurate was IgM (76%) [83]. Finally, an increase in sialylation, core-fucosylation, lewis x antigen (Le[x]), and sialy Lewis [x] (sle[x]) antigen were reported in acute-phase proteins, which included the haptoglobin, α-1-antichymotrypsin, α-1-antitrypsin, α-1-acid glycoprotein, hemopexin, and C1 esterase inhibitor in OC patients [86, 87, 213]. Saldova et al. reported that a combination of sle[x] and agalactosyl biantennary glycans significantly improved the discrimination of ovarian cancer from BOD subjects [87].

1.4 Approaches for *N*-glycan and IgG glycopeptide analysis

There are three approaches with which glycans can be analyzed, and they include at the level of released glycans, glycopeptides, or at the intact glycoprotein level [8]. The choice of a particular strategy is dependent on factors such as the nature and purity of the sample, level of technology available, expert knowledge, and the nature of output information envisaged [8]. In this study, we used the released *N*-glycans analysis and the glycopeptide analysis.

1.4.1 Analysis at the released *N*-glycan level

Analysis of the released *N*-glycans is the most preferred method for studies involving compositional and structural characterization of the *N*-glycome. The advantage of this method is its lower complexity output compared to either the glycopeptide or the intact glycoprotein samples [143]. The release of *N*-Glycans from the glycoproteins is carried out using either chemicals or the enzyme approach. One of the guiding factors in choosing either of the two *N*-Glycan release strategies is

whether the sample will be needed for further analyses. For instance, the release of *N*-Glycans by chemicals causes partial degradation of the sample, making its further use for analyses unsustainable [214]. Once the free *N*-glycans have been released, they are taken through a purification process to remove the peptide, salts, and excess reagents. This is followed by derivatization with the singular purpose of stabilizing the negatively charged sialic acids to prevent their loss during measurement by MALDI-TOF mass spectrometry.

1.4.1.1 Enzymatic release

Different types of enzymes are available for releasing *N*-glycans from the glycoprotein. The information that is hoped to be generated from the analysis of the released *N*-glycans determines the choice of one enzyme from another. The most commonly used enzymes are the *N*-glycosidases, such as peptide-N-glycosidase A and F (PNGase A and PNGase F), and endoglycosidase H (Endo H). PNGase F was first isolated from *Flavobacterium meningosepticum* in 1984 by Plummer et al. [215] and is currently used in a recombinant form, as expressed in *E.coli*. Both PNGase F and PNGase A are used when the release of the complex-type, hybrid-type, and high-mannose *N*-glycans is required. However, PNGase F is the most frequently used, unless the core of the glycan is α1,3-fucosylated, which is only released by PNGase A. However, when the selective release of the hybrid and high-mannose is required, Endo H is the one of choice [214]. The mode of action for PNGase F is by cleaving all the *N*-glycans by the internal glycosidic bond of the asparagine and the terminal GlcNAc residue, leaving aspartic acid in place of the asparagine at the *N*-linked site of the protein. Conversely, Endo H, cleaves the glycans between the two innermost GlcNAc residues in the chitobiose core of the high-mannose or hybrid *N*-glycans, which leaves aspartic acid in place of the asparagine at the N-linked site of the protein [216, 217]. The advantage of using the enzymatic release is that the glycans released are specifically *N*-glycans in exclusion of O-glycans. Furthermore, the resultant free *N*-glycans and the peptide can still be utilized in other analyses as their integrity is usually preserved during separation, unlike when using chemicals for glycan release.
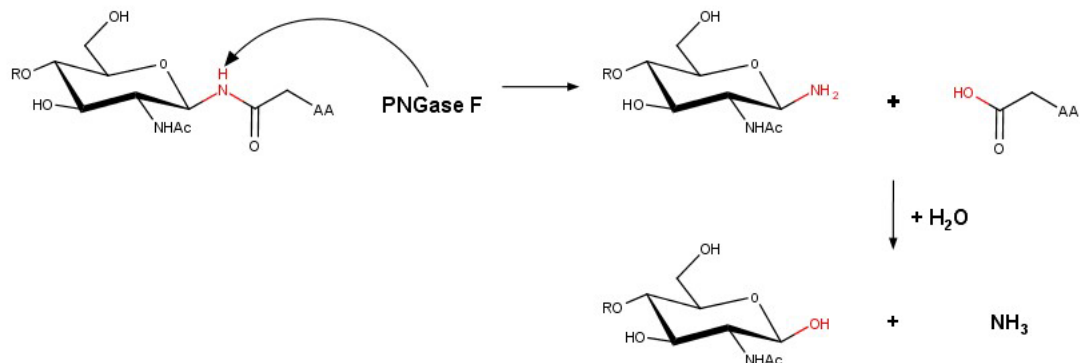
Figure 9. Deglycosylation of the glycoprotein by PNGase F [218].

1.4.1.2 Chemical release

Glycans can also be released from the protein by the hydrazinolysis approach, where both *N*- and O-glycans are released simultaneously under controlled conditions. However, O-glycans can also be released alone through a method known as β-elimination, whereby the glycoprotein is treated with alkaline borohydride under controlled conditions [29]. The shortcoming of glycan release by hydrazinolysis is that it cleaves all the peptide bonds causing damage to the peptide, hence preventing its further experimental utility. Furthermore, the glycans may also exhibit partial degradation [214].

1.4.1.3 Purification methods

The purpose of this step is to isolate and desalt glycans by removing peptides and impurities such as excess reagents and salts [8]. Several strategies such as solid-phase extraction (SPE) and size-exclusion chromatography (SEC) are available for purifying the free *N*-glycans and the glycopeptides. The commonly used SPE methods are hydrophilic graphitized carbon columns, hydrophobic C18 cartridges and hydrophilic interaction liquid chromatography (HILIC). The working principle of C18 is based on the separation of the hydrophilic *N*-glycans by eluting them through the hydrophobic C18 cartridge, while the peptides are adsorbed on the C18 stationary phase. By contrast, the hydrophilic graphitized carbon columns, *N*-glycans adsorb onto the carbon columns, while the salts and excess reagents are removed in the acidified water flow-through [219]. The HILIC method consists of a polar stationary phase and a less polar mobile phase known as the "normal- or straight- phase" [220]. The stationary phase may comprise silica gel and polymers bearing functional groups, but sepharose and microcrystalline may also be used as stationary phases. The advantage of the HILIC method is that it enriches the tryptic glycopeptides by increasing their ionization efficiency and detectability [221].

1.4.1.4 Derivatization

Derivitization of the released *N*-glycans aims at stabilizing the negatively charged sialic acids to prevent their loss during mass spectrometric measurement. The most frequently used methods of derivatization are permethylation and peracetylation [8]. Both permethylation and peracetylation act by decreasing the intermolecular hydrogen bonding, hence converting the hydrophilic glycans into hydrophobic molecules. To do so, free hydroxyl groups are irreversibly converted into methoxy groups using iodomethane [222, 223]. Permethylation of *N*-glycans improves ionization and sample volatility, resulting in high quality spectra [209]. Another way to derivatize *N*-glycans is at the reducing end, using fluorescent tags in order to increase the sensitivity of the measurements and also the limit of detection (LOD) [222].

1.4.2   Analysis at the glycopeptide level

Analysis at the level of glycopeptides is the most preferred for immunoglobulin studies. Its bottom-up approach enables the characterization of the glycans by site-specific heterogeneity and their linkage sites on the protein [8, 224]. Analysis of the glycopeptides involves sequential steps that include purification, enzymatic digestion, derivatization, and finally measurement. Of the immunoglobulin classes, IgG, the major immunoglobulin of humoral immunity, is the most studied in glycomic-based research [8]. Sample preparation for IgG glycopeptide analysis involves the use of Protein A alone or together with protein G in affinity chromatography, followed by tryptic enzymatic digestion. Whereas Protein A is derived from *Staphylococcus aureus,* protein G is obtained from *streptococcus* bacteria groups C and G [225]. Both Protein A and Protein G are commercially available in the immobilized form of sepharose beads and magnetic beads. The difference between the two is that Protein A only binds IgG_1, IgG_2, and IgG_4 to the exclusion of IgG_3, while Protein G captures all the four IgG subclasses including IgG_3 [57, 82]. The separation advantage of the two proteins is achieved by sequential capturing of IgG_1, IgG_2, and IgG_4 using Protein A, followed by the capture of IgG_3 by Protein G in the flow-through. Affinity-separated IgG products are then subjected to a proteolytic digestion to form the glycopeptides.

The types of proteolytic enzymes available for the digestion of the glycoprotein to generate the glycopeptides include specific endopeptidases such as trypsin, which is the most commonly used, and others such as chymotrypsin, lysyl endopeptidase (Lys-C), endoproteinase Gluc-8 (V8 protease), and pronase [226]. Trypsin acts by cleaving the peptides at the carboxyl sides of arginine and the lysine, but not when the succeeding amino acid is proline [227]. Trypsin is well suited for the detection of glycosylation sites because proteolytic cleavage leaves the larger peptide attached to the oligosaccharide and results in glycopeptides that ionize well on both the MALDI (matrix-assisted laser desorption/ionization mass spectrometry) and electrospray ionization (ESI) [226]. The resultant glycopeptides containing Asn-297 are of distinct sequences, enabling discrimination of the different IgG subclasses (Table 1). The peptide sequence of the IgG_3 has two allotypes that vary in the amino acid at position Asn296. Analysis of IgG_3 among Caucasians showed allotype G3m(b*) with a phenylalanine in position 296, while the Asian ethnicity revealed a tyrosine residue in position 296 [228]. Consequently, the Caucasian IgG_3 allotype EEQFNSTFR is identical to IgG_2. However, the allotype EEQYNSTFR is identical to IgG_4, and it is the predominant one in Asian populations, including African populations [139, 229]. After tryptic digestion, the preparation is cleaned up using hydrophilic interaction liquid chromatography (HILIC), in readiness for measurement on an MS platform.

Table 1. Tryptic IgG glycosylation analysis

IgG subclass	Separation protein	Glycopeptide sequence	Swis-prot	Registered m/z-value $[M + 2H]^{2+}$
IgG_1	A	EEQYN-STYR	P01857	595.78
IgG_2	A	EEQFN-STFR	P01859	579.79
IgG_3^a	G	EEQYN-STFR	Q8N4Y9	587.75
IgG_3^b	G	EEQFN-STFR	Q86TTz	579.79
IgG_4	A	EEQFN-STYR	P01861	587.75

IgG_1, IgG_2 and IgG_4 are captured by protein A, while protein G captures IgG_3, which exists in the form of two allotypes of peptide sequences: IgG_3^a and IgG_3^b

1.4.3   Analysis at the level of intact protein

In this approach, analysis of the target proteins is made directly or indirectly without purification or isolation by immuno-purification methods [53,57–59]. This is the method commonly used in routine diagnostic clinical chemistry laboratories, research laboratories, and by pharmaceutical companies in product quality controls.

1.4.4   Mass spectrometric techniques for glycome analysis

Mass spectrometry (MS) relies on the ionization of biomolecules in the sample and identifies substances by sorting them based on their mass-to-charge ratio (m/z) [230]. Generally, there are three main parts that constitute the mass spectrometer. They include the ion source (electrospray ionizer for ESI or laser for MALDI), whose task is to ionize the analyte, the mass analyzer, which separates the ions according to their mass-to-charge ratio, and the detector, which outputs the resultant ions as relative intensities that form the mass spectrum [230-232]. Mass spectrometry platforms are normally used in hyphenated forms with different types of spectrometers, which include time of flight (TOF), orbitrap, ion trap (IT), quadrupole (Q), and Fourier transform ion cyclotron resonance (FTICR). The use of a particular spectrometer, as opposed to another, is a function of the nature of work, the mass range required, laser power and the accuracy levels of the analyzer [230, 233]. The most commonly used mass spectrometry platforms are matrix-assisted laser desorption/ionization mass spectrometry-time of flight (MALDI-TOF-MS), surface-enhanced laser desorption and ionization mass spectrometry (SELDI-TOF-MS), and electrospray ionization mass spectrometry (ESI).

MALDI-TOF-MS is the most commonly used MS platform in glycobiology. It is a soft ionization technique that utilizes a matrix mixed with the sample to identify substances by sorting them on the basis of their m/z. The preferential usage of the MALDI-TOF-MS in glycobiology is founded on its good sensitivity and reproducibility, which enables seamless profiling and characterization of aberrant protein glycosylation resulting from disease or malignant processes. The MALDI-TOF-MS measurement begins by directing a laser beam to the co-crystals formed by mixing equal volumes of

the matrix and sample. The matrix absorbs the energy so that the in-source analyte fragmentation of protein/peptide is minimized [8, 232], which leads to an explosive transition from solid crystal to atomized plume, causing the sample to ionize. In the TOF tube, the atomized sample ions move quickly through the region of the free-flight electric field (Figure 10). Molecular separation is then achieved when energy gain during acceleration drops with an increase in mass, meaning smaller ions fly faster because they are lighter than the bigger ions [232]. Finally, a mass spectrum is then plotted from the mass-to-charge ratio (*m/z*) computed from the period of time between laser ionization and when it is detected at the opposite end. The most commonly used matrix in glycan analysis is 2,5-dihydroxybenzoic acid (DHB), whereas 4-chloro-α-cyanocinnamic acid is mostly used in measurements of the glycopeptides.

Sialic acid moieties are hence labile during measurement by the MALDI-TOF-MS, and some are lost due to in-source and post-source fragmentation producing broad metastable peaks [234]. Moreover, sialylated glycans tend to show different adducts of salts, which interferes with measurements due to a lot of noise being generated for a single glycan composition [234]. It is for this reason that glycans are derivatized to eradicate post-source fragmentation [80, 235]. The advantage of MALDI-TOF-MS over other methods is that it has a high tolerability of salts; nevertheless, thorough desalting and clean-up of excess reagents is necessary for good reproducibility. Other advantages include the ease of studying the structural profile, and the ability to measure a broad spectrum of ranges.

Another platform that is frequently used is the SELDI-TOF-MS platform, which is a high-throughput platform that analyses the biomolecules using a chromatographic protein chip coupled to the TOF separator and MS detector. It selectively binds the glycoprotein or glycan by fractionating and isolating them based on their charge and hydrophobicity, and then analyzes using laser desorption ionization TOF-MS [236, 237]. The advantage of SELDI-TOF-MS is that it only requires a minuscule amount of the sample, and has a high turnaround time.

The ESI operates either by direct infusion of the sample solution or by connecting to liquid chromatography. A voltage of positive or negative polarity forces the liquid sample containing the analytes to come through the needle as nebulized particles, which readily evaporate to form analyte ions [180, 202, 238].
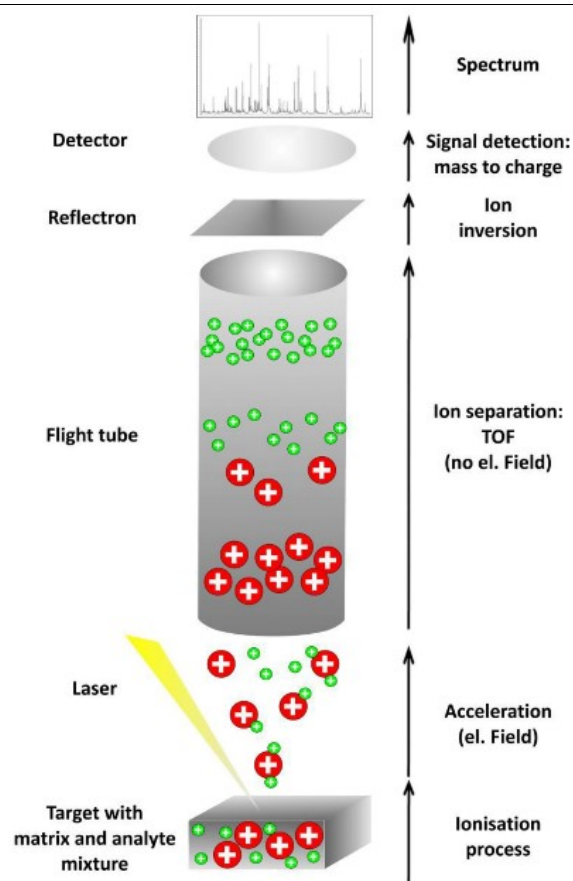
Figure 10. Illustration of the method of MALDI-TOF-MS measurement [218].

1.5   Study aims and rationale

The high mortalities associated with OC are largely ascribed to late-stage definitive diagnosis that takes place at a time when cancer has extensively spread and the treatment options are therefore greatly depleted. Late diagnosis is in the majority a consequence of a lack of effective early diagnostic biomarkers and the unspecific OC symptoms that often lead to the misdiagnosis of patients, yet patient survival is strongly linked to the tumor stage at which the diagnosis is made. The resistance to chemotherapy agents by the cancer cells as well as the lack of an effective biomarker for monitoring the response or the efficacy of anti-cancer agents also contribute to the high mortalities of EOC patients. The current routine biomarker for EOC, CA125, lacks the requisite specificity and sensitivity for early diagnosis of EOC as well as in monitoring patients' response to chemotherapy. It is against the backdrop of the weaknesses inherent in CA125 regarding early OC diagnosis and in monitoring the efficacy of anti-cancer treatment agents that new effective biomarkers are urgently needed to improve patient survival and reduce mortalities. Routine screening of underlying infectious agents such as HIV, HBV and HCV in EOC patients scheduled for chemotherapy is one of the

approaches that may help contribute to enhanced patient survival through the informed application of optimal treatment strategies.

Glycomics, the systematic study of glycans expressed in biological systems, has opened new leads in the arena of biomarker research. Modulations of *N*-glycosylation are increasingly being observed in various diseases and cancers, including ovarian cancer. Independent research groups, including our own, have previously reported malignant transformation of *N*-glycosylation and IgG Fc glycosylation in many cancer types including OC, which showed great potential for new biomarkers. Previous findings of our research group showed alterations in serum total *N*-glycosylation, where four high-mannose *N*-glycans were significantly downregulated, while 7 complex-type structures with conformational features of sialylation and fucosylation were upregulated in Caucasian cohorts of EOC patients compared to their control counterparts (BOD and healthy women). Furthermore, analysis of IgG_1, IgG_2 and IgG_3 glycopeptides revealed an increase in the degree of agalactosylation, with a corresponding decrease in galactosylation among the EOC patients compared to the controls. Interestingly, reports from previous studies indicated inter-ethnic variations in *N*-glycosylation, although the available research data on glycome-based biomarkers for OC have only focussed on cohorts of patients of Caucasian or Asian origin, while to my knowledge, no data is available to date with samples from African ethnicity. In addition, there is no research data available on glycome-based biomarkers for monitoring clinical response to chemotherapy using EOC cohorts of African ethnicity.

Thus, the first objective of this study addressed the identification of *N*-glycan-based biomarkers for effective monitoring of EOC patients' response to chemotherapy in a cohort of African ethnicity. Further, in the second objective, for the first time, an analysis was carried out on total *N*-glycome to identify markers for the diagnosis of EOC in the African cohort. Additionally, an analysis was carried out on the IgG Fc glycosylation in EOC patients of African and European cohorts to identify the potential common IgG_1 and IgG_2 glycosylation signatures that could robustly discharge an improved diagnosis of EOC.

Other than the early diagnosis of OC and effective chemotherapy monitoring strategies, optimal management of OC patients is equally fundamental in reducing mortalities and improving patient survival. Underlying infectious diseases such as HIV and HBV have also been linked with early deaths of OC patients when optimal treatment strategies that simultaneously curtail viral reactivation/replication and suppress tumor cells are not applied. It was reported in previous studies that patients of gynecological cancers infected with HIV and/or with HBV had reduced survival when they were subjected to chemotherapy without first establishing their infective status for guidance on the most beneficial treatment strategies, or where adherence to antiretroviral treatment alongside chemotherapy was not followed. Currently, there is no data available on the prevalence rates of HIV,

HBV, and HCV in Kenyan OC patients, although the prevalence of HIV and HBV in the general population is in the range of high endemicity. Moreover, the question of establishing the seropositivity status of HBV, HIV and HCV among EOC patients scheduled for chemotherapy was not routinely carried out in the study setting at the time of this study. Consequently, the third objective of this thesis was to determine for the first time the burden of HIV, HBV, and HCV in Kenyan EOC patients, to give direction as to whether it was relevant to carry out routine testing of EOC patients scheduled for chemotherapy. By establishing the infective status of EOC patients before chemotherapy, patients can benefit from the selection of optimal management strategies aimed towards the overall desire to reduce early deaths and increase of the quality of life.

In summary, the current study sought to identify new effective glycomic-based biomarkers that will complement CA125 to improve early diagnosis of EOC as well to monitor the response to chemotherapy towards reducing mortalities and improving the survival of EOC patients. The choice of BOD subjects as the controls for these experiments was based on the fact that BOD is one of the most prominent causes of EOC misdiagnosis.

2.0 Materials and methods

2.1   Study setting

Recruitment of the African cohort was done in multi-centers within Nairobi, Kenya as per the ethical vote obtained from the Kenyatta National Hospital / University of Nairobi ethics review committee (KNH/UON-ERC), reference no P701/12/2017. The recruitment centers were Kenyatta National Hospital (KNH), St. Mary's Hospital Langata, and Texas Cancer Centre. The ethical commission of the Charité-Universitätsmedizin Berlin granted the use of the samples for laboratory analysis (approval number EA4/071/19). A total of 136 samples were collected. The European samples (n= 57) were obtained from the Department of Gynecology of the Charité-Universitätsmedizin Berlin, as per the Charité Institutional Ethics Committee (approval number EA4/073/06).   Written consent was obtained for all the study participants.

2.2   Study cohort

In this study, African adult women of 18 years and above attending the Gynecology clinics and wards within the study setting were recruited as participants and their venule blood samples were collected after giving their written informed consent. They comprised 86 women with histologically confirmed EOC, with their age-matched (±5 years) control counterparts comprising 50 women with BOD. Of the 86 EOC patients, 19 patients were primary, whereas 67 had received a varying number of chemotherapy sessions. Recruitment and collection of blood samples from the African cohort commenced in April 2018 and ended in April 2020. For the European (Caucasian) cohort, they comprised 35 women with confirmed EOC and their age-matched (±5 years) counterparts of 22 BOD subjects. The inclusion and exclusion criteria for the study subjects were: women of age 18 years and above, not pregnant, having a definitive diagnosis of ovarian cancer or benign ovarian disease, being free from any other cancers and not suffering from liver or kidney insufficiency. These samples were used to address three different research questions in this thesis, as explained in the rest of this section.

To analyze the prevalence of HBV, HIV, and HCV in Kenyan OC patients, all 86 EOC and 50 BOD Kenyan subjects were used. At this stage, all the serum samples that tested positive for HBV, HIV, and HCV were excluded from the analysis of the remaining research questions. In the second study question, which involved an evaluation of the total *N*-glycome for the diagnostic biomarker of OC and monitoring chemotherapy response, a total of 53 EOC and 46 BOD serum samples were analyzed after excluding the HBV, HIV, and HCV seropositive samples. Of the 53 EOC patients, 19 were primary EOC patients (pre-treatment group), while 34 were chemotherapy responders who had already undergone a varying number of chemotherapy sessions (1 - 6 cycles). The same African cohorts of 19

primary EOC and 46 BOD patients (minus one EOC sample whose volume was insufficient) were carried over to address the question of potential common IgG_1/IgG_2 glycopeptide biomarkers for multi-ethnic diagnosis of EOC.
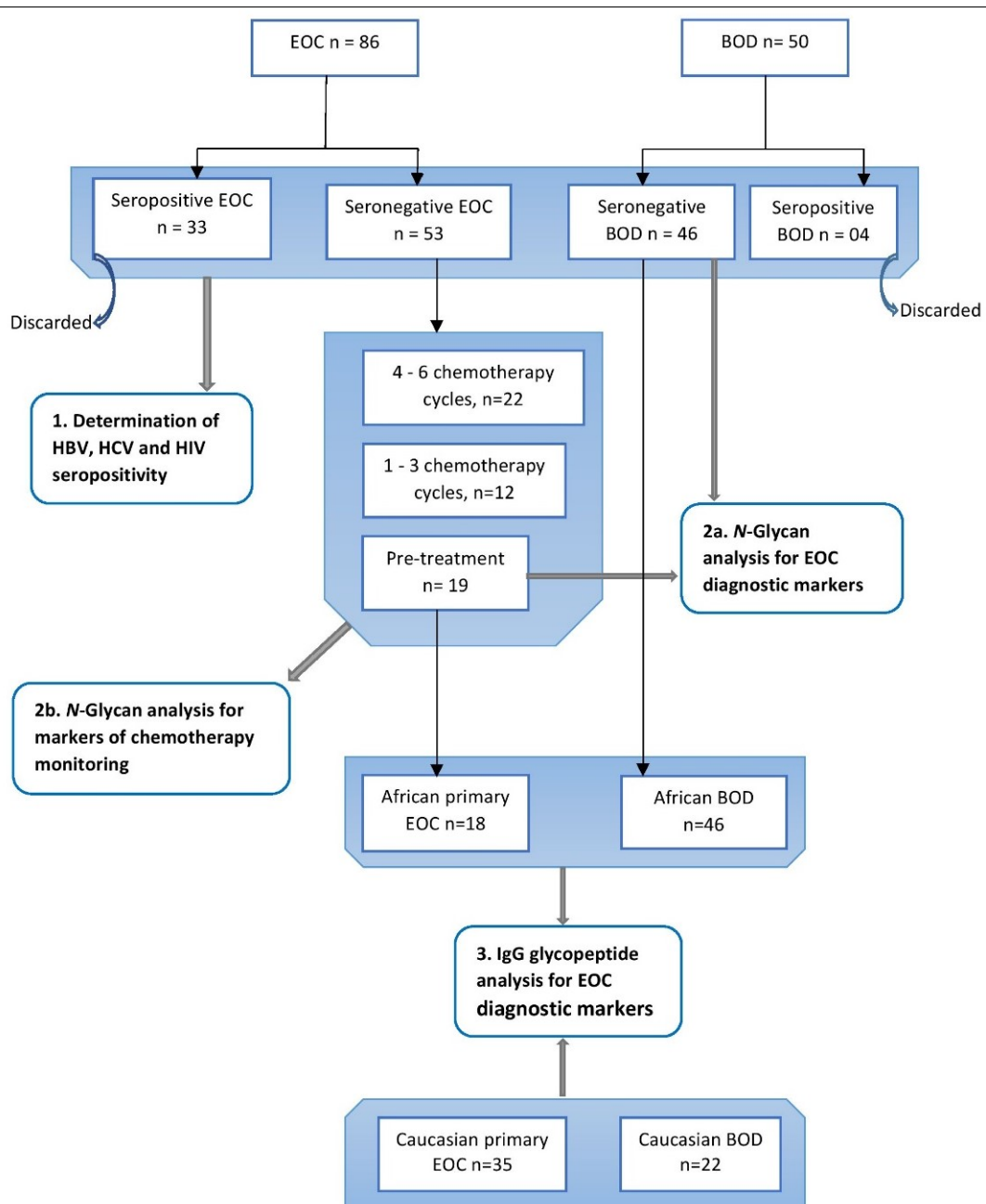
Figure 11. Flow chart diagram of study cohorts and activities

2.3   Specimen collection and handling

Blood samples were collected in 5 ml vacutainers with a serum clot activator (Becton, Dickinson GmbH, Heidelberg, Germany). Samples were then left to stand in a test tube holder for around 30 minutes to 2 hours at room temperature to allow clot retraction, followed by centrifugation at 1,200g for 15 minutes to separate serum from the cells. Approximately 2 ml of serum was aliquoted into the

Eppendorf tubes and stored at -80 ^{0}C for the entire sample collection period before shipment to Prof. Dr. Véronique's laboratory at the Charité-Universitätsmedizin Berlin, Germany. Shipment was done in accordance with the International Air Transport Association (IATA) protocol on shipment of biological samples and also after obtaining the necessary approvals in Kenya. These included; KNH/UoN-ERC Ref no., KNH-ERC/shipment/40, Ministry of Health Kenya, letter of no objection Ref. No. MOH/F/HRD/01/VOL.11 and the Kenya Pharmacy and Poisons board export permit Ref. no. CD2021000PPB321J0002550623.

2.4   Instruments and reagents

All the materials, that were used in this study, which included instruments, reagents and consumables, are listed in Table 2 and Table 3.

Table 2. Instruments and consumables

Instrument	Manufacturer	Town, Country
Cobas e 801 immunoassay system	Roche Diagnostics GmbH	Penzberg, Germany
MALDI-TOF-MS equipped with a smart beam laser	Bruker Daltonics	Bremen, Germany
Flexanalysis	Bruker Daltonics	Bremen, Germany
Centrifugal evaporator	Uniequip	Martinsried, Germany
Thermo mixer	Eppendorf Netheier	Hamburg, Germany
Ground steel target	Bruker Daltonics	Bremen, Germany
Vortex mixer	Cole Palmer	St. Neots, United Kingdom
Centrifuge	Eppendorf AG.	Hamburg, Germany
Mini centrifuge	VWR	Atlanta, GA, U.S.A
Incubator	Eppendorf AG.	Hamburg, Germany
Consumables		
C18 cartridges	Alltech,	Deerfield, IL USA
Graphitized carbon columns	Alltech,	Deerfield, IL USA
Vacutainers	Becton, Dickinson GmbH,	Heidelberg, Germany
Eppendorf tubes	Eppendorf AG.	Hamburg, Germany
10 µl filter tips (for affinity chromatography)	Greiner Bio-One	Kremsmünster, Austria
100% cotton wool pads	Greiner Bio-One	Kremsmünster, Austria
Pipette tips	Sarstedt AG	Numbrecht, Germany

Table 3. Reagents

Type of Analysis	Reagent	Manufacturer	Town, Country
CA125	Elecsys CA125 II	Roche diagnostics GmbH	Penzberg, Germany

HBV	Elecsys HBsAg II	Roche diagnostics GmbH	Penzberg, Germany
HCV	Elecsys anti-HCV II	Roche diagnostics GmbH	Penzberg, Germany
HIV	Elecsys HIV Duo (HIV-Ag and anti-HIV)	Roche diagnostics GmbH	Penzberg, Germany
IgG	Milli-Q water	Merck	Darmstadt, Germany
	Protein A sepharose beads	GE Healthcare	Eindhoven, The Netherlands
	Phosphate buffer	Honeywell Fluka	Steinheim, Germany
	Formic acid	Honeywell Fluka	Steinheim, Germany
	Ammonium bicarbonate	Merck	Darmstadt, Germany
	Trypsin	Promega	Madison, WI, United States
	4-chloro-α-cyanocinnamic acid (CICCA)	Sigma-Aldrich	Steinheim, Germany
	peptide calibration standard II	Bruker Daltonics	Bremen, Germany
N-glycan	PNGase F	Promega	Wisconsin, USA
	Acetonitrile (ACN)	VWR chemicals	France
	sodium hydroxide (NaOH)	Merck Millipore	Darmstadt, Germany
	Dimethylsulfoxide (DMSO)	Applichem	Germany
	2,5-dihyroxybenzoic acid (sDHB)	Sigma-Aldrich	Steinheim, Germany
	Peptide calibration standard II	Bruker Daltonics	Bremen, Germany
	Dithioerythrytol (DTE)	Sigma-Aldrich	Steinheim, Germany
	Iodoacetamide	Sigma-Aldrich	Steinheim, Germany
	graphitized carbon columns	Alltech	Deerfield, IL
	Chloroform	Merck Millipore	Darmstadt, Germany
	Ethanol	Merck Millipore	Darmstadt, Germany
	Trifluoroacetic acid	Merck Millipore	Darmstadt, Germany

2.5   Analytical methods

2.5.1   *N*-Glycan analysis

N-Glycans were enzymatically released from serum glycoproteins, then purified using the C18 cartridges and porous graphitized carbon columns, and finally permethylated before MALDI-TOF-MS measurements.

2.5.1.1 *N*-Glycan release

N-Glycans from total serum glycoproteins were released and purified as described [209]. Briefly, 10 μl of serum was diluted in 25 μl of 200 mM phosphate buffer (pH 6.5). The preparation was then reduced by mixing with 2.5 μl of 200 mM dithioerythrytol (DTE) and incubated at 60^0C on a shaker for 45 minutes. The purpose of DTE is to denature the glycoprotein by disrupting the disulfide bridges

to unfold it to allow the enzyme access to the glycosylation sites. The preparation was then alkylated in 10 μl of 200 mM iodoacetamide for one hour at room temperature in darkness. The purpose of alkylation is to prevent the reformation of disulfide bonds by blocking the sulfhydryl groups, hence bringing stability to the unfolded glycoproteins. Subsequently, the reaction was stopped by the addition of the excess DTE. *N*-Glycans were then enzymatically released from serum glycoproteins after adding 5 μl of 100 mU PNGase F, 2.5 μl of 200 mM DTE, 300 μl of Milli-Q water, and 100 μl of 200 mM phosphate buffer by incubating for 16 hours at 37^0C.

2.5.1.2 Purification and enrichment of *N*-glycans

The released *N*-glycans were then purified by reverse-phase chromatography using the C18 cartridges to remove peptides from the glycan-containing fraction. This was followed by desalting the free *N*-glycans and removing the excess reagents by porous graphitized carbograph columns. The protocol was as follows: C18 and the porous graphitized carbon columns were conditioned simultaneously using 600 μl (carbograph columns) or 400 μl (for C18) of 80% ACN + 0.1% trifluoroacetic acid (TFA). This was repeated twice. This was followed by an equilibration step consisting of either 3 x 600 μl (carbon) or 400 μl (C18) of Milli-Q water + 0.1% TFA. Afterward, the C18 cartridge was placed on top of the carbograph column. The free *N*-glycan serum sample, previously acidified to pH<4 using 1% TFA, was applied to the C18 cartridge. Additionally, 3 x of 400 μl of 0.1% TFA in Milli-Q water was passed through the C18. This step was aimed at removing unbound *N*-glycans from the C18 cartridge by being adsorbed on the carbograph column. The C18 columns were then removed and discarded. The carbograph columns underwent a further 3 x 600 μl of 0.1% TFA in Milli-Q water wash, to remove excess reagents and salts. Then, *N*-glycans were eluted with 500 μl 25% ACN +0.1% TFA (x2) and then with 500 μl 50% ACN +0.1% TFA (x2). The eluates were then dried under a reduced atmosphere in a vacuum concentrator in preparation for permethylation.

2.5.1.3 Permethylation

To stabilize the negatively charged sialic acids, samples were derivatized by permethylation using the protocol reported by Wedepohl et al. [219]. Briefly, a volume of 40 μl of DMSO saturated with NaOH was added into the dried *N*-glycan preparation, which was then incubated for 30 minutes on a shaker at room temperature. Subsequently, the free hydroxyl groups of glycans were methylated by the addition of 40 μl methyl iodide and incubated at room temperature for I hour on a shaker. Permethylated *N*-glycans were finally cleaned up using chloroform/water as solvents, as purified by liquid/liquid extraction. The chloroform phase was washed with water to a neutral pH. The chloroform phase was evaporated under reduced pressure in the centrifugal evaporator to end up with the permethylated *N*-glycans ready for MALDI-TOF measurement.

2.5.1.4 MALDI-TOF measurements of *N*-glycans

The dried *N*-glycans were dissolved in 10 μl of 75% aqueous acetonitrile. Equal volumes (0.5 μl) of *N*-glycans and the super 2,5-dihyroxybenzoic acid (sDHB) matrix were spotted in triplicate on the MALDI-TOF-MS ground steel target. A glucose ladder was used for calibration and the measurements were made in reflector positive ionization mode in the mass range of 1000 – 5000 Da. For each spectrum generated, at least 4,000 laser shots were made and baseline correction and peak picking were done by Flexanalysis. The raw mass spectra were exported as ASCII text for further processing to generate the assigned *N*-glycan structures with the GlycoWorkbench software.

2.5.2   Analysis of IgG$_1$ and IgG$_2$ glycopeptides

Preparation of serum samples for measurement of IgG$_1$ and IgG$_2$ glycopeptide was done through a series of steps that included affinity purification, tryptic digestion, cotton HILIC purification and finally MALDI-TOF measurement.

2.5.2.1 Affinity purification

IgG glycopeptides in the subjects' serum were isolated as previously carried out by Wuhrer et al. and Wieczorek et al. [57, 82]. Briefly, 30 µl of immobilized protein A 42 Sepharose beads in an Eppendorf tube were equilibrated by 400 µl of phosphate buffer (1x PBS) by spinning on a mini centrifuge for 30 sec. The beads were then allowed to settle to enable separation from the supernatant (about 350 µl), which was then aspirated from the tube and discarded. The process was repeated two more times using 350 µl (1XPBS). The volume of the beads was then brought up to 400 µl by the addition of 350 µl of 1 x PBS, followed by the addition of 10 µl of the serum. The sample mixture was then incubated on a rotor for 650 revolutions per minute for one hour at room temperature. After incubation, the sample mixture underwent a series of washes in phosphate buffer solution (4 x 200 µl of 1 x PBS), and 3 x 200 µl Milli-Q water through a shortened 200 µl filter tip. Finally, captured IgG was eluted using 3 x 100 µl of 100 mM formic acid, followed by drying under a vacuum centrifuge. It is of note that protein A sepharose completely captures the three IgG sub-classes - IgG$_1$, IgG$_2$, and IgG$_4$ - in the serum.

2.5.2.2 Tryptic digestion

The dry IgG$_1$ and IgG$_2$ isolated from serum were dissolved in 50 µl of 50 mM ammonium bicarbonate. This was followed by the addition of 5 µl of 0.2 µg/µl lyophilized sequencing-grade modified trypsin (reconstituted as per the manufacturer´s instructions) and subsequent overnight (16 hours) incubation at 37^{0}C. The digested glycopeptides were then dried under reduced pressure.

2.5.2.3 Cotton HILIC purification

The IgG glycopeptides were enriched by self-made micro-spin cotton hydrophilic interaction liquid chromatography (HILIC) columns of 10 µl filter tip, as previously described by Selman et al. [221]. This step is aimed at removing salts, non-glycosylated peptides, and reagents through the flow-through while capturing the glycopeptides. The filter tips were filled with fibers of 100% cotton from the cotton wool pads and then conditioned with 3 x 50 µl Milli-Q water and 3 x 50 µl of 80% ACN. The dried samples were then reconstituted in 50 µl of 80% ACN and loaded into the columns. The columns then underwent 3 washes of 80% ACN containing 0.1% TFA, followed by another three washes of 50 µl of 80% ACN without TFA. The glycopeptides, adsorbed on the stationary phase, were subsequently eluted using 50 µl of Milli-Q water (X 6) and then dried under reduced pressure in a vacuum centrifuge. At this point, the samples were ready for MALDI-TOF-MS measurements.

2.5.2.4 MALDI-TOF-MS measurements of IgG glycopeptides

The measurements of serum IgG glycopeptides were done with MALDI-TOF-MS. The serum IgG glycopeptide samples were reconstituted in 50 µl of Milli-Q water. One microliter of the sample was spotted on the MALDI target and allowed to dry before overlaying with 1µl of 2.5 mg/ml 4-chloro-α-cyanocinnamic acid matrix in 70% ACN + 0.1% TFA. The MALDI-TOF instrument was first calibrated using peptide calibration standard II. The measurements were made within the mass range of 1000 – 5000 Da in reflectron negative ionization mode. The measurements were done in the reflectron negative ionization mode to enable the detection of both negatively charged and neutral sialylated IgG_1/IgG_2 glycopeptides. For each sample spotted, 4000 laser shots were taken by random walk laser movement mode. The raw mass spectra data produced by the MALDI were then retrieved as ASCII text for further processing that included recalibration, baseline subtraction and peak extraction using the Massy tools software [239]. Recalibration was done using six IgG_1 glycopeptides, which included fucosylated agalactosyl (G0F), fucosylated monogalactosyl (G1F), Bisected-fucosylated agalactosyl (G0FN), fucosylated digalactosyl (G2F), bisected-fucoslated monogalactosyl (G1FN) and Sialylated fucosylated digalactosyl (G2FS1). The absolute intensities of the detected glycopeptides were normalized to the total area for IgG_1 and IgG_2. IgG_4 only accounts for a very small percentage of the total IgG (4%) and produces a very low signal in the order of the femtomole and subfemtomole range [57], and moreover, some of its masses overlap with the afucosylated IgG_2 glycopeptides, hence it was not considered in the current analysis.

2.5.3   Repeatability of the *N*-glycan and IgG glycopeptide workflows

The total *N*-glycans and IgG glycosylation methods were evaluated for their ability to reproduce results. Repeatability was evaluated as per previous studies [221, 240, 241], which involved intra-day

(within a day) and inter-day (in between day) analysis. Intra-day repeatability was assessed by preparing a healthy serum sample 3 times in parallel on the same day, while for inter-day repeatability, serum sample was prepared every day for three consecutive days. The mean, standard deviation (SD), and coefficient of variation (CV) were evaluated.

2.5.4   Analytical method for CA125, HBV, HCV and HIV

Serum CA125 values of the study subjects and their HBV, HCV, and HIV status were determined using the enzyme-linked immunoassay technique on a Cobas e 801 immunoassay system. The Cobas e 801 immunoassay system is a high-throughput, fully-automated immunochemistry module designed to carry out electrochemiluminescence sandwich immunoassays. The reagents used were Elecsys CA125 II, Elecsys HBsAg II, Elecsys anti-HCV II, and Elecsys HIV Duo (HIV-Ag and anti-HIV) reagent. All the reagents are packed in closed cassettes and ready to use. The analytical processes for CA125, HBV, HCV, and HIV were conducted in a fully automated mode, as per the flow chart in Figure 12.

Elecsys – CA125/HBsAg II/Anti-HCV II/ HIV Duo on Cobas e 801 using sandwich

First incubation

HBV
30ul sample was added into reagent containing 2 biotinylated monoclonal anti-HBsAg, and a mixture of monoclonal anti-HBsAg and polyclonal anti-HBsAg labeled with ruthenium complex[a] to form a

20ul sample was added into reagent containing 2 biotinylated monoclonal anti-CA125 and monoclonal anti-CA125 labeled with ruthenium complex[a] to form a sandwich complex.

HCV
30ul of sample was added into reagent containing biotinylated HCV-specific antigens and another reagent containing HCV-specific antibodies labeled with a ruthenium complex[a], then reacted to form a sandwich complex

HIV
30ul of sample was reacted - with biotinylated monoclonal anti-p24 antibodies and ruthenylated monoclonal anti-p24 antibodies (for HIV Ag detection) -or with biotinylated HIV-specific recombinant antigens with ruthenylated[a] HIV-specific recombinant antigens (for anti-HIV detection) forming a sandwich complex in both case (in separate vessels)

CA125

Second incubation

HBV/HCV/HIV/CA125
After addition of streptavidin-coated microparticle, the complex was eventually bound to the solid phase via interaction of biotin and

Measuring cell

HBV, HCV/HIV/CA125
The reaction mixture was transferred to measuring cell where microparticles are captured onto the surface of the electrode. Unbound substances were removed with ProCell/Procell M or ProCell II M (for HIV)

Readout

HBV, HCV/HIV/CA125
A voltage was applied to the electrode to induce chemiluminescent emission, which was then measured by a photomultiplier by comparing electrochemiluminescence signal of the sample product with the signal of the calibrant.

HBV, HCV and HIV result output – (reactive or non-reactive)/ CA125 in concentration

Figure 12. Methods for CA125, HBV, HCV and HIV measurement

2.6 Statistical analysis

The participants' demographic and clinical data were expressed in the form of mean, median, standard deviation (SD), range, interquartile range, total counts and percentages depending on the nature of the variable (continuous or non-continuous). The Mann-Whitney U test was used to compare the expression patterns of the detected *N*-glycans and IgG glycopeptides in the EOC patients against the BOD (control) subjects, which were then presented as medians, range, and p-values. The biomarker signatures were selected by plotting the peak intensities of *N*-glycans and IgG glycopeptides on to the Receiver Operating Curves (ROC) and the corresponding values of the area

under the curve (AUC, 95% C.I) were subsequently used to describe the accuracy levels of discriminating EOC patients from the control. Furthermore, the ROC curves were used in evaluating the performance of the *N*-glycan signatures in determining the EOC patients' response to chemotherapy by discriminating patients stratified into three categories based on chemotherapy use. The three categories included the pre-treatment group (primary EOC patients), 1-3 chemotherapy cycles and 4-6 chemotherapy cycles. AUC values of the ROC curve greater than 0.9 indicated an outcome of "excellent accuracy", while values from 0.8 to 0.9 indicated "good accuracy". AUC values from 0.7 to 0.8 indicated moderate accuracy, while values between 0.5 and 0.7 were interpreted as "uninformative". Box plots were generated to describe the distribution of the values of the *N*-glycan index (GLYCOV) and the CA125 in the primary EOC patients, EOC patients on chemotherapy (carboplatin and paclitaxel) and control subjects. To corroborate the previous studies findings on the effect of age on IgG galactosylation [54, 149], box plots were generated of the sum of the relative peaks of galactosylated IgG glycopeptides of the control subjects using different age clusters to describe the distribution patterns of galactosylation.

In addition, the data on the burden of infectious diseases in Kenyan OC patients were analyzed using similar statistics, where mean, median, standard deviation (SD), range, and interquartile range were applied to continuous variables such as age. Furthermore, comparative tests were applied on the non-continuous variables such as tumor stage (late or early), cancer management approaches of surgical debulking (surgical debulking done or not done), and on/or not on chemotherapy. The comparative tests used included the Chi-square test, Fisher's exact test, two proportions Z test, and Mann-Whitney U test as appropriate. The odds ratios (OR) were also determined for risk estimation. Statistical analyses were conducted using SPSS version 28 (SPSS Inc, Chicago, Illinois, USA).

3. Results

3.1 *N*-Glycan signatures for the diagnosis of EOC and monitoring response to chemotherapy

I analyzed for the first time the total *N*-glycome of a cohort of African EOC patients and BOD (control) subjects to identify signatures for monitoring the clinical response to chemotherapy as well as for complementing the CA125 test to improve early diagnosis of EOC. The analyses were performed on serum samples of EOC patients and controls who were HIV, HBV and HCV seronegative.

3.1.1   Subjects' demographics and clinical factors

The study subjects comprised 53 EOC patients and 46 control subjects of African ethnicity above the age of 18 years (Table 4). Of the 53 EOC patients, 15 were in FIGO stages I and II (early stage) while 38 were in stages III and IV (late stage). Further classification of EOC patients according to their status at enrollment into the study showed that 19 were primary EOC cases (pre-treatment), while 12 patients had received between one to three cycles of chemotherapy (1-3 chemotherapy cycles group), and the remaining 22 had received from four to six cycles of chemotherapy (4-6 chemotherapy cycles group). The patients who were on chemotherapy were all responders to carboplatin-paclitaxel combination therapy that was being administered at the time of enrollment. The subjects' serum CA125 levels were determined on a Cobas e 801 immunoassay system. The EOC patients had serum CA125 levels ranging from 5.0 to 31214.0 kU/L, with a median value of 92.4 kU/L. Conversely, the CA125 values for the control subjects ranged from 4.2 to 998.0 kU/L, with a median value of 25.9 kU/L. The reference values for CA125 were up to 35 kU/L.

Table 4. Demographics and clinical factors of the African cohort

Factors	BOD (n=46)	EOC (n=53)
Age (years)		
Mean ±SD	37±11	50 ±16
Range	18 -74	20 – 81
Median	34	50
Interquartile range	28 – 46	38 – 65
Stage		
FIGO stage I + II (early stage)	-	15
FIGO stage III + IV (late stage)	-	38
EOC patient (chemotherapy) categories		
Pre-treatment (Primary EOC patients)	-	19
1 – 3 chemotherapy cycles	-	12
4 – 6 chemotherapy cycles	-	22
CA125 values (kU/L)		
Mean ±SD	76.0 ±165.2	991.6 ±4318.4
Range	4.2 – 998.0	5.0 - 31214
Median	25.9	92.4
Interquartile range	12.1 – 65.5	77.2 – 1017

EOC patients were categorized into the early and late stages, while chemotherapy treatment was stratified into 3 categories depending on the chemotherapy cycles taken at the time of patient enrollment into the study.

3.1.2   Evaluation of method repeatability for the total *N*-Glycans

N-Glycans were released from denatured total serum glycoproteins through the proteolytic activity of the PNGase F. After isolation and purification, they were permethylated prior to MALDI-TOF-MS measurements yielding 46 signals corresponding to *N*-glycans. Inter- and intra-day reproducibility were assessed as described in section 2.6. Four *N*-glycan signals of the high-mannose type, namely N_2H_5 (*m/z* 1579.8), N_2H_6, (*m/z* 1783.9), N_2H_7, (*m/z* 1988.0), and N_2H_8 (*m/z* 2192.1), were selected for this evaluation because they were previously described by our research group as potential signatures of EOC in cohorts of Caucasians. These four *N*-glycans are highlighted in bold letters in the supplementary materials section (Table S1). The relative peak intensities of the four *N*-glycan signals were exported to an Excel worksheet in order to calculate their means, standard deviations (SD), and coefficient of variations (CV) (Figure 13). The CVs for intra-day measurements were 4.5% (N_2H_5), 4.0% (N_2H_6), 4.0% (N_2H_7), and 3.8% (N_2H_8), while for intra-day measurements, the CVs were 4.8% (N_2H_5), 8.8% (N_2H_6), 10.6% (N_2H_7), and 1.7% (N_2H_8). The coefficients of variation of both intra- and inter-day repeatability were low (3.8 – 10.6%) and within the acceptable limit, where CV <10%, (very good) and 10 -20%, (good), despite the fact that the *N*-glycan signals had low relative intensities, which is also an attribute of the good sensitivity of the method.

	N_2H_5	N_2H_6	N_2H_7	N_2H_8
Intra-day CV%	4.5	4.0	4.0	3.8

	N_2H_5	N_2H_6	N_2H_7	N_2H_8
Inter-day CV%	4.8	8.8	10.6	1.7
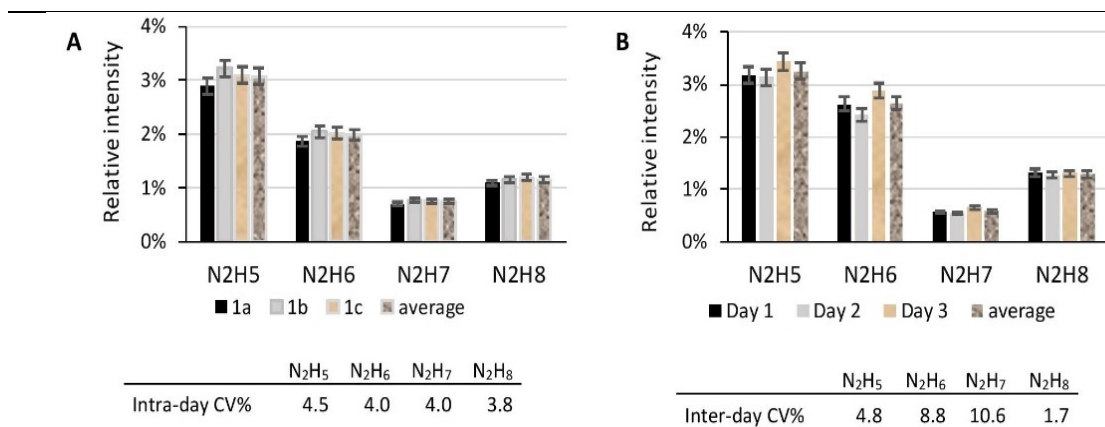
Figure 13.   Evaluation of total *N*-glycan method repeatability. The relative intensities of the four high-mannose *N*-glycans previously described by the Blanchard group as potential signature biomarkers for EOC were used in this experiment. (A) intra-day repeatability was determined with one sample of primary EOC patient that was analyzed in triplicate within the same day. (B) inter-day repeatability was determined with one sample of primary EOC patient that was analyzed in triplicate on three consecutive days.

3.1.3 Comparison of the *N*-glycans expression patterns between primary EOC patients and controls

From the serum samples of 19 primary African EOC patients and 46 control subjects, 46 signals of permethylated *N*-glycans were obtained and assigned in the MALDI-TOF mass spectrum of each individual (Supplementary materials, Table S1). Figure 15 shows a representative MALDI-TOF mass spectra of an African control subject against an EOC patient. In the mass spectra, the signals

corresponding to high-mannose were decreased in the EOC patient when compared to the control subject, while the signals corresponding to the complex-type tri- and tetrantennary *N*-glycans characterized by high sialylation and fucosylation were increased in the EOC patient compared to the control subject. The 46 *N*-glycans identified in the current analysis corresponded with the *N*-glycan structures previously detected by our research group in a Caucasian cohort of similar clinical description [64]. An investigation of the differences in the *N*-glycans expression between the African primary EOC patients and the control subjects was carried out using a non-parametric Mann-Whitney U test. The differences in expression of *N*-glycans were described from the medians and the range of their relative intensities as well as p-values.

Thirty-nine of the 46 detected *N*-glycans were found to be differentially expressed between the primary African EOC patients and their control counterparts (p <0.05) (Supplementary materials, Table S1). Of the 39 *N*-glycans, 27 had their peak intensities significantly decreased in EOC patients compared to controls, and in addition, their *m/z* values were smaller than 3600. They comprised high-mannoses (4), hybrid-type and monoantennary *N*-glycans (4), and complex-type multiantennary (bi-, tri- and tetrantennay) *N*-glycans (19). Conversely, the remaining 12 of the 39 *N*-glycans were significantly upregulated in primary EOC patients compared to the control subjects, apart from having *m/z* values of more than 3600. They were also multiantennary structures, but of either tri- or tetrantennary complex-type *N*-glycans characterized by high sialylation and/or fucosylation.
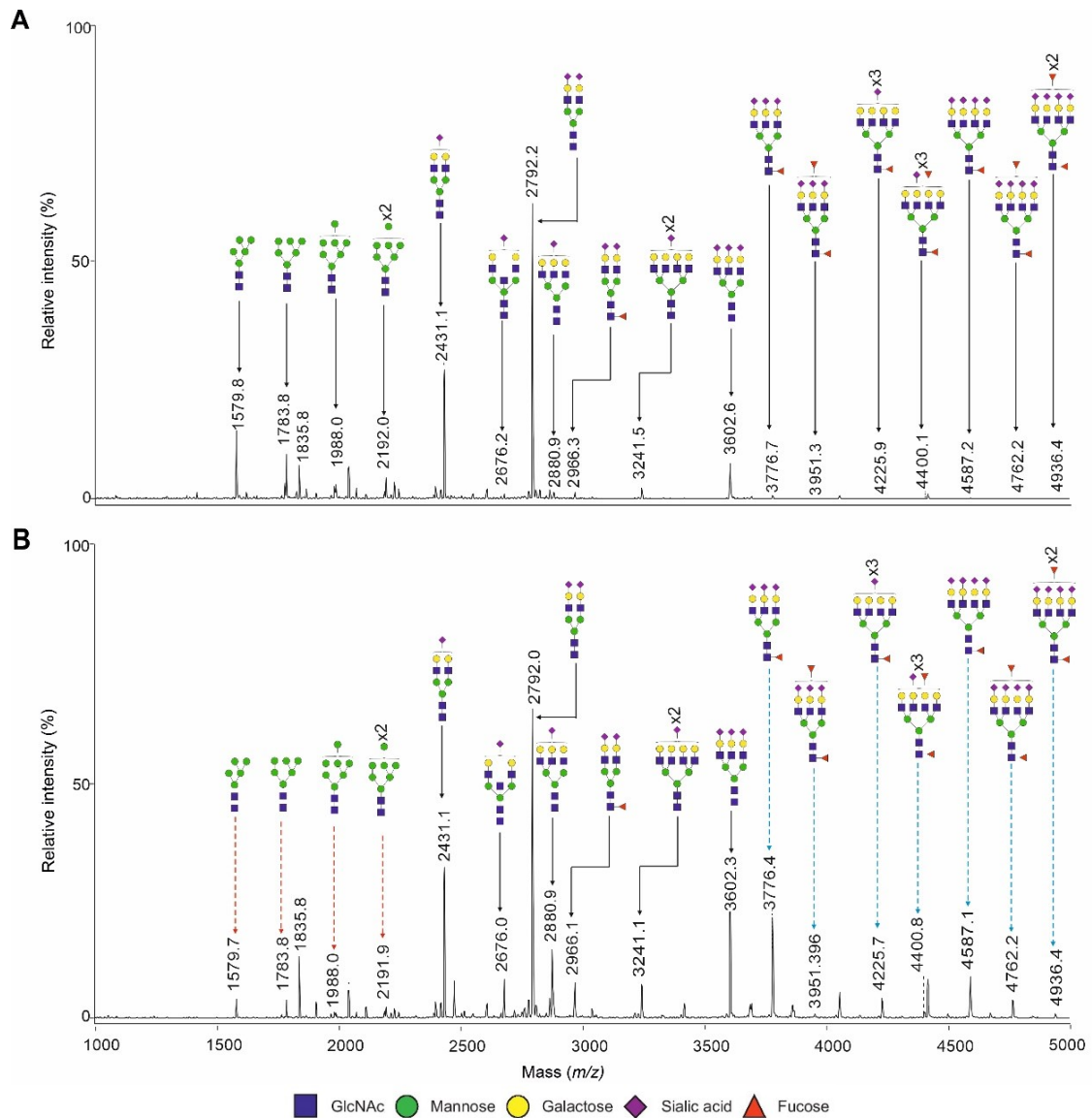
Figure 14. Representative MALDI-TOF mass spectra of the permethylated *N*-glycans of (A) a BOD subject and (B) a primary EOC patient from an African cohort. The red dotted arrows in (B) represent downregulated high-mannose *N*-glycan peaks and the blue dotted arrows in (B) represent upregulated complex-type *N*-glycan peaks in the primary EOC patient compared to the BOD subject. Measurements were carried out in positive-ion mode and molecular ions are present in their [M+Na]+ form.

3.1.4 *N*-Glycosylation changes in epithelial ovarian cancer upon intake of chemotherapy

The permethylated total *N*-glycans in serum samples of EOC patients, broadly categorized as pre-treatment or primary EOC patients and others undergoing chemotherapy, were measured on the MALDI-TOF-MS platform to identify changes in glycosylation caused by an efficacious combination of platin-taxane therapy. EOC patients were stratified into three categories, namely: prior to chemotherapy (primary/pre-treatment), 1-3 and 4-6 chemotherapy cycles to facilitate identification, and description of such glycan changes mediated by the effective chemotherapy strategy. The

alterations of glycosylation were described in terms of how the *N*-glycans were expressed (as either upregulation or downregulation), and the alterations were related with the *N*-glycan type, (high-mannose, hybrid-type and complex-type), antennarity, compositional profile (sialylation and/or fucosylation) and the mass range.

Representative MALDI-TOF mass spectra for patients in various treatment categories are shown in Figure 16, where 16 A is a spectrum for a patient in the pre-treatment category, while 16 B is for a patient in the 1-3 chemotherapy cycles group and 16 C is for a patient in the category of 4-6 chemotherapy cycles. Figure 16 shows that the *N*-glycans of *m/z* <3600, which included the high-mannose (red continuous arrows), had reduced peak intensities in primary EOC patients (Figure 16 A). However, the peak intensities of the same *N*-glycans (high-mannose) are higher in Figure 16 B, (dotted red arrows) compared to Figure 16 A, but lower (pink dotted arrows) than those of patients in the category of 4-6 chemotherapy cycles (Figure 16 C). Conversely, the change in the peak intensities of the *N*-glycans of *m/z* >3600 was the opposite of those with *m/z* <3600, which were solely of complex-type. Their peak intensities were increased (highest) in primary patients (blue continuous arrows), as shown in Figure 16 A, which decreased in patients within subsequent groups of 1-3 chemotherapy cycles (Figure 16 B) (sky-blue dotted arrows), with the lowest peak intensities in Figure 16 C (dotted purple arrows).

The 39 *N*-glycans with altered expression (27 downregulated and 12 upregulated) in primary EOC patients were further analyzed among EOC patients undergoing anti-cancer treatment to determine whether further alterations in glycosylation that resulted from the use of the chemotherapy could be exploited for clinical use in monitoring patients' response to chemotherapy.
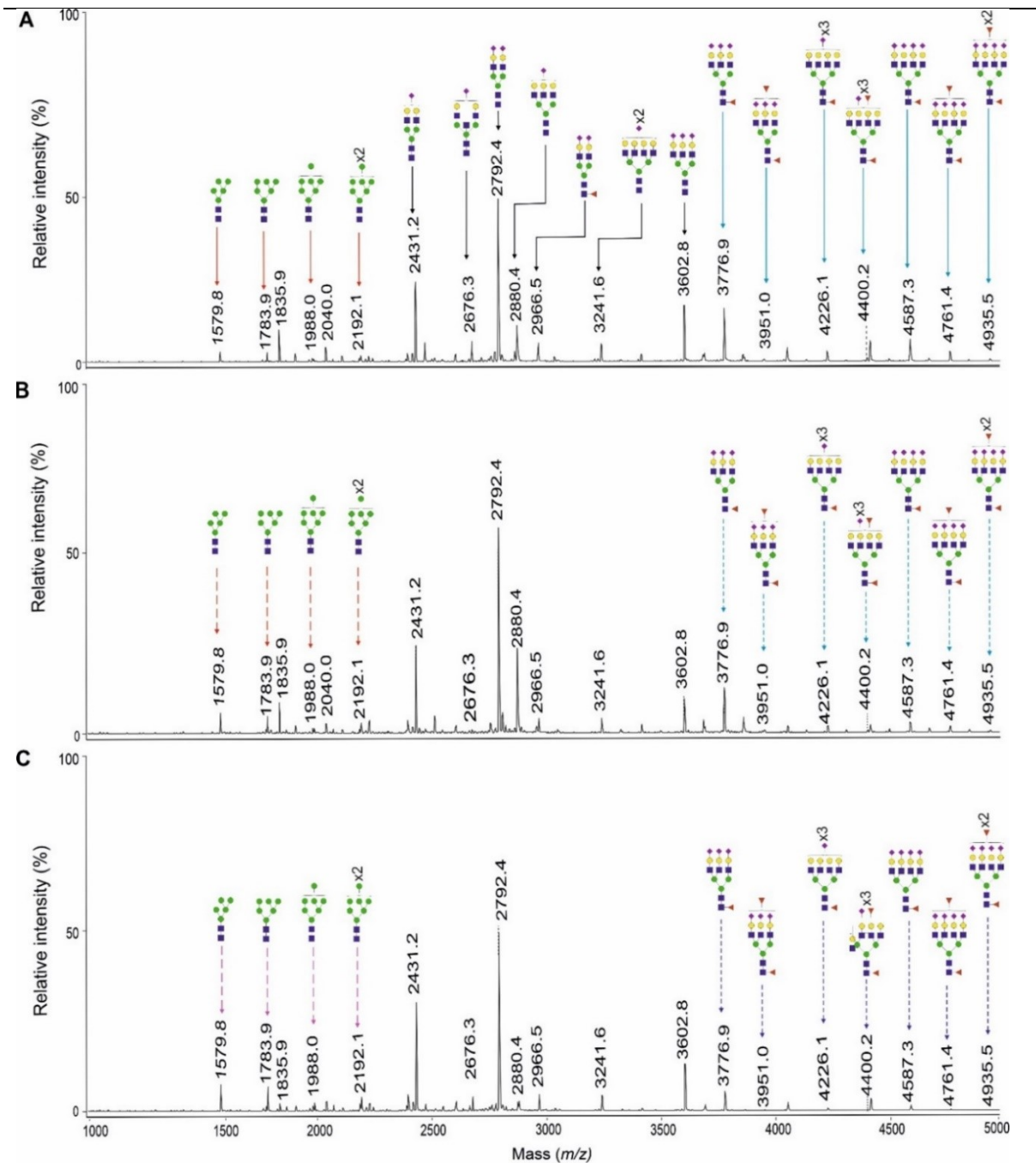
Figure 15. MALDI-TOF mass spectra of the permethylated *N*-glycans of three representative EOC patients in the course of chemotherapy: A, prior to chemotherapy; B, 1–3 chemotherapy cycles; C, 4–6 chemotherapy cycles. Measurements were performed in positive ionization mode and molecular ions are present in their [M+Na]+ form.

3.1.5 Changes in the downregulated *N*-glycans (*m/z* <3600) of primary EOC patients upon intake of chemotherapy

Differential analysis by Kruskal-Wallis test of the 27 *N*-glycans of *m/z* <3600 that were significantly downregulated in the primary EOC patients (Supplementary materials, Table S1) showed significant differences in their expression across the three patients' categories of anti-cancer treatment (p<0.05)

(Table 5). The medians of the peak intensities of the 27 *N*-glycans were shown to increase consistently from the category of pre-treatment, which recorded the lowest median value, to the categories of 1-3 and 4-6 cycles of chemotherapy, which had higher and the highest median values, respectively.

All the 27 *N*-glycans except peak *m/z* 3037.5 discriminated the primary EOC patients (pre-treatment) from those in the category of 4–6 chemotherapy cycles, with acceptable accuracies that ranged from "moderate" to "excellent" (AUC, 0.71 – 0.90). Similarly, the CA125 test discriminated the two patient categories from each other with good accuracy (AUC = 0.80). Conversely, of the 27 *N*-glycans, a total of 17 were able to discriminate patients in the pre-treatment category from those in the 1-3 chemotherapy cycles group (AUC, 0.71 – 0.81). They comprised three high-mannose *N*-glycans (*m/z* 1579.8, 1783.9 and 1988.0), four hybrid-type and monoantennary *N*-glycans (*m/z* 1620.8, 1416.7, 1982.0 and 2390.2), and ten complex-type multiantennary (bi-, tri- and tetrantennary) *N*-glycans (*m/z* 2285.2, 2315.2, 2489.3, 2850.4, 1661.8, 2070.0, 2519.3, 2693.4, 2880.4 and 3054.5). However, CA125 failed to discriminate the two groups of patients (i.e., the pre-treatment category from those in the 1-3 chemotherapy cycles) (AUC, 0.54). Interestingly, all 27 *N*-glycans produced non-informative results (AUCs < 0.70) when they were applied to discriminate patients in the category of 1-3 chemotherapy cycles from those of 4–6 chemotherapy cycles.

Therefore, an AUC value of > 0.70 was used as the criteria for selecting the best-performing *N*-glycan signatures in differentiating primary EOC patients from the two categories of patients that had taken 1-3 and 4-6 cycles of chemotherapy. A total of 17 *N*-glycan signatures (*m/z* <3600) were identified as the best performers and they included all the *N*-glycans that differentiated patients in the pre-treatment group from those that had taken 1-3 chemotherapy cycles. Overall, the *N*-glycan that showed the best indication for anti-cancer agent efficacy in early chemotherapy sessions (within the first three cycles of chemotherapy) was peak *m/z* 1620.8 (AUC, 0.81).

Additionally, it was of interest in the present study that the four high-mannoses (*m/z* 1579.8, 1783.9, 1988.0 and 2192.3) previously described by our group as part of the 11 signatures for diagnosis of EOC in Caucasian ethnicity were a part of the 17 selected signatures. They differentiated the primary EOC patients from those that had taken 4-6 chemotherapy cycles with acceptable accuracies (AUC, 0.71 – 0.87), as well as primary EOC patients from those that had taken 1-3 chemotherapy cycles (AUC, 0.71 – 0. 74), except for one of peak *m/z* 2192.3.

Table 5. Changes in the downregulated *N*-glycans (*m/z* <3600) of primary EOC patients upon intake of chemotherapy

Description	*N*-glycan	Mass (*m/z*)	*N*-glycan structure	Chemotherapy cycles Relative intensity			p-value	Chemotherapy cycles AUC		
				Pre-treatment Median (min-max)	1 - 3 cycles Median (min-max)	4 - 6 cycles Median (min-max)		Pre-treatment / 1-3	1-3 / 4-6	Pre-treatment /4-6
High-Mannose	N_2H_5	1579.8		0.013 (0.003 -0.020)	0.021 (0.005-0.032)	0.022 (0.008-0.038)	<0.001	**0.74**	0.60	**0.87**
	N_2H_6	1783.9		0.013 (0.003 -0.020)	0.019 (0.005-0.033)	0.021 (0.007-0.039)	<0.001	**0.73**	0.59	**0.85**
	N_2H_7	1988.0		0.004 (0.001 -0.009)	0.006 (0.002-0.011)	0.006 (0.003-0.015)	<0.001	**0.71**	0.65	**0.85**
	N_2H_8	2192.1		0.008 (0.001 -0.019)	0.012 (0.004-0.020)	0.014 (0.006-0.026)	0.001		0.68	0.67	**0.82**
Hybrid-type and monoantennary	N_3H_4	1620.8		0.001 0.0003-0002	0.002 0.001-0.003	0.002 0.001-0.004	<0.001	**0.81**	0.51	**0.84**
	N_3H_3	1416.7		0.001 0.0004-0.003	0.002 0.0003-0.003	0.002 0.001-0.004	0.030	**0.73**	0.55	**0.71**
	$N_3H_4S_1$	1982.0		0.004 0.001-0.008	0.007 0.002-0.011	0.006 0.004-0.012	0.005	**0.75**	0.50	**0.79**
	$N_3H_6S_1$	2390.2		0.003 0.001-0.007	0.004 0.001-0.006	0.005 0.003-0.007	<0.001	**0.70**	0.69	**0.90**
Complex-type (≥ biantennary)	N_5H_4	2111.1		0.002 0.001-0.007	0.041 0.001-0.009	0.004 0.003-0.008	0.041	0.66	0.53	**0.73**
	N_5H_5	2315.2		0.001 0.0002-0.002	0.002 0.0003-0.003	0.002 0.001-0.003	0.001	**0.73**	0.54	**0.83**
	$N_5H_4F_1$	2285.2		0.001 0.0002-0.002	0.001 0.0003-0.002	0.001 0.0008-0.002	<0.001	**0.74**	0.60	**0.87**
	$N_5H_5F_1$	2489.3		0.001 0.0002-0.003	0.002 0.0004-0.003	0.002 0.001-0.004	0.001	**0.74**	0.53	**0.84**

	m/z	Structure				p-value			
$N_5H_5S_1$	2676.3		0.004 0.001-0.008	0.005 0.001-0.011	0.007 0.004-0.014	0.017	0.59	0.67	**0.75**
$N_5H_5S_1F_1$	2850.4		0.002 0.0004-0.011	0.005 0.0008-0.011	0.005 0.003-0.010	0.002	**0.73**	0.55	**0.82**
$N_5H_5S_2$	3037.5		0.002 0.001-0.004	0.003 0.0004-0.006	0.003 0.001-0.004	0.027	0.51	0.65	**0.64**
N_4H_3	1661.8		0.001 0.0003-0.005	0.003 0.001-0.010	0.002 0.001-0.006	0.014	**0.75**	0.62	**0.72**
N_4H_5	2070.0		0.003 0.001-0.007	0.006 0.001-0.010	0.007 0.004-0.013	<0.001	**0.75**	0.59	**0.89**
N_5H_6	2519.3		0.001 0.0002-0.002	0.002 0.0004-0.003	0.002 0.001-0.004	<0.001	**0.78**	0.57	**0.88**
N_6H_6	2764.4		0.003 0.001-0.006	0.004 0.001-0.008	0.005 0.003-0.009	0.001	0.68	0.63	**0.84**
$N_4H_4F_1$	2040.0		0.009 0.002-0.024	0.015 0.003-0.026	0.018 0.008-0.0.27	0.032	0.63	0.60	**0.74**
$N_4H_4S_1$	2227.1		0.007 0.002-0.015	0.011 0.003-0.014	0.010 0.006-0.022	0.044	0.64	0.57	**0.73**
$N_4H_5S_1F_1$	2605.3		0.006 0.001-0.015	0.007 0.003-0.011	0.009 0.004-0.017	0.019	0.58	0.65	**0.76**
$N_5H_6F_1$	2693.4		0.001 0.0002-0.002	0.001 0.0003-0.002	0.001 0.001-0.002	<0.001	**0.74**	0.59	**0.86**
$N_5H_6S_1$	2880.4		0.009 0.003-0.022	0.014 0.002-0.027	0.017 0.008-0.032	<0.001	**0.72**	0.60	**0.84**
$N_4H_5F_1$	2244.1		0.002 0.0003-0.006	0.003 0.001-0.008	0.005 0.003-0.008	<0.001	0.67	0.75	**0.86**
$N_4H_5S_1$	2431.2		0.090 0.053-0.152	0.105 0.035-0.165	0.129 0.102-0.172	<0.001	0.62	0.73	**0.86**
$N_5H_6S_1F_1$	3054.5		0.001 0.0004-0.003	0.002 0.001-0.005	0.002 0.001-0.003	0.019	**0.70**	0.54	**0.75**

Medians in each cluster of study subjects were calculated from the relative peak intensities of the *N*-glycans generated from the MALDI-TOF-MS. Min-Max is the minimum and maximum value of the relative intensities of the *N*-glycan peaks in each cluster of study subjects. AUC (area under the curve) values evaluated the accuracy of each *N*-glycan in discriminating EOC from BOD. p-values less than 0.05 were considered as statistically significant. The four high-mannose *N*-glycan signatures for EOC diagnosis previously described by our research group are highlighted in bold.

3.1.6 Changes in the upregulated *N*-glycans (*m/z* >3600) of primary EOC patients upon intake of chemotherapy

The 12 complex-type multiantennary *N*-glycans of *m/z* >3600 that showed significant increases in their peak intensities among the primary EOC patients (Supplementary materials, Table S1) were evaluated for the differences in their expression among the EOC patients in the three categories of chemotherapy treatment. The differential expression of nine *N*-glycans in the three categories of treatment was found to be statistically significant (p<0.05). They comprised *N*-glycan peaks of *m/z* 3415.7, 3776.9, 3864.9, 3951.0, 4226.1, 4400.2, 4587.3, 4761.4, and 4935.5 (Table 6). The median peak intensities of the nine *N*-glycans were lower among patients in the categories that had taken chemotherapy compared to patients that had not commenced chemotherapy (pre-treatment). The category of 4-6 chemotherapy cycles had the lowest decrease, followed by the category of 1-3 chemotherapy cycles.
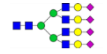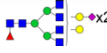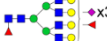
Interestingly, unlike the 17 *N*-glycan signatures (*m/z* <3600) that were downregulated in primary EOC patients, the nine *N*-glycan signatures identified in this section could not discriminate patients in the category of pre-treatment from the patients in the category of 1–3 chemotherapy cycles (AUC, < 0.70), except for peak *m/z* 4587.3 (AUC, 0.74). Instead, eight of the nine *N*-glycan peaks discriminated the patients in categories of 1-3 chemotherapy cycles from those in the category of 4–6 chemotherapy cycles (AUC, 0.70 − 0.81), which was also unlike the *N*-glycans of *m/z* <3600. Interestingly, all the nine *N*-glycan signatures, just like the 17 *N*-glycan signatures of *m/z* <3600, discriminated patients in the pre-treatment category from those in the 4–6 chemotherapy cycles group (AUC, 0.70 − 0.87). Of all the nine *N*-glycan signatures, the *m/z* 4935.5 peak was the best discriminator for both 1-3 against 4-6 chemotherapy cycles and pre-treatment against 4-6 chemotherapy cycles (AUC, >0.80).

Therefore, using the criteria of AUC of > 0.70 for selecting the best-performing *N*-glycan signatures for discriminating EOC patients according to their treatment categories, 1-3 chemotherapy cycles from 4-6 chemotherapy cycles and pre-treatment category from 4-6 chemotherapy cycles, eight *N*-glycans fulfilled the conditions and were subsequently selected as potential signatures. They included peaks of *m/z* 3415.7, 3776.9, 3864.9, 3951.0, 4226.1, 4400.2, 4761.4, and 4935.5.

It was also important to note for the present study that, of the nine complex-type *N*-glycan identified as differentially expressed, seven were among the 11 signatures previously described by our laboratory as biomarkers for EOC among the Caucasian population. The other four (high-mannose) of the 11 *N*-glycan signatures were previously described in section 3.1.5. The seven *N*-glycan peaks were *m/z* 3776.9, 3951.0, 4226.1, 4400.2, 4587.3, 4761.4, and 4935.5, which formed the component

of the upregulated *N*-glycans in the "GLYCOV index". As a result, the 11 *N*-glycan structures previously described by our group were demonstrated to show sufficiency in describing the total *N*-glycome changes associated with EOC and the response to chemotherapy in the present African cohort.

Table 6. Changes in the upregulated *N*-glycans (*m/z* >3600) of primary EOC patients upon intake of chemotherapy

N-glycan	Mass (*m/z*)	*N*-glycan structure	Chemotherapy cycles Relative intensity			p-value	Chemotherapy cycles AUC		
			Pre-treatment Median (min-max)	1 - 3 cycles Median (min-max)	4 - 6 cycles Median (min-max)		Pre-treatment /1-3	1-3 /4-6	Pre-treatment /4-6
$N_6H_7S_4$	4413.2		0.026 0.010-0.042	0.018 0.004-0.076	0.018 0.002-0.051	0.071	0.69	0.58	0.68
$N_6H_7S_3$	4052.0		0.013 0.005-0.024	0.014 0.002-0.036	0.012 0.003-0.021	0.818	0.56	0.50	0.54
$N_5H_6S_3$	3602.8		0.067 0.032-0.091	0.060 0.015-0.106	0.077 0.017-0.108	0.391	0.62	0.63	0.54
$N_5H_6S_2F_1$	3415.7		0.006 0.001-0.011	0.005 0.001-0.018	0.003 0.001-0.012	0.023	0.53	**0.71**	**0.73**
$N_5H_6S_3F_1$	3776.9		0.056 (0.011 -0.095)	0.035 (0.002-0.088)	0.018 (0.001-0.082)	<0.001	0.68	**0.74**	**0.84**
$N_6H_7S_2F_1$	3864.9		0.003 0.0001-0.008	0.003 0.001-0.019	0.002 0.001-0.008	0.043	0.51	**0.70**	**0.70**
$N_5H_6S_3F_2$	3951.0		0.002 (0.0004- 0.008)	0.002 (0.001-0.009)	0.001 (0.0003-0.007)	0.008	0.53	**0.77**	**0.74**
$N_6H_7S_3F_1$	4226.1		0.009 (0.002-0.017)	0.005 (0.001-0.016)	0.002 (0.0004-0.014)	0.001	0.63	**0.70**	**0.83**
$N_6H_7S_3F_2$	4400.2		0.003 (0.0004-0.014)	0.002 (0.001-0.009)	0.001 (0.001-0.007)	0.004	0.63	**0.78**	**0.76**
$N_6H_7S_4F_1$	4587.3		0.020 (0.003-0.051)	0.014 (0.001-0.025)	0.005 (0.0004-0.026)	<0.001	**0.74**	0.66	**0.87**
$N_6H_7S_4F_2$	4761.4		0.008 (0.001-0.047	0.005 (0.0004-0.021)	0.001 0.0003-0.021)	<0.001	0.69	**0.74**	**0.83**
$N_6H_7S_4F_3$	4935.5		0.002 (0.0003-0.023)	0.001 (0.0003-0.007)	0.0001 (0.0002-0.009)	<0.001	0.60	**0.81**	**0.80**

Medians in each cluster of study subjects were calculated from the relative peak intensities of the *N*-glycans measured by MALDI-TOF-MS. Min-Max is the minimum and maximum value of the relative intensities of the *N*-glycan peaks in each category of study subjects. AUC values evaluated the potential of each *N*-glycan to discriminate EOC from BOD. p-values less than 0.05 were considered statistically significant. The complex-type *N*-glycan signatures for EOC diagnosis in Caucasians previously described by our research group are highlighted in bold.

3.1.7 Performance of the *N*-glycan index in discriminating patients within categories of anti-cancer treatment.

Further analysis was carried out to evaluate the performance of the 11 *N*-glycan signatures, in the form of an "*N*-glycan index" (GLYCOV) in correctly assigning the EOC patients into their respective treatment categories. The *N*-glycan index was computed as a ratio of the sum of relative intensities of the seven upregulated complex-type *N*-glycans to the sum of relative intensities of the four downregulated high-mannose *N*-glycans in EOC patients, namely: (sum of relative areas of *m/z* 3776.9, 3951.0, 4226.1, 4400.2, 4587.3, 4761.4, and 4935.5)/ 7*4/ (sum of relative areas of *m/z* 1579.8, 1783.9, 1988.0 and 2192.3) [64]. The box plot of log2-transformed values of the *N*-glycan index in the three treatment clusters showed a linear decrease of median *N*-glycan index values from the highest in the pre-treatment category, to the category of 4-6 chemotherapy cycles, which had the lowest median value (Figure 17). The differences in the median values of the *N*-glycan index in each of the three groups were statistically significant (the category of pre-treatment against 1-3 chemotherapy cycles was p=0.012, 1-3 chemotherapy cycles against 4-6 chemotherapy cycles was p=.037, and pre-treatment against 4-6 chemotherapy cycles was p=1.3E-4).

Conversely, the box plot of the median values of the log2-transformed CA125 values in the three treatment clusters produced an irregular pattern that made it difficult to determine the efficacy of chemotherapy in early sessions. This was supported by the lack of significant difference between the log2-transformed median CA125 values of the pre-treatment categories against those in 1-3 chemotherapy cycles (p=0.734). However, the differences in the log2-transformed median values of CA125 of the categories 1-3 chemotherapy cycles against 4-6 chemotherapy cycles (p=0.021) and the pre-treatment against 4-6 chemotherapy cycles (p=0.001) were statistically significant.

The *N*-glycan index was more accurate in discriminating EOC patients according to their categories of treatment compared to the CA125 values or individual *N*-glycan signatures. For discriminating pre-treatment patients from 1-3 chemotherapy cycles, the AUC value of the *N*-glycan index was 0.77 against 0.54 for CA125. Furthermore, comparison of the *N*-glycan index against CA125 in differentiating the category of 1 -3 cycles from 4-6 chemotherapy cycles yielded an AUC of 0.72 (*N*-glycan index) against 0.74 (CA125), and finally, pre-treatment against 4-6 chemotherapy cycles was 0.89 (*N*-glycan index) against 0.80 (CA125) (Figure 18).
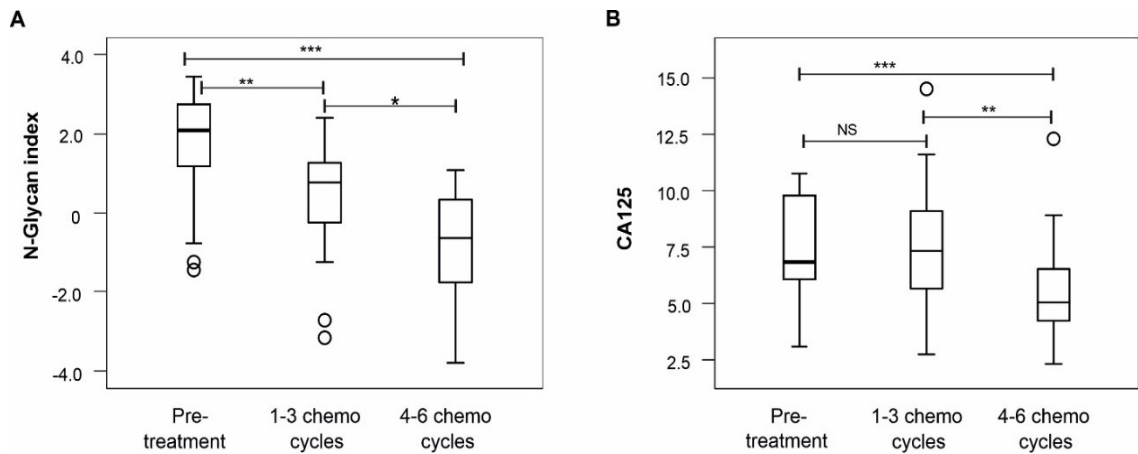
Figure 16. Box plots of the log2 transformed values of (A) the *N*-glycan index (GLYCOV) and (B) CA125 showing their comparative distribution in the three categories of chemotherapy. The Mann-Whitney U test was applied to test for the difference in the distribution of *N*-glycan index and CA125 values across the patient treatment clusters (NS-not significant, *p<0.05, **P<0.01 and ***p<0.001). Pre-treatment/1-3 chemotherapy cycles, *N*-glycan index (p=0.012), CA125 (p=0.734); 1-3/4-6 chemotherapy cycles, *N*-glycan index (p=.037), CA125 (p=0.021); and pre-treatment/4-6 chemotherapy cycles, *N*-glycan index (p=1.3E-4), CA125 (p=0.001).
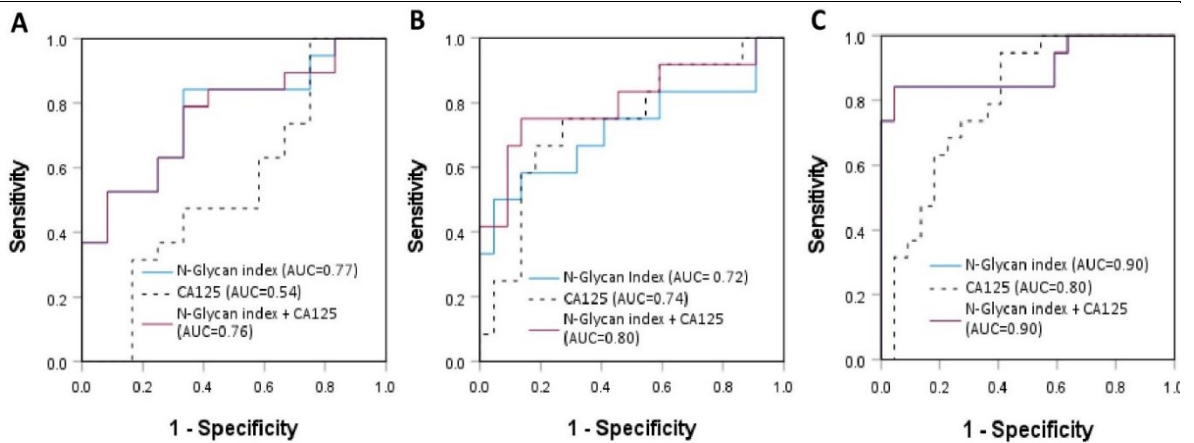
Figure 17. ROC curves of the *N*-glycan index, CA125 and a combination of *N*-glycan index and CA125 in assessing response to chemotherapy. (A) Pre-treatment versus 1-3 chemotherapy cycles, (B) 1-3 versus 4-6 chemotherapy cycles and (C) pre-treatment versus 4-6 chemotherapy cycles. The *N*-glycan index had a superior discriminative performance compared to CA125 (AUCs, 0.77 vs. 0.54, 0.72 vs. 0.74 and 0.90 vs. 0.80).

3.1.8   *N*-Glycan profiling for signatures of epithelial ovarian cancer diagnosis in an African cohort

Evaluation of altered *N*-glycosylation in primary EOC for potential diagnostic *N*-glycan signatures was carried out on the 39 *N*-glycans that were differentially expressed between primary EOC patients and the controls as shown in section 3.1.4 and in Supplementary materials, Table S1. Analytical descriptions of the changes in glycosylations were anchored on classifications such as *N*-glycan type, antennarity and the compositional profile. The ROC curves for each of the 39 *N*-glycans were built, representing 27 (*m/z* <3600) and 12 (*m/z* >3600) *N*-glycans that were downregulated and upregulated in EOC patients, respectively, and their corresponding AUC values used to describe the accuracy of differentiating the primary EOC patients from the controls.

All the 27 *N*-glycans that were downregulated among the primary EOC patients had acceptable discriminatory accuracies of EOC patients from the controls (AUC, > 0.70) (Table 7). Of the 27 *N*-glycans, 12 produced excellent discriminatory accuracies (AUC, 0.90 – 0.96), and they included four high-mannose (at *m/z* 1579.8, 1783.9, 1988.0, and 2192.1), two monoantennary (*m/z* 1620/8 and 1982.0), and six complex-type *N*-glycans (*m/z* 2285.2, 2070.0, 2519.3, 2244.1, 2431.2 and 2693.4). Another set of 11 *N*-glycans comprising one monoantennary (*m/z* 1426.7), one hybrid-type (*m/z* 2390.2) and nine complex-type *N*-glycans (*m/z* 2315.2, 2489.3, 2850.4, 1661.5, 2040.0, 2227.1. 2605.3, 2764.4 and 2880.4) discriminated primary EOC patients from the controls with "good accuracies" (AUC, 0.82 – 0.89). Finally, four complex-type *N*-glycans (*m/z* 2111.1, 2676.5, 3037.5 and 3054.5) had moderate discriminatory accuracies of EOC patients from the controls (AUC, 0.70 – 0.79).

By contrast, the 12 upregulated *N*-glycans in primary EOC patients were all complex-type multiantennary (tri- or tetrantennary) *N*-glycans characterized by high sialylation and/or fucosylation (Table 8). Eleven of the 12 *N*-glycans had the acceptable discriminatory performance of EOC patients from the controls (AUC, >0.70). Their discriminatory accuracies ranged from "excellent" for peak *m/z* 4587.3 (AUC, 0.91) to "good" for five peaks at *m/z* 3776.9, 4400.2, 4413.2, 4226.1 and 4761.4 (AUC, 0.80 – 0.89), while another five had moderate accuracies (AUC, 0.75 – 0.79). They included peak *m/z* of 3415.7, 3864.9, 3951.0, 4052.0 and 4935.5.

The *N*-glycans with downregulated peaks in primary EOC were better discriminators of EOC from controls compared to those that were upregulated. However, to guarantee the versatility and dynamism of the new biomarker in a wider scope of EOC patients, it is pertinent that the different attributes reflected by each of the glycans in the two groups are captured in the envisaged biomarker. Therefore, the selection criteria of the potential signature biomarkers for EOC diagnosis were based on the pattern of alteration in primary EOC, and more importantly with AUC values of at least >0.70, which implied acceptable accuracy. The cut-off AUC values for the downregulated *N*-glycan

signatures (*m/z* <3600) was therefore established at 0.90, while for the upregulated *N*-glycan signatures (*m/z* >3600) it was at 0.75. Consequently, a total of 23 *N*-glycan signatures for the diagnosis of EOC were selected in an African cohort. They included 12 *N*-glycan signatures for those downregulated in primary EOC (*m/z* 1579.8, 1783.9, 1988.0, 2192.1, 1620/8, 1982.0. 2285.2, 2070.0, 2519.3, 2244.1, 2431.2 and 2693.4). The other 11 *N*-glycan signatures (upregulated in primary EOC) were of *m/z* 4587.3, 3776.9, 4400.2, 4413.2, 4226.1, 4761.4, 3415.7, 3864.9, 3951.0, 4052.0 and 4935.5.

Interestingly, the 11 *N*-glycan signatures (4 high-mannose and 7 complex-type *N*-glycans) that were used to compute the GLYCOV biomarker of EOC diagnosis in Caucasians, as described by our research group [64, 107], were part of the current 23 EOC diagnostic *N*-glycan signatures identified in the African cohort. Evaluation of the 11 *N*-glycan signatures showed that the four high-mannoses had excellent accuracies (AUCs, 0.91 – 0.94) in discriminating EOC patients from the controls, while the seven complex-type *N*-glycans had accuracies ranging from "moderate" to "excellent" (0.75 - 0.91) (Tables 7 and 8). Therefore, the 11 *N*-glycan signature peak intensities for all the study subjects were computed into indices in section 3.1.9 and subsequently, their values were plotted in ROC curves to establish their discriminative performance of primary EOC patients from the controls.
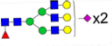
Table 7. Statistically significant downregulated *N*-glycans (*m/z* <3600) in primary EOC patients

N-glycan description	*N*-glycan Formula	*m/z*	*N*-glycan structure	BOD (n=46) Relative intensity		EOC (n=19) Relative intensity		p-value	AUC
				Median	Min-max	Median	Min-max		
High-Mannose	N_2H_5	1579.8		0.031	0.008-0.059	0.013	0.003-0.020	<0.001	0.94
	N_2H_6	1783.9		0.028	0.007-0.065	0.013	0.002-0.020	<0.001	0.91
	N_2H_7	1988.0		0.010	0.003-0.021	0.004	0.001-0.009	<0.001	0.92
	N_2H_8	2192.1		0.018	0.006-0.037	0.008	0.001-0.019	<0.001	0.91
Hybrid-tsype and monoantennary	N_3H_3	1416.7		0.002	0.001-0.005	0.001	0.0004-0.003	<0.001	0.85
	N_3H_4	1620.8		0.003	0.001-0.007	0.001	0.0003-0.002	<0.001	0.96
	$N_3H_4S_1$	1982.0		0.008	0.004-0.015	0.004	0.001-0.008	<0.001	0.92
	$N_3H_6S_1$	2390.2		0.005	0.002-0.012	0.003	0.001-0.007	<0.001	0.89
Compl	N_5H_4	2111.1		0.004	0.002-0.009	0.002	0.001-0.007	0.003	0.74

$N_5H_4F_1$	2285.2		0.002	0.001-0.003	0.001	0.0002-0.002	<0.001	0.91
N_5H_5	2315.2		0.002	0.001-0.004	0.001	0.0002-0.002	<0.001	0.89
$N_5H_5F_1$	2489.3		0.003	0.001-0.005	0.001	0.0002-0.003	<0.001	0.89
$N_5H_5S_1$	2676.3		0.007	0.004-0.016	0.003	0.001-0.010	<0.001	0.79
$N_5H_5S_1F_1$	2850.4		0.005	0.003-0.013	0.002	0.0004-0.011	<0.001	0.87
$N_5H_5S_2$	3037.5		0.002	0.002-0.006	0.002	0.001-0.004	0.002	0.74
N_4H_3	1661.8		0.003	0.001-0.010	0.001	0.0003-0.004	<0.001	0.83
N_4H_5	2070.0		0.009	0.003-0.023	0.003	0.001-0.007	<0.001	0.95
N_5H_6	2519.3		0.002	0.001-0.004	0.001	0.0002-0.002	<0.001	0.92
N_6H_6	2764.4		0.005	0.002-0.010	0.003	0.001-0.006	<0.001	0.82
$N_4H_4F_1$	2040.0		0.018	0.006-0.052	0.010	0.002-0.024	<0.001	0.81
$N_4H_4S_1$	2227.1		0.012	0.008-0.022	0.007	0.002-0.015	<0.001	0.88
$N_4H_5F_1$	2244.1		0.007	0.002-0.017	0.002	0.0003-0.006	<0.001	0.94
$N_4H_5S_1$	2431.2		0.144	0.085-0.189	0.090	0.053-0.152	<0.001	0.91
$N_4H_5S_1F_1$	2605.3		0.010	0.005-0.021	0.006	0.001-0.014	<0.001	0.86
$N_5H_6F_1$	2693.4		0.002	0.001-0.002	0.001	0.0002-0.002	<0.001	0.90
$N_5H_6S_1$	2880.4		0.016	0.007-0.029	0.009	0.003-0.022	<0.001	0.83
$N_5H_6S_1F_1$	3054.5		0.002	0.001-0.005	0.001	0.0004-0.003	0.012	0.70

Medians in each cluster of study subjects were calculated from the relative peak intensities of the *N*-glycans generated from MALDI-TOF-MS. Min-Max is the minimum and maximum value of the relative intensities of the *N*-glycan peaks in each cluster of study subjects. AUC (area under the curve) values evaluated the potential of each *N*-glycan to discriminate EOC from BOD. P-values were computed using the Mann-Whitney U test, where a value less than 0.05 was considered statistically significant. The *N*-glycan signatures for EOC previously described by our research group are highlighted in bold.

Table 8. Statistically significant upregulated complex-type multiantennary *N*-glycans (*m/z* >3600) in primary EOC patients

N-glycan Formula	*m/z*	*N*-glycan structure	BOD (n=46) Relative intensity		EOC (n=19) Relative intensity		p-value	AUC
			Median	Min-max	Median	Min-max		
$N_5H_6S_2F_1$	3415.7		0.002	0.001-0.016	0.006	0.001-0.011	0.001	**0.75**
$N_5H_6S_3$	3602.8		0.052	0.025-0.111	0.067	0.032-0.091	0.016	0.69
$N_5H_6S_3F_1$	3776.9		0.011	0.001-0.076	0.056	0.011-0.095	<0.001	**0.89**
$N_6H_7S_2F_1$	3864.9		0.001	0.0004-0.013	0.003	0.001-0.008	0.001	**0.76**
$N_5H_6S_3F_2$	3951.0		0.001	0.0003-0.006	0.002	0.0004-0.008	0.002	**0.75**
$N_6H_7S_3$	4052.0		0.008	0.004-0.019	0.013	0.005-0.024	<0.001	**0.77**
$N_6H_7S_3F_1$	4226.1		0.001	0.0003-0.015	0.009	0.002-0.017	<0.001	**0.88**
$N_6H_7S_3F_2$	4400.2		0.001	0.0004-0.006	0.003	0.0004-0.014	<0.001	**0.80**
$N_6H_7S_4$	4413.2		0.010	0.004-0.056	0.026	0.010-0.042	<0.001	**0.83**
$N_6H_7S_4F_1$	4587.3		0.002	0.0003-0.025	0.020	0.003-0.051	<0.001	**0.91**
$N_6H_7S_4F_2$	4761.4		0.001	0.0002-0.015	0.008	0.001-0.047	<0.001	**0.87**
$N_6H_7S_4F_3$	4935.5		0.0001	0.0002.005	0.002	0.0003-0.023	<0.001	**0.79**

Medians in each cluster of study subjects were calculated from the relative peak intensities of the *N*-glycans generated from MALDI-TOF-MS. Min-Max is the minimum and maximum value of the relative intensities of the *N*-glycan peaks in each cluster of study subjects. AUC (area under the curve) values evaluated the potential of each *N*-glycan to discriminate EOC from BOD. P-values were computed using the Mann-Whitney U test, where a value less than 0.05 was considered statistically significant. The *N*-glycan signatures for EOC previously described by our research group are highlighted in bold.

3.1.9 Performance of the *N*-glycan index in discriminating primary African EOC patients from the controls

The *N*-glycan index values of the primary EOC patients ranged from 0.36 - 10.9, with a median value of 3.91, while for the control subjects the *N*-glycan index values ranged from 0.05 - 5.05, with a median value of 0.21 (Table 9). The cut-off for the reference value of the *N*-glycan index was previously set at 0.63 in Caucasian cohorts. This value was adjusted to 1.04, as obtained from the coordinates of the ROC curve in the present study due to adjustments in the methodology. The methodology differences between the current and previous analysis happened in the final analysis,

where two fractions of acidified 25% and 50% acetonitrile elutions of the released *N*-glycans were used in the final analysis, as opposed to the use of only one fraction of acidified 25% acetonitrile elution in the previous analysis. On the contrary, CA125 values of the EOC patients ranged from 30.1 - 1733 kU/L, with a median value of 122 kU/L, while CA125 values for the control subjects ranged from 4.2 - 998 kU/L, with a median of 25.9 kU/L. The upper limit of normal CA125 values was 35 kU/L.

Both the log2 transformed *N*-glycan index and CA125 values showed significantly higher values in primary EOC patients compared to the controls (*N*-glycan index, p= 2.8E-8; CA125, P= 2.0E-6) (as shown in Figures 19A, 19B and Table 6). The *N*-glycan index was a better discriminator of EOC patients from the controls (AUC, 0.94; specificity (SP), 98% and sensitivity (SN), 84%) compared to the CA125 test (AUC, 0.88; SP, 67%, SN, 95%) (Figure 19C and Table 6). A combination of both the *N*-glycan index and CA125 produced similar accuracy and specificity as the *N*-glycan index (AUC, 0.94; SP, 98%) and a sensitivity of 90%.
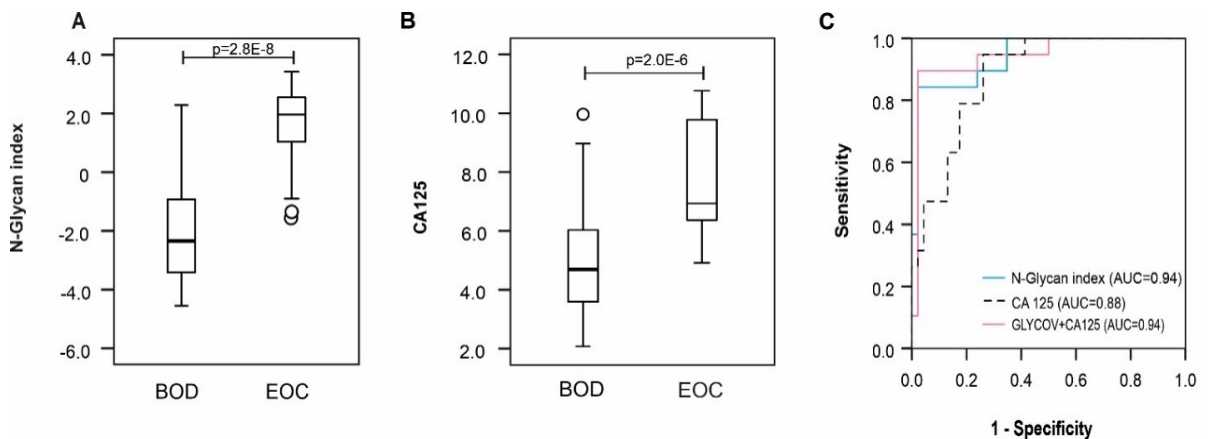
Figure 18. Box plots of log2 transformed values of **(A)** *N*-glycan index and **(B)** CA125 showing their distribution in primary EOC and BOD groups. The data were log2 transformed to reduce the skewness in distribution. The Mann-Whitney U test was used to compare the difference in the distribution of *N*-glycan index and CA125 values within the primary EOC and BOD cohorts (p≤0.001). **(C)** The ROC curve of the *N*-glycan index and CA125 values showing their performance in discriminating EOC from BOD subjects.

Table 9: Comparative performance of the *N*-glycan index against the CA125 in discriminating primary EOC patients from the controls

Statistics	*N*-Glycan index		CA125 (kU/L)	
	BOD (Control)	EOC	BOD (Control)	EOC
Mean ±SD	0.52 ±0.84	3.97 ±2.75	76.3 ±165.3	512.5 ±590.5
Median	0.21	3.91	25.9	122.0
Range	0.05 – 5.05	0.36 – 10.9	4.2 – 998	30.1 – 1733
Diagnostic performance				
Cut-off value	1.04		35	
Sensitivity (SN) (%)	84		95	
Specificity (SP) (%)	98		67	
AUC (95% C.I)	0.94 (0.880 - 1.000)		0.88 (0.801- 0.961)	
Combined *N*-glycan index + CA125				
Sensitivity (SN) (%)		90		
Specificity (SP) (%)		98		
AUC (95% C.I)		0.94 (0.880 - 1.000)		

Data showing the comparison of the *N*-glycan index against the CA125 in discriminating primary African EOC patients from BOD control subjects. The data were presented as mean, SD (standard deviation), range (showing minimum and maximal values of the biomarker in each cohort), AUC (area under the curve, which showed the discriminatory accuracy), sensitivities and specificities of the *N*-Glycan index and the CA125 biomarkers.

3.2 Determination of common serum IgG_1 and IgG_2 glycopeptide signatures for diagnosis of EOC in African and Caucasian ethnicities

This section aimed at finding common glycopeptide signatures from the two ethnic cohorts (African and Caucasian) for complementing CA125 in improving accuracy in the diagnosis of EOC. To begin with, IgG_1, IgG_2, and IgG_4 were isolated from human serum by affinity purification using Protein A Sepharose, which does not bind IgG_3. The captured IgGs were then eluted and tryptically digested. The glycopeptides were then enriched using self-made micro-spin cotton HILIC columns. MALDI-TOF-MS measurements were carried out in the negative ionization mode to measure both neutral and sialylated glycopeptides in a single experiment. The generated mass spectra were exported to the Massy tools software as ASCII text for recalibration, baseline subtraction and peak extraction. A representative MALDI-TOF mass spectrum is shown in Figure 20. The peptide sequences containing the *N*-glycosylation site at Asn-297 were: $E_{293}EQYNSTYR_{301}$ for IgG_1, $E_{293}EQFNSTFR_{301}$ for IgG_2, and $E_{293}EQFNSTYR_{301}$ for IgG_4. A total of 17 glycoforms and 11 glycoforms could be assigned unambiguously for IgG_1 and IgG_2, respectively (Table 8). Unfortunately, signals assigned to IgG_4 glycopeptides had the same mass as some non-fucosylated IgG_2 glycopeptides; as a result, these peaks were excluded from further analyses.
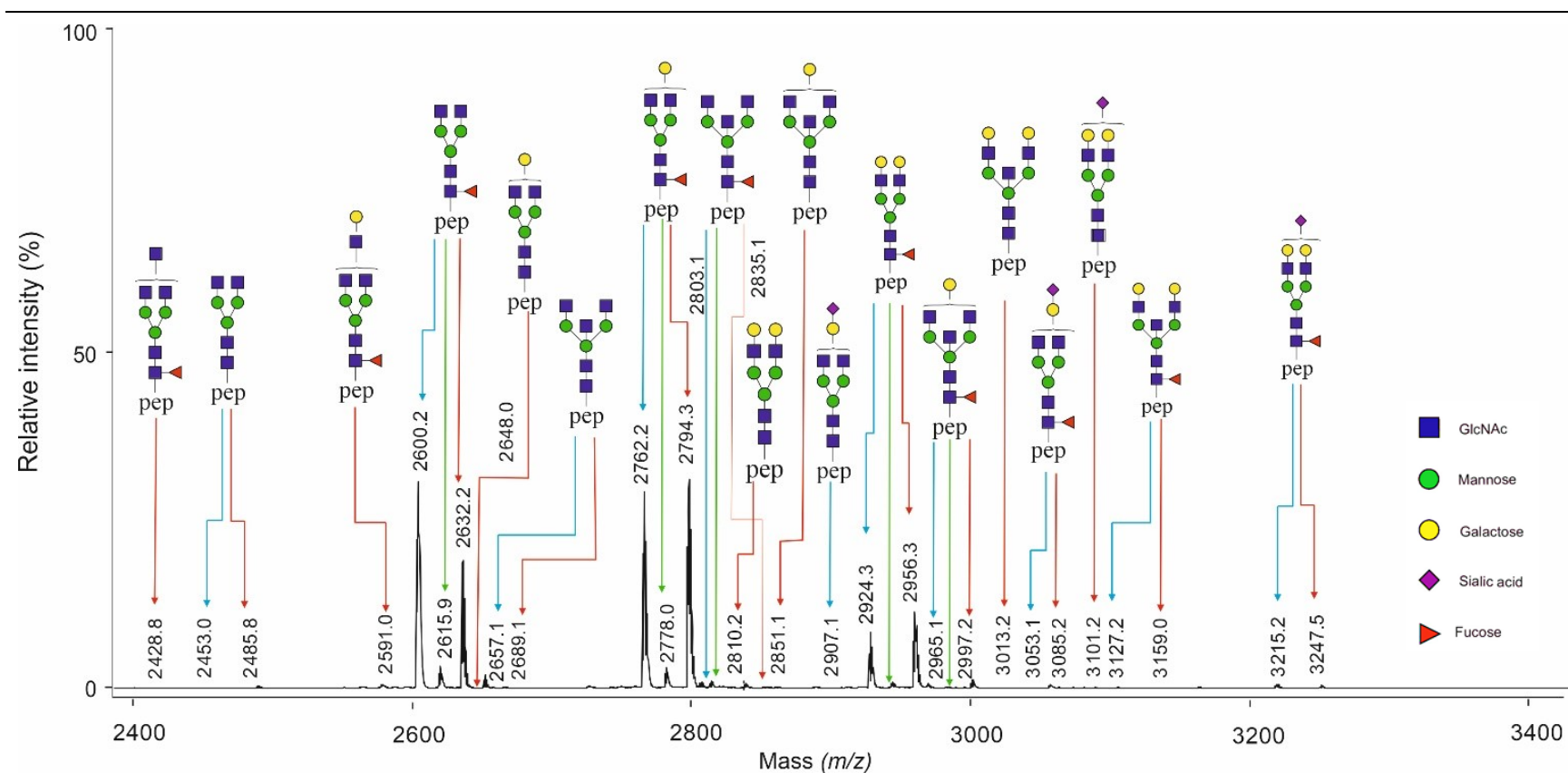
Figure 19. A representative MALDI-TOF mass spectrum of tryptic IgG glycopeptides was obtained from a BOD subject. Red arrows indicate IgG_1, while blue arrows indicate IgG_2 glycopeptide peaks. IgG4 glycopeptides (green arrows) were not analyzed in this study because of their very low serum concentrations, hence producing very low signals which also overlapped with some non-fucosylated structures of IgG_2 glycopeptides which are highly abundant in serum.

3.2.1   Repeatability of the glycopeptide analytical method

Intra and inter-day repeatability were evaluated for the glycopeptide method using one serum sample from a healthy subject. The mean, standard deviation (SD), and coefficient of variation (CV) for the six major IgG_1 and IgG_2 glycoforms, namely G1F, G0F, G0FN, G2F, G1FN, and G2FS1, were calculated (Figure 21). For intra-day repeatability, CVs for both IgG_1 and IgG_2 were less than 10% (very good), as they ranged from 0.4 – 3.8%. Similarly, the inter-day CVs for both IgG_1 and IgG_2 glycoforms were also within the rating of "very good" (1.0 – 7.3%) except for one. The IgG_1 glycoform G2FS1 had 16.7% CV which is rated as "good" (CV <20%) and within the acceptable limit.
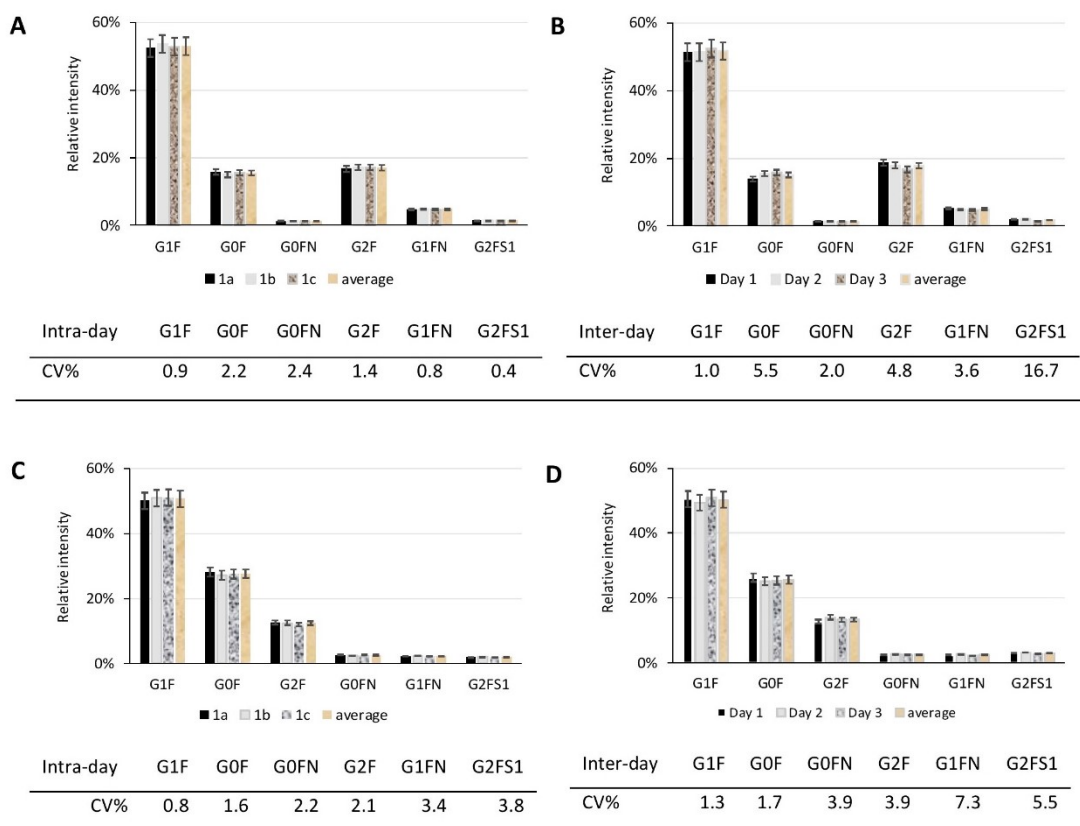
Figure 20. Evaluation of IgG_1/IgG_2 method repeatability using the percentage relative intensities of the six major IgG_1 and IgG_2 glycoforms. Intra-day repeatability is represented by (A) for IgG_1 and (C) for IgG_2, where one sample of a healthy patient was taken through the same analytical procedure in triplicates within the same day. Inter-day repeatability is represented by (B) for IgG_1 and (D) for IgG_2, where the same serum sample as used for intra-day repeatability was analyzed, but once in three consecutive days. Acceptable repeatability (<20%).

3.2.2   The subjects' demographic and clinical factors

The study subjects comprised confirmed EOC and BOD controls of African and Caucasian ethnicities aged 18 years and above. The age of African control subjects' ranged from 18 to 74 years, with an

average of 37 years, while the Caucasian cohort averaged 41 years, with ages ranging from 18 to 75 years (Table 10). However, for the EOC patients of the African cohort, the age ranged from 25 – 80 years, while the Caucasians ranged from 21 – 78 years. The sum of both African and Caucasian BOD cohorts was 68 subjects, where 46 were Africans and 22 Caucasians, while the sum of both African and Caucasian EOC cohorts was 53 patients accounting for 18 Africans and 35 Caucasians. For both ethnicities, the majority of the EOC patients were in the late stage of the disease (FIGO stage III and IV, 83.3% for the African cohort and 74.3% for the Caucasian cohort). The average serum CA125 levels for both ethnicities in their respective categories of BOD and EOC were above the cut-off upper limit of 35 kU/L. However, the median CA125 values for BOD subjects in both ethnicities were below the cut-off limit (African, 25.9 kU/L and Caucasians, 15.5 kU/L), while the median CA125 values for the EOC patients were above the cut-off limit for both ethnicities (African, 202.5 kU/L and Caucasian, 252 kU/L).

Table 10. Demographics and clinical factors of the participants used for IgG glycosylation analysis

Factors	BOD (n=68)		EOC (n=53)	
	African (n=46)	European (n=22)	African (n=18)	European (n=35)
Age (years)				
Mean ±SD	37±11	41 ±16	54 ±14	59 ±13
Median	34	38	48	65
Range (min – max)	18 -74	18 – 75	25 – 80	21 - 78
Interquartile range	28 – 46	29.50 – 50.25	43 – 65	49 - 69
Tumor stage % (n)				
FIGO stage I+II (early stage)	-	-	16.7 (3)	25.7 (9)
FIGO stage III+ IV (late stage)	-	-	83.3 (15)	74.3 (26)
CA125 (kU/L)				
Mean ±SD	77.0 ±165.2	46.55 ±87.8	536.69 ±597.8	1024.4 ±1588.1
Range	4.2 – 998.0	6.0 – 386	30.1 – 1733.3	11.0 – 7343.0
Median	25.9	15.5	202.5	252
Interquartile range	12.1 – 67.15	11 – 40	81.55 – 1064.5	63.45 – 1604.5

Summary of the data of EOC and BOD subjects. Age is presented in years, EOC staging is classified as early and late stage, and CA125 concentrations are presented in kU/L.

3.2.3   Evaluation of the association between galactosylation and age

The serum samples for IgG$_1$ and IgG$_2$ glycopeptide were prepared as stated in section 2.5.2. MALDI-TOF-MS spectra yielded 17 glycopeptide signals for IgG$_1$ and 11 for IgG$_2$. The relative intensities of the detected glycopeptides were normalized to the total intensity for IgG$_1$ and IgG$_2$. As IgG glycosylation is said to be age-dependent and directly related to the menopausal status of women [54, 149], an evaluation of the degree of galactosylation was therefore carried out on the combined control cohort of 68 Africans and Caucasians by classifying them into six age-group clusters (Figure 22). The degree

of glycosylation for each glycopeptide was determined by summing up the relative intensities of all the galactosylated glycoforms of each IgG glycopeptide:

- ◦ Galactosylation (IgG_1) = G1 + G1F + G1FN + G1FS1 + G1N + mono G1F + G2F + G2FN + G2 + G1N + G2N + G2S1 + G2FS1
- ◦ Galactosylation (IgG_2) = G1F + G1FN + G1FS1 + G1S1 + G2F + G2FN + G2FS1

The corresponding boxplots are shown in Figure 19. IgG_1 galactosylation showed near constant values, ranging between 70% - 80% in the age groups of 18-28, 29-38, and 39-48 years. A fall in the degree of IgG_1 galactosylation in the range of 60% - 70% was later observed in the patients that had 49 years and above. Conversely, IgG_2 showed a decrease in the degree of galactosylation as patients age advanced from 18 years and above. However, a steeper decrease in IgG_2 galactosylation was observed between the age groups 39 – 48 and 49 – 54 years, although between ages 55 – 59 years there was a rise in galactosylation, which could be attributed to large effect sizes due to sampling issues within the group. Regarding the present IgG galactosylation data, the study subjects were classified into two age categories, namely 18 – 48 years (young) and ⩾ 49 years (elderly).
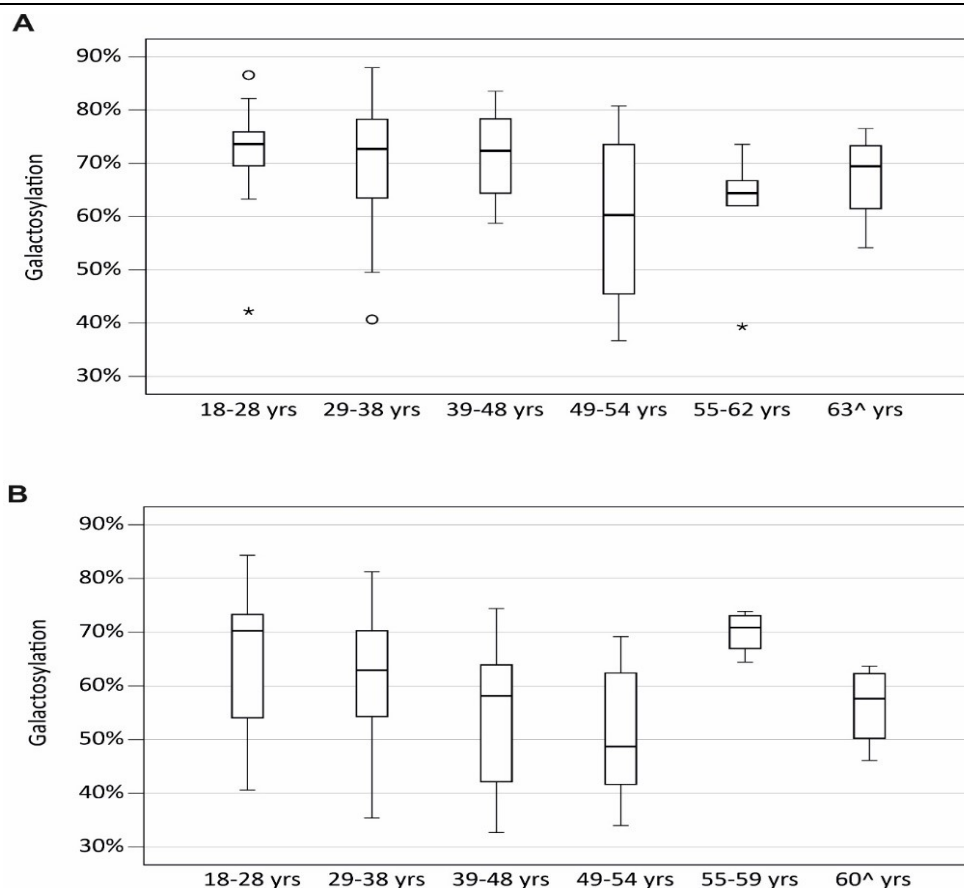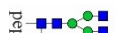
Figure 21. Box plots of the degree of galactosylation according to age in the BOD subjects in (A) IgG_1 and (B) IgG_2 glycopeptides. Galactosylation was largely stable in IgG_1 up to the age of 48 years, then it started to decline. In IgG2, the decrease of galactosylation was gradual from 28 years, but the biggest dip happened after 48 years. The rise in galactosylation in the age group of 55 -59 years is most likely due to the small number of patients (type two error).
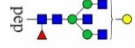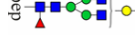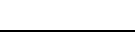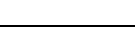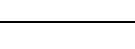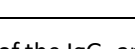
3.2.4 Glycopeptide profiling of IgG$_1$ and IgG$_2$ at Asn-297 in African and Caucasian BOD cohorts and their comparative inter-ethnic expression patterns

IgG$_1$ and IgG$_2$ glycopeptide profiles for age-matched African and Caucasian BOD cohorts were carried out at the Asn-297 level, and their expression patterns were described through a comparative analysis to exclude differentially expressed glycoforms, which would be irrelevant when addressing a common biomarker strategy for EOC diagnosis (Table 11).

Five IgG$_1$ and two IgG$_2$ glycopeptides were identified as differentially expressed between the two ethnicities at a significance of p <0.01. All the five differentially expressed IgG$_1$ glycopeptides had galactosylated moieties and they included the fucosylated bisecting monogalactosyl (G1FN), fucosylated bisecting digalactosyl (G2FN), fucosylated monosialylated monogalactosyl (G1FS1), monosialylated digalactosyl (G2S1) and mono N-acetylglucosamine fucosylated monogalactosyl (mono G1F). Four of these glycans (G1FN, G2FN, G2SI, and G1FS1) were abundantly expressed in the Caucasian compared to the African ethnicity. Conversely, the expression of mono G1F in the Caucasians was lower in abundance compared to the African ethnicity. For the IgG$_2$ glycoforms, the differentially expressed glycoforms were also of galactosylated species. They included monosialylated monogalactosyl (G1S1) and G1FN, which were abundantly expressed in Caucasians compared to Africans. Consequently, the five IgG$_1$ and two IgG$_2$ differentially expressed glycopeptides and were therefore excluded from the subsequent analyses that aimed at finding a common EOC glycopeptide-based biomarker for the two ethnicities

Table 11. Comparison of IgG$_1$ and IgG$_2$ Fc-glycosylation between African and Caucasian benign ovarian disease cohorts

Glycoform	Structure	Mass (m/z)	IgG$_1$ E$_{293}$EQYNSTYR$_{301}$				IgG$_2$ E$_{293}$EQFNSTFR$_{301}$			
			African	Caucasian	p-value	Mass (m/z)	African	Caucasian	p-value
			R. intensity Median (max-min)	R. intensity Median (max-min)			R. intensity Median (max-min)	R. intensity Median (max-min)	
G0		2485.9	0.005 0.002-0.031	0.006 0.002-0.014	0.495	2453.9	0.004 0.001-0.013	0.002 0.001-0.013	0.125
G0F		2632.0	0.260 0.136-0.379	0.185 0.096-0.340	3.6E-3*	2600.0	0.311 0.211-0.579	0.248 0.119-0.497	0.002
G1		2648.0	0.017 0.008-0.055	0.021 0.008-0.036	0.125	-*			-
G1F		2794.1	0.391 0.318-0.565	0.374 0.314-0.423	0.704	2762.1	0.391 0.261-0.598	0.379 0.287-0.499	0.665
G2		2810.1	0.013 0.003-0.040	0.019 0.012-0.033	0.002*	-*			-
G2F		2956.1	0.136 0.038-0.241	0.200 0.069-0.307	0.002*	2924.2	0.130 0.038-0.239	0.175 0.057-0.294	0.008
G0N		2689.1	0.003 0.001-0.010	0.004 0.002-0.008	0.521	2657.1	0.007 0.003-0.026	0.006 0.003-0.015	0.132
G0FN		2835.1	0.035 0.014-0.140	0.035 0.015-0.094	0.521	2803.1	0.043 0.023-0.120	0.043 0.024-0.096	0.896
G1N		2851.1	0.006	0.008	0.004	-*			-

				0.003-0.013	0.006-0.014				
GIFN		2997.2	0.047 0.024-0.096	0.075 0.056-0.144	1.02E-7***	2965.2	0.033 0.017-0.061	0.046 0.028-0.086	8.0E-5***
G2N		3013.2	0.004 0.001-0.008	0.005 0.003-0.007	0.002*	-*			-
G2FN		3159.2	0.006 0.001-0.013	0.012 0.005-0.017	3.0E-6***	3127.2	0.007 0.002-0.018	0.007 0.003-0.020	0.021
G1FS1		3085.2	0.007 0.003-0.013	0.011 0.005-0.020	1.02E-7***	3053.2	0.021 0.011-0.049	0.028 0.017-0.040	0.004*
G1S1					-	2907.1	0.006 0.001-0.013	0.013 0.004-0.020	3.40E-5***
G2S1		3101.2	0.002 0.001-0.009	0.004 0.002-0.011	1.70E-5***	-*			-
G2FS1		3247.2	0.022 0.003-0.066	0.028 0.009-0.055	0.02	3215.2	0.029 0.007-0.100	0.006 0.010-0.091	0.051
Mono G0F		2428.9	0.003 0.001-0.006	0.003 0.001-0.006	0.803	-			-
Mono GIF		2591.0	0.008 0.001-0.170	0.003 0.001-0.012	3.0E-6***	-			-

Comparison of the IgG$_1$ and IgG$_2$ glycopeptide expression between African and Caucasian control subjects. Medians of each glycopeptide in each ethnicity were calculated from the relative peak intensities of all the BOD subjects as generated by MALDI-TOF-MS. Samples were ionized in the [M-H]$^-$ form. The differences in the expression of the IgG$_1$ and IgG$_2$ glycopeptides between the two ethnicities were evaluated by their median peak intensities and p-values. The glycopeptides differentially expressed between the African and Caucasian controls are highlighted in bold letters. The Bonferroni correction (*) was calculated for p–values as follows: IgG$_1$ (0.05/17 = *, 0.01/17 = **, and 0.001/17 = ***), IgG$_2$ (0.05/11 = *, 0.01/11 = **, and 0.001/11 = ***). A pH < 0.01** was considered significant. * Isomeric structures of overlapping masses of IgG$_2$ and IgG$_4$ glycopeptides.

3.2.5   Comparison of IgG$_1$ and IgG$_2$ glycopeptide expression between EOC and BOD subjects based on ethnicity and age

An evaluation of the differential expression of IgG$_1$ and IgG$_2$ glycopeptides was carried out by comparing primary EOC patients against the controls at the levels of inter-ethnic expression and age (for combined cohorts of African and Caucasian ethnicities) (Figure 23). A representative MALDI-TOF mass spectrum of IgG$_1$ and IgG$_2$ glycopeptides as was expressed in a BOD control individual and an EOC patient, both of African ethnicity, is shown in Figure 22.

There was a similar trend in the expression of the major IgG$_1$ and IgG$_2$ agalactosylated glycoform, the G0F, and the two galactosylated glycoforms (G1F and G2F), in BOD and EOC subjects at the analytical levels of inter-ethnic expression and when the two ethnicities are combined (Afro for African/Euro for Caucasian). In both levels of analysis (inter-ethnic and combined ethnicity-age categorization), G0F was abundantly expressed among the EOC patients compared to the BOD subjects. Conversely, the G1F and G2F, which constitute the largest percentage of the galactosylated species, were less abundantly expressed in EOC patients compared to BOD subjects. Other agalactosylated glycoforms that were abundantly expressed among the EOC patients compared to their BOD counterparts were G0N (for both IgG$_1$ and IgG$_2$) and G0 (for IgG$_2$), while the G2FS1 of IgG$_2$ was less in abundance among EOC patients compared to BOD. Glycoforms, which had peak intensities of less than 1%, were not taken into account in further analyses. They included G2N and G2S1 for IgG$_1$, and G1S1, G1FS1 and G2FN for IgG$_2$. Therefore, only ten IgG$_1$ and seven IgG$_2$ glycopeptide peaks were used in the final analysis of determining the common multi-ethnic signature biomarkers for diagnosis of EOC within the age classifications of 18 – 48 years and 49 years and above. They included G0, G0F, G1F, G2F, G0N, G0FN, G1N, G1, G2 and mono G0F for IgG$_1$, while for IgG$_2$ they were G0, G0F, G1F, G2F, G0N, G0FN and G2FS1.
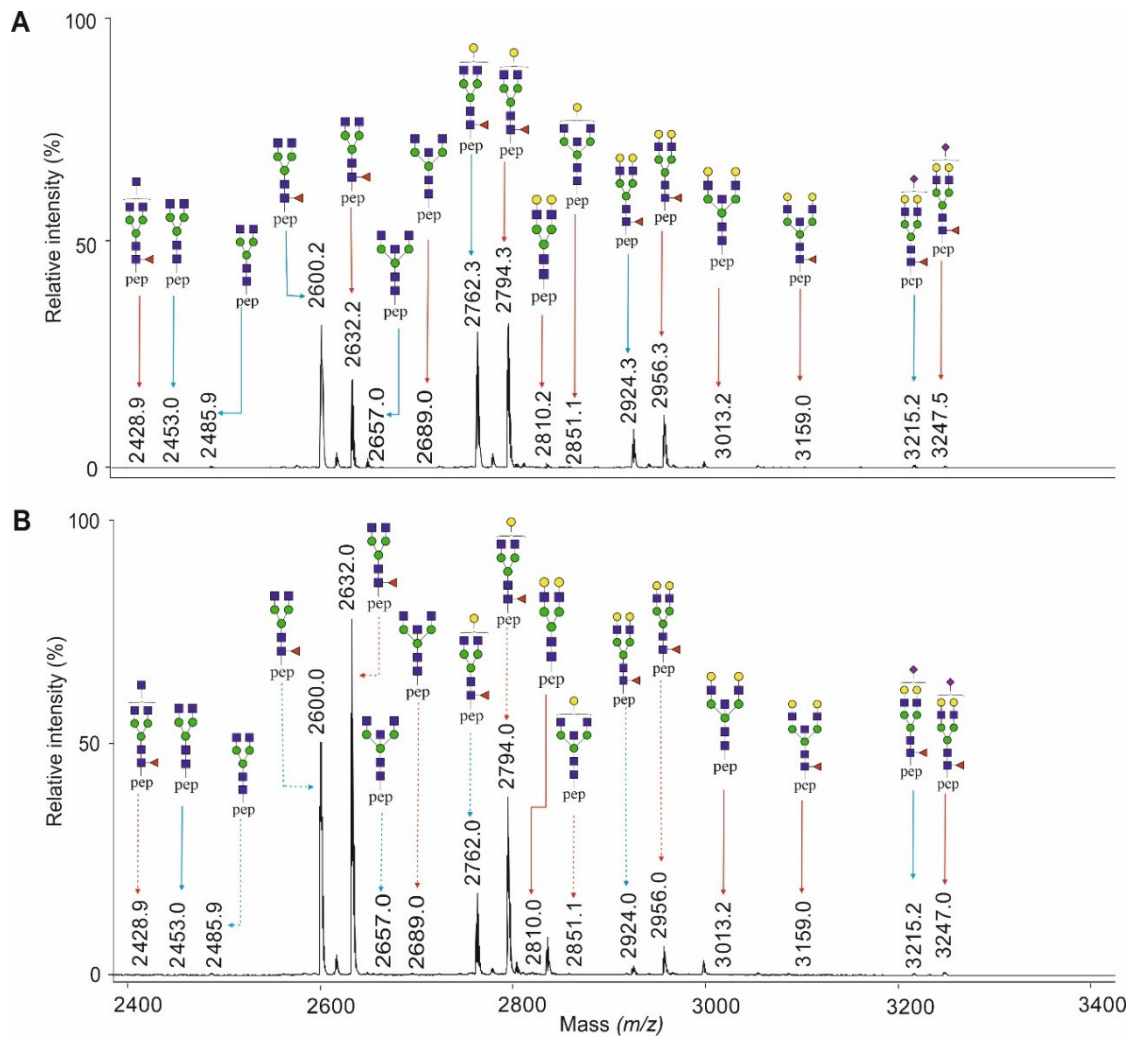
Figure 22. Representative MALDI-TOF mass spectra of IgG$_1$ and IgG$_2$ glycopeptide expression in (A) a BOD subject and (B) an age-matched EOC subject drawn from African cohort. Red arrows indicate IgG$_1$, while blue indicates IgG$_2$ glycoforms. The significantly altered IgG$_1$ glycoforms are marked with red dotted lines, and for IgG$_2$ with blue dotted lines.
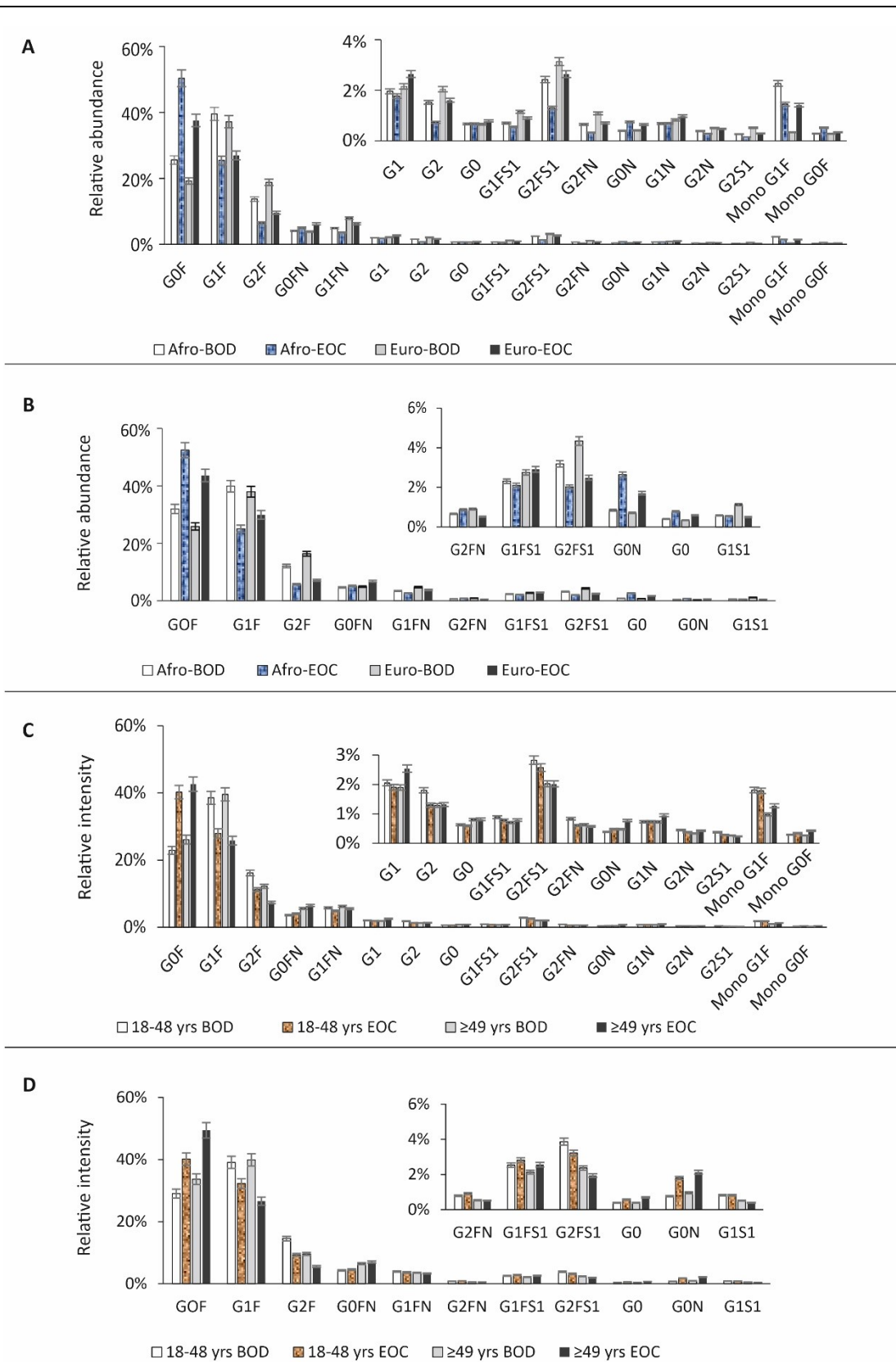
Figure 23. Comparison of the average relative intensities of the IgG$_1$ (A, C) and IgG$_2$ (B, D) glycopeptide expression between BOD and EOC subjects classified according to ethnicities (A, B) and age (C, D). Error bars represent the standard deviation. Afro: African ethnicity, Euro: European (Caucasian) ethnicity.

3.2.6   IgG$_1$ and IgG$_2$ glycopeptide signatures for discriminating combined multi-ethnic cohorts of EOC patients from BOD subjects

Statistical analysis was performed using the Mann-Whitney U test to identify IgG$_1$ and IgG$_2$ glycopeptides that showed significant differential expression between EOC patients and BOD controls in combined African and Caucasian cohorts (Table 12). ROC curves of the relative peak abundances of patients' IgG$_1$ and IgG$_2$ glycopeptides were built, and their corresponding AUC values were used to select the signatures of IgG$_1$ and IgG$_2$ glycopeptides for discriminating EOC patients from the controls. The criteria for selecting a glycopeptide for inclusion as a signature biomarker for the diagnosis of EOC was based on an AUC value of ≥ 0.70 in both young (18-48 years) and elderly (≥49 years) subjects, and a significance of $p < 0.05$.

Evaluation of the elderly subjects yielded seven IgG$_1$ and four IgG$_2$ glycopeptides that could discriminate primary EOC patients from the controls, which were hence possible signatures for diagnosis of EOC. For the IgG$_1$ signatures, they comprised three agalactosylated species [G0F (AUC=0.76), G0N, (AUC=0.81), and mono G0F (AUC=0.80)], in addition to four galactosylated moieties [G1F (AUC=0.70), G2F (AUC=0.71), G1N (AUC=0.78) and G1 (AUC=0.74)]. IgG$_2$ glycopeptide signatures for the elderly subjects included G0 (AUC=0.80), G0N (AUC=0.85), G1F (AUC=0.70, and G2F (AUC=0.71).

In younger subjects, three IgG$_1$ glycopeptides were identified as signatures for discriminating EOC patients from their control counterparts. They included one agalactosylated type, G0F (AUC=0.75), and two galactosylated moieties, G2F (AUC=0.76) and G2 (AUC=0.72). For the IgG$_2$ glycopeptides, five glycoforms, namely G0 (AUC=0.70), G0F (AUC=0.70), G0N (AUC=0.88), G1F (AUC=0.73) and G2F (AUC=0.78), were sufficiently identified as signatures for discriminating younger EOC patients from their control counterparts.

Table 12. IgG$_1$ and IgG$_2$ glycoforms and their discriminative performance of EOC from BOD in a cohort of combined African and Caucasian ethnicities

Glycopeptides	Age	IgG$_1$ (E$_{293}$EQYNSTYR$_{301}$)				IgG$_2$ (E$_{293}$EQFNSTFR$_{301}$)			
		BOD (n=68) Median (R.I.)	EOC (n=53) Median (R.I.)	P-value	AUC	BOD (n=68) Median (R.I.)	EOC (n=53) Median (R.I.)	p-value	AUC
G0	18-48	0.005	0.006	0.856	0.50	0.003	0.006	0.008	0.70
	≥49	0.006	0.007	0.688	0.54	0.004	0.006	0.001*	0.80
G0F	18-48	0.217	0.307	0.003*	0.75	0.303	0.391	0.015	0.70
	≥49	0.277	0.429	0.004*	0.76	0.410	0.499	0.055	0.66
G1F	18-48	0.375	0.331	0.127	0.65	0.378	0.334	0.008	0.73
	≥49	0.372	0.271	2.61E-4***	0.84	0.324	0.278	0.024	0.70
G2F	18-48	0.168	0.117	0.004*	0.76	0.148	0.090	7.89E-4***	0.78
	≥49	0.118	0.064	0.004*	0.76	0.082	0.057	0.017	0.71
G0N	18-48	0.003	0.004	0.251	0.58	0.007	0.014	1.1E-5***	0.88
	≥49	0.004	0.007	7.33E-4***	0.81	0.008	0.016	1.12E-4***	0.85
G0FN	18-48	0.033	0.031	0.856	0.54	0.040	0.044	0.314	0.58
	≥49	0.052	0.069	0.213	0.61	0.059	0.070	0.387	0.57
G1N	18-48	0.007	0.006	0.585	0.52				
	≥49	0.007	0.009	0.005	0.78				
G1	18-48	0.018	0.018	0.933	0.53				
	≥49	0.020	0.024	0.014	0.74				
G2	18-48	0.016	0.010	0.009	0.72				
	≥49	0.013	0.013	0.933	0.52				
Mono G0F	18-48	0.003	0.003	0.158	0.59				
	≥49	0.002	0.003	9.2E-4***	0.80				
G2FS1	18-48					0.036	0.028	0.240	0.60
	≥49					0.021	0.017	0.688	0.52

Comparison of the IgG$_1$ and IgG$_2$ glycopeptide expression of combined African/Caucasian controls against combined African/Caucasian EOC subjects. Medians of the glycopeptides in each age group were calculated from the relative peak intensities of BOD and EOC subjects as measured by the MALDI-TOF-MS. Only IgG$_1$ and IgG$_2$ glycopeptides with similar trends of expression in African and Caucasian controls were considered in this analysis. The differences in the expression of the IgG$_1$ and IgG$_2$ glycopeptides between controls and the EOC patients in each age category were evaluated by their median peak intensities and p-values. The AUC values were used to determine the accuracy levels of each glycopeptide in discriminating the EOC patients from the controls (AUC > 0.70 was deemed acceptable). The p-values of BOD against EOC were obtained by the Mann-Whitney U test followed by Bonferroni correction (*) p-value: IgG$_1$ (0.05/10*, 0.01/10**, and 0.001/10***) IgG$_2$ (0.05/7*, 0.01/7**, and 0.001/7***), and the AUC value for each glycoform as obtained when differentiating EOC from BOD.

3.2.7   Performance of the IgG$_1$/IgG$_2$ glycopeptide signatures in discriminating EOC patients from controls compared to the CA125

The IgG$_1$ and IgG$_2$ glycopeptide signatures for EOC were subjected to ROC curve analysis to assess the discriminative performance of EOC patients from control subjects. The analysis was performed based on age categories, where in each age category two-biomarker sets of both IgG$_1$ and IgG$_2$ glycopeptide signatures were used, as presented in Table 13.

Table 13: The selected IgG$_1$ and IgG$_2$ glycopeptide signature for EOC primary diagnosis

IgG sub-type	Glycopeptide signature sets	
	Young (18 – 48 years)	**Elderly (> 49 years)**
IgG$_1$	G0F, G2F and G2	G0F, G0N, mono G0F, G1F, G2F, G1N, and G1
IgG$_2$	G0, G0F, G0N, G1F and G2F	G0, G0N, G1F and G2F

The criterion for selecting a glycopeptide for inclusion as a signature biomarker for diagnosis of EOC was an AUC value of ≥ 0.70 in both young and elderly subjects.

Evaluation of the category of younger subjects showed that the glycopeptide signature set for IgG$_2$ had better discriminative accuracy of EOC from the controls (AUC = 0.92) compared to the CA125 (AUC = 0.86) (Figure 25). In addition, the IgG$_1$ glycopeptide signature set produced a moderate accuracy (AUC = 0.79) under the category of younger subjects. When the two glycopeptide signature sets (IgG$_1$/ IgG$_2$) and the CA125 were jointly applied to discriminate the primary EOC patients from the controls, improved accuracy was achieved (AUC = 0.93).

Similarly, evaluation of the performance of both glycopeptide signature sets in the elderly subjects showed good discriminatory accuracies of EOC from control subjects (IgG$_1$, AUC = 0.89 and IgG$_2$, AUC = 0.87) compared to CA125 (AUC = 0.83). The discriminative accuracy of EOC patients from the controls improved further when the IgG$_1$/IgG$_2$ glycopeptide signature sets were tested in complement to CA125 (AUC = 0.94).
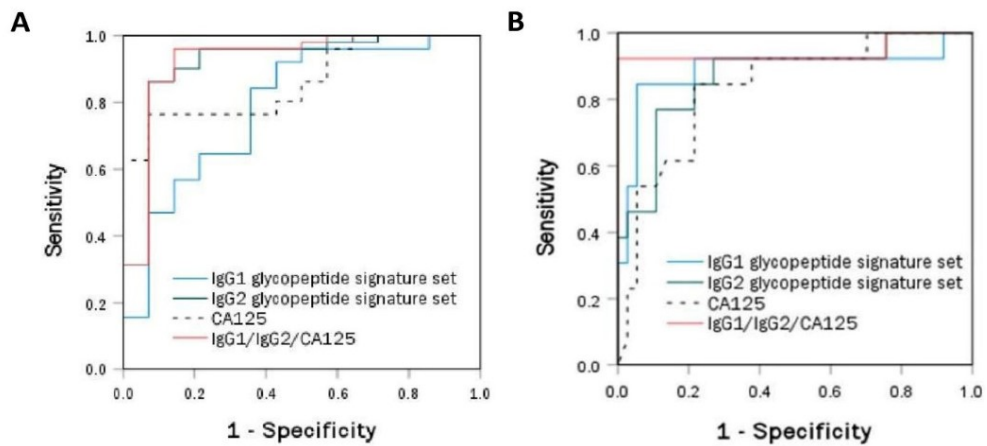
EOC biomarkers	18 – 48 years		≥ 49 years	
	AUC	95% C.I	AUC	95% C.I
CA125	0.86	0.772 – 0.954	0.83	0.710 – 0.961
IgG$_1$ glycosignatures	0.79	0.648 – 0.931	0.89	0.753 – 1.000
IgG$_2$ glycosignatures	0.92	0.820 – 1.000	0.87	0.747 – 0.991
IgG$_1$/IgG$_2$/CA125	0.93	0.830 – 1.000	0.94	0.831 – 1.000

Figure 24. ROC curves of the performance of IgG$_1$ and IgG$_2$ glycopeptide signatures against CA125 in discriminating EOC from BOD based on age clusters. [A] 18 – 48 years and [B] ≥49 years. [C] The corresponding information on AUC values and confidence intervals (CI).

3.3 Determination of HBV, HIV, and HCV seropositivity status in Kenyan ovarian cancer patients

The third question, whose data has already been published in a peer-reviewed journal, sought to determine for the first time the burden of HBV, HIV, and HCV in Kenyan ovarian cancer patients, so as to find out the necessity for routine testing of the HBV, HIV, and HCV in OC patients scheduled for chemotherapy in Kenya. In this analysis, all 86 EOC patients and 50 BOD subjects enrolled in the African cohort were evaluated for their seropositivity status of HBV, HIV, and HCV on the Cobas e 801 immunoassay system. Subsequently, the prevalence of infection of EOC patients was determined as the seropositivity status of HBV (HBsAg), HIV (simultaneous qualitative detection and differentiation of HIV-1 p24 antigen and antibodies to HIV-1 (groups M and O) and HIV-2)), HCV (anti-HCV) and possible coinfection of either formation. A comparison of the HBV, HCV, HIV, and HBV/HIV coinfection seroprevalence rates of OC patients was then made against that of the control subjects.

The findings showed seroprevalence rates of 29.1% (HBV), 26.7% (HIV), and 1.2% (HCV) in OC patients, while the seroprevalence rates for the control group were 3.9% each for both HBV and HIV. None of the control participants tested positive for HCV, while the only coinfection type reported in OC subjects was for HBV and HIV at a seroprevalence rate of 17.4%.

4.0 Discussion

The major findings of the present study are summarized below:

1. The current analysis identified 25 *N*-glycan signatures with the potential for use in monitoring the chemotherapy response. They included 17 *N*-glycan signatures of *m/z* <3600, which showed indications of early response to chemotherapy, and a further eight of *m/z* >3600 as late indicators of chemotherapy response among the African ethnicity. On one hand, the 17 *N*-glycans signatures (*m/z* <3600) indicated response within the first three cycles of chemotherapy through increases in their peak intensities that were initially downregulated in primary EOC. They included three high-mannose (*m/z* 1579.8, 1783.9 and 1988.0), four hybrid-type and monoantennary *N*-glycans (*m/z* 1620.8, 1416.7, 1982.0 and 2390.2) and ten complex-type multiantennary *N*-glycans (*m/z* 2285.2, 2315.2, 2489.3, 2850.4, 1661.8, 2070.0, 2519.3, 2693.4, 2880.4 and 3054.5). On the other hand, the eight *N*-glycan signatures (*m/z* >3600) were indicators for the late response to chemotherapy through their significant decrease of peak intensities from 4 - 6 cycles of chemotherapy from the initially upregulated peak intensities in primary EOC. All of them were complex-type multiantennarity *N*-glycans (triantennary or tetrantennary), with conformational features of sialylation and fucosylation at a peak *m/z* of 3415.7, 3864.9, 3951.0, 4400.2, 3776.9, 4226.1, 4761.4 and 4935.5.

2. In determining *N*-glycan-based signatures of EOC diagnosis at the level of *N*-glycome, a total of 23 *N*-glycan signatures were found to be differentially expressed between primary EOC patients and the controls. Of the 23 *N*-glycans, 12 *N*-glycans of *m/z* <3600 had their peak intensities decreased in primary EOC patients compared to controls and discriminated EOC patients from the controls with excellent accuracies (AUC, >0.90). They included four high-mannose (at *m/z* 1579.8, 1783.9, 1988.0, and 2192.1), two monoantennary (*m/z* 1620/8 and 1982.0), and six complex-type *N*-glycans (*m/z* 2285.2, 2070.0, 2519.3, 2244.1, 2431.2 and 2693.4). The remaining 11 *N*-glycan signatures were of *m/z* >3600, with increased peak intensities in primary EOC. They were complex-type multiantennary *N*-glycans, with features of sialylation and/or fucosylation at a peak *m/z* of 4587.3, 3776.9, 4400.2, 4413.2, 4226.1, 4761.4, 3415.7, 3864.9, 3951.0, 4052.0 and 4935.5.

3. Investigation of EOC diagnostic biomarkers at the level of IgG (IgG1 and IgG2) yielded age-specific glycopeptide signatures in a multi-ethnic cohort of Africans and Caucasians. In the young age bracket (18 – 48 years), the glycopeptide signature sets for EOC were comprised of G0F, G2F and G2 (for IgG_1) as well as G0, G0F, G0N, G1F and G2F (for IgG_2), while for the elderly (from 49 years), the glycopeptide signature sets were G0F, G0N, mono G0F, G1F, G2F, G1N, and G1 (for IgG_1) as well as G0, G0N, G1F and G2F (for IgG_2).

4. An investigation to determine the HIV, HBV and HCV seroprevalence rates among Kenyan OC patients showed a heavy burden of HBV (29.1%) and HIV (26.7%), with a low HCV (1.2%) seropositivity. Higher HBV and HIV seroprevalence were found in Kenyan OC patients compared to the healthy control group, while HCV prevalence was reflective of the general population. Therefore, the heavy burden of HBV and HIV in Kenyan OC patients revealed by the current analysis shows the necessity for routinely establishing the HBV and HIV status in all OC patients scheduled for chemotherapy. This will in turn allow the selection of optimal treatment strategies that suppress the viral load, prevent viral reactivation and curtail tumor growth for patients' enhanced survival and quality of life.

4.1 Glycomic signatures for the diagnosis of EOC and monitoring chemotherapy response

Independent research groups including our own have extensively published on *N*-glycosylation changes associated with ovarian cancer [64, 78-80, 107, 116, 242, 243], but mostly on Caucasian and Asian populations. *N*-Glycosylation changes of high-mannosylation, sialylation, fucosylation, and antennarity have previously been reported in cells, tissues, ascites, and blood samples of ovarian cancer patients [8, 106, 186, 244]. Suffice to say that many of these discoveries have been projected as future biomarkers for the improved diagnosis of many malignancies, prediction of prognosis and disease monitoring among other possible utilities, but are yet to be taken through the full rigor of the validation process for consideration in clinical application. Of all other sources of biomarkers, blood remains the most preferred specimen for biomarker analysis, due to its relative abundance and continuous perfusion through all body tissues including sites of emerging tumors, which generates a rich biological milieu of various substances that are secreted into the blood [245]. Furthermore, its acquisition procedure is minimally invasive, only requiring an insignificant fraction of the whole.

So far, the data on EOC glycosylation changes in African populations is limited in contrast to the Caucasian and Asian ethnicities. To the best of my knowledge, there is no data in existence to date on *N*-glycome changes that arise from the effect of chemotherapy intake and the ovarian tumor process on the total *N*-glycosylation and IgG glycosylation of African EOC patients. To underscore this point, previous studies have shown variations in *N*-glycosylation that are ethnically influenced [246], making it of fundamental importance to undertake ethnic-based glycosylation studies.

Therefore, for the first time, this study addresses the total *N*-glycan changes associated with the ovarian tumor process as well as upon intake of chemotherapy by analyzing EOC and BOD cohorts of African ethnicity to identify signature biomarkers for improving early diagnosis of EOC as well as for monitoring the response to chemotherapy. The study additionally explored for the first time a cross-

ethnic approach, to identifying common IgG_1 and IgG_2 glycopeptide biomarkers that result from changes associated with EOC processes in African and Caucasian ethnicities for improved diagnosis of EOC in the two ethnicities.

4.1.1   Glycome changes in EOC on chemotherapy intake: signatures for monitoring response

The high mortalities associated with OC are not only because of late diagnosis but also due to resistance to chemotherapy as well as its propensity to relapse with resistant tumors after treatment. CA125, the current standard biomarker for monitoring the efficacy of chemotherapy, is unreliable because there is no established standard for interpreting serial measurements of CA125 to predict sensitivity or resistance [247, 248]. The lack of a systematic guide for interpretation of CA125 values in determining the true state of tumor progression during patient management impedes the full realization of the benefits accrued by the hierarchical treatment options of EOC. Effective biomarkers are therefore needed for monitoring the efficacy of chemotherapy agents to facilitate the identification of patients that require reduced dosage or are in need of more aggressive chemotherapy, including even a change of treatment strategy altogether.

Resistance to chemotherapy agents is a major drawback in the quest to attain enhanced survival of EOC patients and improved quality of life [6]. Chemotherapy resistance can emerge from intrinsic factors inherent in cancer cells and causes resistance through mechanisms such as increased drug efflux, inhibition of cellular apoptosis and inhibited uptake of drugs [6, 249]. Other factors causing chemotherapy resistance are acquired in nature and they include genetic as well as epigenetic alternations [249, 250].

The types of glycosylation changes that take place in cancer biological processes are either pro-cancer or anti-cancer. The pro-cancer glycosylation changes are those promoting activities such as evasion of host immune response, inflammation, and adhesion of tumor cells to the endothelium, which in their totality support cancer progression and metastasis [36, 61, 79, 109]. By contrast, anti-cancer glycosylation changes such as bisection are protective against cancer aggression through tumor suppression [8]. Previous studies reported aberrations of glycan structures in various pathological conditions with changes appearing sometime very early in the acute phase of the disease [251, 252] and persisting during the progression of the disease [253, 254]. Glycosylation was also shown to play a role in anti-cancer drug resistance [112, 210], a characteristic that can be exploited for new effective biomarkers for monitoring chemotherapy response. The present study investigated changes in *N*-glycosylation attributed to the intake of combination therapy of carboplatin and paclitaxel by EOC patients. This was achieved by comparing their *N*-glycosylation patterns to another category of EOC patients that had not commenced treatment (primary EOC patients).

Prior to running the study experiments, test runs for the methodology that was used for measuring the permethylated *N*-glycans were carried out and found to be repeatable for both intra- and inter-day analyses. Method repeatability is fundamentally important if a new biomarker is to gain acceptance as well as authorization for application in clinical settings. In this analysis, *N*-glycosylation changes of African EOC patients categorized into the three treatment groups of pre-treatment, 1-3 chemotherapy cycles and 4-6 chemotherapy cycles were investigated for potential signature biomarkers of monitoring response to chemotherapy. A total of 46 permethylated *N*-glycans were detected by the MALDI-TOF-MS from each of the study subjects' serum samples as reflected in their mass spectrum.

A total of 25 *N*-glycan signatures were identified as biomarkers that can monitor response to chemotherapy in EOC patients. Of the 25 *N*-glycan signatures, 17 *N*-glycans of *m/z* less than 3600 and downregulated in primary EOC were identified as signatures for early detection of carboplatin-paclitaxel response within the first three cycles. This conclusion was arrived at because their relative peak intensities were able to discriminate primary EOC patients from patients that had taken between one to three chemotherapy cycles (AUC, >0.70). They included three high-mannoses (*m/z* 1579.8, 1783.9 and 1988.0), one hybrid-type and three monoantennary *N*-glycans (*m/z* 2390.2, 1620.8, 1416.7 and 1982.0), and ten complex-type *N*-glycans, mainly biantennary asialylated and afucosylated (*m/z* 2285.2, 2315.2, 2489.3, 2850.4, 1661.8, 2070.0, 2519.3, 2693.4, 2880.4 and 3054.5). This analysis took note of the decreased peak intensities of the 17 *N*-glycans among the pre-treatment EOC patients as signature biomarkers of monitoring chemotherapy response following their significant progressive increases in categories of EOC patients according to the number of chemotherapy cycles received.

The set of eight complex-type *N*-glycan signatures of *m/z* greater than 3600, that were upregulated in primary EOC patients, distinguished themselves as indicators of clinical response in late cycles of chemotherapy intake. They were complex-type triantennary or tetrantennary *N*-glycans, with features of sialylation and/or fucosylation at *m/z* 3415.7, 3864.9, 3951.0, 4400.2, 3776.9, 4226.1, 4761.4 and 4935.5. The peak intensities of the eight signature *N*-glycans decreased progressively from the category of patients with less chemotherapy cycles (1-3) to the one with the highest number of cycles (4-6), although the most significant decreases took place in the category of 4-6 chemotherapy cycles. These decreases in peak intensities with increments in the number of chemotherapy cycles enabled discrimination of the patients who had taken upwards of four cycles of chemotherapy from those that had not commenced chemotherapy and those that had taken less than three cycles. Consequently, the eight signatures were classified as late markers of chemotherapy response. In summary, the changing expression patterns of the peak intensities of the 25 *N*-glycan

signatures of EOC in the patients on chemotherapy from the reference primary EOC patients are indicative of the glycosylation restorative activity of the carboplatin-paclitaxel anti-cancer agents, and hence, suppression of the cancer activities.

Although very limited data exists on glycomic-based biomarkers for monitoring response to chemotherapy, the available data supports the findings of the current analysis. Evidence from a previous study by Miyahara et al. proposed a high-mannosylated glycan of *m/z* 2010.69 as a potential biomarker for the efficacy of gemcitabine monotherapy treatment in unresectable advanced pancreatic cancer patients as well as for predicting patient survival [255]. Gemcitabine, a nucleoside analog with a broad-spectrum effect on solid tumors, is widely used for managing advanced pancreatic cancer [256]. They identified a high concentration of high-mannose glycan (*m/z* 2010.69) as an independent risk factor of reduced duration of tumor progression and a diminished period of survival. Similarly, the present analysis also identified three high-mannose *N*-glycan signatures (*m/z* 1579.8, 1783.9 and 1988.0) that were among the 17 *N*-glycan signatures that could indicate the early response to/or efficacy of carboplatin-paclitaxel treatment among EOC patients. In addition, Saldova et al. also observed an increase of the otherwise decreased high-mannose (Man$_6$) glycan in primary breast cancer patients upon initiation of neoadjuvant chemotherapy [211]. Their observation was in agreement with the current findings, where the high-mannose (N2H6, *m/z* 1783.9) was equally reduced in primary EOC patients but progressively increased in categories of patients that were receiving chemotherapy based on the number of cycles received. The reduction of the high-mannose intensity in cancer and its subsequent increase in response to chemotherapy among responders suggests its possible involvement with cancer activities. However, the specific role played by the high-mannose in the biology of cancer is yet to be explicitly explained to date, and therefore requires to be addressed by future research. The other 14 *N*-glycan signatures that were identified as early indicators of chemotherapy response exhibited similar physical characteristics as the high-mannose discussed herein. For instance, they were of *m/z* <3600, largely asialylated and afucosylated, and decreased in their peak intensities among the primary EOC patients.

A review of the literature generated from previous findings on the expression of fucosylated and sialylated complex-type *N*-glycans during chemotherapy are concordant with the current findings. Reports from independent research groups associated aberrant complex-type sialylated and core-fucosylated *N*-glycans with monitoring the efficacy of treatment and prediction of drug resistance. As previously reported in various studies, alterations in the expression of sialic acid and fucose are functions of either upregulation or downregulation of the sialic acid transferases, ST6Gal 1, and fucosyltransferases [31, 166]. Schultz and coworkers particularly demonstrated this fact when they found out that tumor cells were sensitized to cisplatin treatment in OC patients when ST6Gal 1 was

knocked out, while the converse was the case when overexpression of ST6Gal 1 caused resistance to the cisplatin therapy [257]. Their findings indicated that increased sialylation in OC can predict cancer resistance, while from the perspective of the current analysis, the reduction of sialylated glycans during chemotherapy signifies sensitivity of tumor cells to the platin and taxen drugs. Hence, treatment of intrinsically platin-resistant EOC individuals with the carboplatin-paclitaxel regime means that the state of upregulation of sialylated signature glycans will continue to perpetuate regardless of testing.

In another independent study, Kamiyama and colleagues reported elevation of the complex-type sialylated and core-fucosylated glycans *m/z* 2890.05 (triantennary) and 3560.29 (tetrantennary) among hepatocellular carcinoma (HCC) patients who had undergone hepatectomy [258]. They associated an increased relative peak of glycan *m/z* 2890.05 with recurrent disease, while glycan *m/z* 3560.29 was described as a significant factor of disease prognosis. Furthermore, they reported that the two glycans were promising biomarkers that can monitor characteristics of the HCC after hepatectomy, including patient survival as well as disease-free survival. Miyahara et al. investigated the same glycans (*m/z* 2890.05 and 3560.29) among hepatocellular patients, and they further demonstrated a correlation between their peak elevations with the progressive disease six weeks after starting Sofarenib treatment [258]. In an interesting turn of events, Miyahara et al. associated elevated levels of glycan, *m/z* 2890.05 with poor survival, and hence a predictor of prognosis, unlike Kamiyama et al. who had associated the glycan at *m/z* 3560.29 with patient prognosis. Their findings on the two glycans (*m/z* 2890.05 and 3560.29) suggested them as promising biomarkers for monitoring the response to Sofarenib among hepatocellular cancer patients.

In both the studies of Kamiyama et al. and Miyahara et al., increased levels of the two complex-type sialylated and core-fucosylated glycans in HCC patients were associated with poor patient survival. Poor patient survival is a consequence of increased sialylation, which promotes tumor progression through increased integrin activity, suppression of galectin-3-induced apoptosis of tumor cells, and immune evasion [111]. Another mechanism through which increased sialylation in cancer contributes to poor survival is by induction of resistance to some anti-cancer agents, as was demonstrated by Schultz et al. [257]. The findings of the two studies on HCC patients with different approaches to patient intervention lend credence to the findings of the current study of new glycan-based biomarkers that may find their application in the clinical monitoring of response to chemotherapy agents in EOC patients. In another study, significant increases of three glycans comprising lewis type biantennary (*m/z* 2401.36), triantennary trisialylated (*m/z* 2986.44) and lewis type triantennary (*m/z* 3086.39) were reported among the platin-resistant patients, unlike in patients who were platin sensitive [210]. These findings are in agreement with the current analysis, which found a reduction

in the peak intensities of the tri- and tetrantennary complex-type glycan signatures among EOC patients, with the same glycans likely to remain elevated in chemoresistant patients.

The Zahradnikova group associated some specific glycans with resistance to platin and taxane chemotherapy using tissue and serum samples of OC patients. In sera, the glycan at peak *m/z* 1595.81 was associated with chemoresistance, while six tissue *N*-glycans of bi-, tri- and tetrantennary structures with features of sialylation and fucosylation were correlated to chemotherapy resistance in OC patients [112]. They included glycans of *m/z* 3053.54, 3880.96, 4055.04, 4128.09, 4603.20 and 4777.39. Their findings showed a difference in the expression of serum and tissue glycans, where each specimen yielded specific glycans that can be applied for the purpose of predicting resistance to chemotherapy among OC patients. Suffice to say that the same glycans can be deployed to monitor response to platin-taxane-based chemotherapy, as they exhibited similar conformation features as the ones described in the current analysis. Another independent research group examined glycosylation changes in OC patients at the level of acute-phase proteins and reported that high sialylated and branched glycans as well as those with sLe^x tended to stay longer in circulation [109]. Relating this data with the findings of the present analysis, especially regarding the ability of the complex-type multiantennary structures to stay in circulation for a long time, explains why many sessions of chemotherapy were needed to impart a significant change in their expression. Hence, they are late indicators of response to chemotherapy, unlike the highly chemotherapy-responsive high-mannoses, whose peak intensities corrected in early cycles of chemotherapy.

Research findings of other independent groups on non-cancer diseases have also shown various potential utilities of glycosylation at the immunoglobulin level. Decreased sialylation of IgG$_1$ was reported as an important precursor of the relapse of granulomatosis with polyangiitis, and hence a significant indicator for pre-emptive treatment [167]. From this finding, measurement of sialic acid levels in polyangiitis could therefore be useful in predicting or detecting a recurrence as well as for monitoring the disease course during treatment. Another study that involved patients with Kawasaki disease and Guillain-Barre syndrome reported a good correlation between the restoration of the serum sialylation of the autoantibodies with the response to treatment [168, 169].

In summary, the findings of the current analysis are supported by data from previous studies on the total glycome and its potential utility in providing new biomarkers for monitoring the efficacy of treatment of many pathologies, including malignancies. The changes in the relative intensities of the 25 *N*-glycan signatures shown in patients on chemotherapy were reflective of the restorative activity of an efficacious treatment strategy. The identified *N*-glycan signatures of *m/z* <3600 were distinctively early indicators of platin-taxane drug response, and their performance in that role was much better compared to the CA125 (AUC, $0.71 - 0.81$ against 0.54). Conversely, the multiantennary

complex-type *N*-glycan signatures of *m/z* >3600 distinguished themselves as late indicators of drug response after the patient has taken upwards of four chemotherapy cycles. Suffice to say that the patterns of response exhibited by the 25 *N*-glycan signatures in the three treatment categories of patients implied their possibility to detect early resistance due to the intrinsic factors as well as resistance due to acquired factors during treatment and in remission.

Important findings were also made on the 11 *N*-glycan signatures, previously described by the Blanchard research group for diagnosis of EOC in Caucasian populations, when integrated into an *N*-glycan index (GLYCOV, as previously described) [64]. The *N*-glycan index carried on board the unique characteristics of each of the 11 *N*-glycan signatures, and hence, as a single marker, it was more robust in determining response both in early and late cycles of chemotherapy as opposed to either the individual glycan or CA125. Another advantage that accrues from using the index as a single marker is the reduced coefficient of variation and increased precision compared to when all the 11 *N*-glycan signatures are used individually. The use of the *N*-glycan index in complementarity with CA125 to determine chemotherapy response produced nearly similar findings to when the *N*-glycan index is used alone. Hence, the current findings support the application of the 11 *N*-glycan signature index to complement CA215 in improving the monitoring of the clinical response of carboplatin-paclitaxel treatment of EOC patients.

4.1.2   Glycome-based diagnosis of EOC: total *N*-glycosylation and IgG glycosylation signatures

The delay in the diagnosis of EOC due to a lack of effective early diagnostic biomarkers is a major contributor to the high mortalities ascribed to ovarian cancers. CA125, the routine biomarker, has the limitations of low sensitivity and specificity, which makes early detection of OC a challenge necessitating the need for new effective biomarkers. Aberrant glycosylation is an established feature of disease development and progression that appears to manifest in all types of human cancers including OC. Whereas it is known that in the normal homeostatic condition the process of protein glycosylation is fairly stable in an individual, the same is disturbed in disease situations or malignancies, resulting in alterations that are disease or cancer-specific. In Hakomori's review, he underscored the genesis of aberrant glycosylation as being a consequence of initial oncogenic transformations and an important step in the start of tumor invasion and metastasis spread [69].

Therefore, the changes in protein glycosylation associated with cancer processes have provided a platform from which new-cutting age biomarkers are hoped to emerge in the endeavor to revolutionize the diagnosis of various cancers, and specifically OC. Many independent research groups including our own have demonstrated altered glycosylation in various cancers at the levels of total glycome, immunoglobulin and acute-phase proteins. In the current analysis, the serum samples

of African and Caucasian cohorts of EOC and BOD subjects were analyzed by MALDI-TOF-MS at the levels of *N*-glycans (only African) and IgG glycosylation (both African and Caucasian), to prospect for biomarkers that will complement CA125 in improving the diagnosis of EOC.

4.1.2.1 Total serum *N*-glycans

Analysis at the level of serum total *N*-glycan identified a total of 23 *N*-glycan signatures which could differentiate the BOD subjects from primary EOC patients. The 23 *N*-glycan signatures were classified into two sets depending on the nature of the alterations in primary EOC. The first set comprised 12 *N*-glycan signatures which had decreased peak intensities in EOC, and also had a *m/z* less than 3600. They included four high-mannose (*m/z* 1579.8, 1783.9, 1988.0 and 2192.1), two complex-type monoantennary (*m/z* 1620/8 and 1982.0), and six bi- or triantennary complex-type *N*-glycans (*m/z* 2285.2, 2070.0, 2244.1, 2431.2, 2519.3, and 2693.4). The second set of *N*-glycan signatures comprised 11 multiantennary (tri- or tetrantennary) complex-type with features of high sialylation and/or fucosylation, which were upregulated in primary EOC patients and were of *m/z* >3600. They included peaks at *m/z* 4587.3, 3776.9, 4400.2, 4413.2, 4226.1, 4761.4, 3415.7, 3864.9, 3951.0, 4052.0 and 4935.5. The problem of late diagnosis of EOC largely arises from misdiagnoses while EOC is still in the early stages. Benign conditions of the ovary are often major causes of misdiagnosis because of their shared signs and symptoms with EOC. Furthermore, the routine biomarker for EOC, CA125, has the limitations of inaccuracies and a lack of specificity, whereby in some BOD cases, abnormal values of CA125 are reported, as was shown in the current analysis. Also, as highlighted in the literature and subsequently shown in the current analysis, some EOC patients (appr. 20%) still exhibit normal values of CA125 (<35 KU/L), causing a delay in diagnosis [201]. The current findings of the 23 *N*-glycan signatures are therefore hoped to complement the CA125 in improving the diagnosis of EOC through precision testing to establish the true state of the patient at a timely stage that allows sufficient time and strategies to mitigate the condition.

The findings of the present study are consistent with previous reports from other independent research groups, which largely carried out their analyses on non-African ethnic cohorts. Many studies using different methodological approaches came to similar findings on the alteration of the high-mannose and hybrid-type *N*-glycans in OC, where their relative peak abundances were found to be decreased compared to either the healthy or BOD controls. By contrast, the multiantennary complex-type *N*-glycans characterized by a high degree of sialylation and/or fucosylation were increased in the EOC patients [64, 79, 107, 116, 126, 259, 260].

The findings of the current analysis on some particular glycans replicated the previous work by Blanchard's research group, which found 11 *N*-glycan signatures as well suited for the diagnosis of

EOC in Caucasian cohorts as well as tumor staging upon collapsing them into an *N*-glycan index (GLYCOV) [64, 107]. The 11 *N*-glycans included four high-mannosylated glycans (*m/z* 1579.8, 1783.9, 1988.0, and 2192.1), which were downregulated in primary EOC patients, and seven multiantennary complex-type *N*-glycans (*m/z* 3776.9, 3951.0, 4226.1, 4400.2, 4587.3, 4761.4, and 4935.5) that were upregulated. The *N*-glycan index, used as a sole biomarker in the current analysis, had excellent accuracy (AUC, 0.94 and a specificity of 84%) compared to the CA125 (AUC, 0.88 and specificity, 67%). Specificity is one of the major concerns regarding CA125 in discharging an accurate diagnosis of EOC as many other conditions, both physiological (pregnancy and menstruation) [261] and with pathological (liver disease and heart conditions) failure [262], present with abnormal CA125 levels. The specificity improved even more when the *N*-glycan index was run on the patients in complementarity with the CA125 (SP, 0.98 and AUC, 0.94). Consequently, the 11 *N*-glycan structures, which were previously described by Blanchard's group as biomarkers for EOC among Caucasian populations, demonstrated sufficiency in discriminating the EOC patients from the controls in the present African cohort. This finding is interesting in light of a report from a past study that reported ethnicity as a factor that influences the way glycans are expressed [246]. The authors found variations in glycan expressions among healthy cohorts drawn from Ethiopian, Japanese, Indian and USA populations. However, the good diagnostic performance of the *N*-glycan index in the African cohort, just as it previously did in the Caucasian cohorts, implies a lack of significant difference in the expression of the 11 *N*-glycan signatures between the two ethnicities, African and Caucasians. Therefore, the current findings justify the use of a common *N*-glycan biomarker strategy for the diagnosis of EOC for the African and Caucasian populations. It is also important to note that the total serum glycome, unlike the IgG glycome, is not significantly influenced by age [263], hence the stratification of patients according to age is not compelling.

Other studies on non-African cohorts that generally agree with the current findings include one by Hua and colleagues, who used a nano-HPLC-chip-TOF-MS method and reported 26 differentially expressed *N*-glycans between the EOC patients and the healthy controls. In their observations, the high-mannose and hybrid structures were downregulated in EOC, while the complex-type *N*-glycans were increased in their abundances [116]. Alley et al. also reported increased abundances of tri- and tetrantennary complex-type *N*-glycans, with varying degrees of sialylation and/or fucosylation in EOC patients, while bisecting *N*-glycans were reduced [79].

Other malignancies in which decreases in the expression of high-mannose *N*-glycans were reported include gastric cancer [68, 107, 116], duodenal cancer [68], and colorectal cancer [264]. Conversely, some studies of breast cancer reported increases in the abundance of some specific high-mannose *N*-glycans. Liu et al. reported a dramatic increase of a high-mannose containing eight mannose (Man_8)

in the tissue of breast cancer patients [114]. However, de Leoz and colleagues, using serum samples of breast cancer patients, reported an increase in the intensity of the high-mannose Man_9 (*m/z* 1905.63) [265]. There is a hypothesis that explains the scenario of upregulation of the high-mannose in cancer to be a result of accumulated precursor structures occasioned by incomplete maturation during the biosynthetic pathway of *N*-glycosylation [266]. In the case of breast cancer, therefore, increased Man_8 in the tissue [114], and Man_9 in the serum of patients [265], could be attributed to an incomplete *N*-glycosylation biosynthetic pathway, where, in normal circumstances, the high-mannoses are trimmed and continue with the maturation process to form the hybrid-type or the complex-type *N*-glycans. Altered expression of high-mannose is a key feature of many cancers, as has already been reported by many studies including the current one; however, the exact role played by the high mannose in tumor etiology is yet to be overtly elucidated. Although Haankesen and colleagues linked the hexamannosylated (Man_6) glycan with decreased mitochondrial fatty acid beta-oxidation [267], Saldova and coworkers suggested that there was an increase in cellular energy production when this glycan decreases in cancer [211].

The increases and decreases of multiantennary complex-type *N*-glycans with features of high sialylation and/or fucosylation among EOC patients in the present analysis is a feature that has not only been reported in many OC studies, but indeed in many other malignancies and pathologies. In a past study, increases of the core-α-1, 6-fucosylated triantennary glycan was reported in breast cancer patients, as well as decreases in bigalacto-biantennary glycan and bigalacto-core-α-1, 6-fucosylated bisecting biantennary [114]. Indeed, conformational features of sialylation and fucosylation expressed on modified glycans in many cancers have been associated with the progression and metastasis of cancers including OC. Increased branching, fucosylation, and sialylation have also been reported in cancers of the prostate and pancreas [81, 268]. Fujita and coworkers also previously found increased fucosylation of the prostate-specific antigen (PSA), for which they indicated a possible utility in the diagnosis of prostate cancer [65]. It is imperative to note that the current clinical biomarker for prostate cancer is a PSA test, albeit measured by immunochemistry as opposed to mass (*m/z*) measurement. Overexpression of Fut8 and core-fucosylation were associated with high metastasis and mortalities in patients with breast cancer [125], and non-small cell lung cancer [124]. However, in gastric cancer patients, expression of core-fucosylation was reported as decreased [129]. To corroborate the evidence of fucosylation and sialylation in breast cancer progression, Kyselova and coworkers showed progressive increases in the degree of sialylation and fucosylation, from a low in the early-stage to a high in late-stage cancer [209]. The alterations in the degree of sialylation and fucosylation that followed the path of advancing cancer also implied their probable use as biomarkers for staging.

The significance of multiantennarity in the biology of cancer also cannot be gainsaid. In another past study, patients with prostate cancer were reported to have high levels of fucosylated biantennary and α2,3-linked sialic acids compared to benign prostate hyperplasia (BPH) subjects. The same patients also exhibited decreased levels of triantennary trigalactosylated glycans and the bisected core-fucosylated biantennary [81]. The authors further correlated tetrantennary tetrasialylated glycans and the decrease in abundance of the triantennary trigalactosylated glycans as well as the bisected core-fucosylated biantennary with the pleural effusion [81]. The correlation with pleural effusion is more likely due to the positive association of decreased bisecting *N*-glycans with the invasiveness of cancer. The role played by bisecting *N*-glycan structures and branching in cancer biology is interlinked, and has so far been well elucidated. Bisecting *N*-glycan structures have been correlated with tumor suppression through inhibition of the further addition of antennas to the biantennary structure [8]. Increases in the expression of branching (complex β- 1,6 linked N-glycans) and MGAT5 are key components of tumor growth, invasion and the formation of metastases [269, 270]. Therefore, a decrease of bisecting *N*-glycans in EOC predicts a poor prognosis for the patient, as it promotes tumor progression and metastasis by default. In the current analysis, the EOC patients were found to exhibit increased multiantennary complex-type *N*-glycan structures and high sialylation as a consequence.

Increased antennarity or branching is at the center of promoting cancer growth. Firstly, through increases of tri- and tetra-antennary structures, which inhibit both the N-acetylglucosamintransferase III (GnT III) and GnT IV enzymes that trigger the formation of the bisecting structures, which are anti-cancer progression. The inhibitory role of the bisecting structures towards more branching is achieved through GlcNAcT-IV, GlcNAcT-V and MGAT5, which catalyzes the transfer of GlcNAc to the growing *N*-glycan to form the tri- and tetraantennary structures. Nonetheless, the multiantennary *N*-glycans negative response towards the formation of the bisecting structures is by inhibiting the GlcNAcT-III [69, 79]. The second pathway through which increased antennarity promotes cancer growth is through providing more sites for terminal sialic acids, and hence increased sialylation, which has been associated with aiding the spread of the malignant cells [103].

Although high sialylation has been associated with cancer progression and metastasis [104, 105], the information that presently exists about the molecular changes that result in increased sialylation during such tumor events is still scarce [77]. Increased sialylation in cancer is a consequence of multiple factors that include upregulation of the sialyltransferases, increased availability, and possibly, reduced sialidases, which are supposed to cleave terminal sialic acids [101, 271, 272]. Deviation from a normal sialic acid expression is linked to the alterations that take place in the epithelial-mesenchymal transition, which is an essential enabler for the invasion of the nearby tissues

by the tumor cells, and hence, the spread [77]. The events of tumor progression and metastasis arise from the modification of cell signaling and adhesion, upregulation of integrin, cadherins and matriptase activities, and suppression of galectin-3-induced apoptosis as modulated by MGAT5 [101, 273]. Others include the induction of resistance to some chemotherapy agents, and evasion of tumor cells from the attack by the immune system [103-105, 111].

Increased sialylation in cancer includes upregulation of sialylated derivatives of Lewis antigens (sialyl-lewis x, sLe^x and sialyl-lewis a, sLe^a), which are known to promote metastasis as they are the ligands of selectins [100, 101]. Several cancer studies have reported increased sialylated derivatives of Lewis antigens. In a study that involved colorectal cancer patients, an increase of sLe^a expression was reported in the patients' lymph node metastases in comparison to the primary lesions, while also being correlated with relapse as well as poor patient outcomes [134]. Moreover, in a separate study involving lung cancer patients, increased highly-sialylated glycans and sLe^x were reported [67]. Sialyl lewis x, (sLe^x) and sLe^a are the most common terminal glycans, which are the ligands of selectins whose interaction causes the formation of aggregates of tumor cells and platelets in the blood vessels, thereby accelerating tumor cell migration and disease progression through extravasation [61, 274]. Nakayama et al. correlated increased expression of sLe^a with tumor metastases in patients with colorectal cancer [134], while in another study of colorectal cancer, a shortened survival of patients that expressed an increase of both Le^x and sLe^x epitopes was demonstrated [131]. This is because expression of both Le^x and sLe^x provides the tumor cell with an enhanced, conducive microenvironment for cancer spread, which is incompatible with prolonged survival. Additionally, in another study involving breast cancer patients, increased terminal fucosylation and sLe^x expressions were reported and correlated with poor patient prognosis [122].

A study that carried out immunohistochemical examination of the sLe^x and sLe^a expression in esophageal, gastric, pancreatic and colorectal cancer patients found an association between the aggressiveness of esophageal, gastric and colorectal cancers with the expression of either sLe^x or sLe^a positive tumors compared to the patients with negative tumors. Pancreatic cancer showed no association [62]. Furthermore, the study also reported the poor survival of esophageal and colorectal cancer patients who expressed positive sLe^x tumors, and also gastric cancer patients who had sLe^a positive tumors, when compared to patients with negative sLe^x or sLe^a tumors [62]. These findings show the possibility of applying sLe^x and sLe^a as biomarkers for predicting the prognosis of patients with esophageal, gastric and colorectal cancers. Another study also used an immunohistochemical approach to analyze sLe^x antigens and associated the incidence of positive sLe^x with the severity of malignancy, recurrence and poor survival of colorectal cancer patients. The authors demonstrated a positive correlation between sLe^x presence and the depth of tumor invasion, lymphatic invasion,

lymph node metastasis as well as the disease stage [61]. From these findings, it can be deduced that applications can be found for sialylation derivatives of sLex in predicting patient prognosis and detecting relapse immunohistochemically.

4.1.2.2 Serum IgG$_1$ and IgG$_2$ glycopeptides

The second level of evaluation of the glycomic-based diagnostic biomarkers for EOC was carried out on the IgG glycome by measuring serum samples of African and Caucasian EOC and BOD cohorts with MALDI-TOF-MS for the IgG$_1$ and IgG$_2$ glycopeptides. This analysis sought for the first time to find altered IgG$_1$ and IgG$_2$ glycopeptides in African and Caucasian EOC patients that can be exploited as a common strategy to improve the accuracy of EOC diagnosis within the two ethnicities. The report of a previous study showed inter-ethnic variations of total *N*-glycome [246], while the process of glycosylation process within an individual remains stable in normal homeostatic conditions [56]. At the IgG glycome level, variations in glycosylation have been reported to occur in the physiological states of aging and pregnancy or in cases of homeostatic perturbations resulting from diseases and cancer processes [53]. Therefore, to develop an IgG glycome-based biomarker that commands wider approval and broad-spectrum utility, factors responsible for IgG glycosylation variations, such as diversities of ethnicities/populations, age and geographical areas, are fundamentally important to be taken into account. For instance, Parekh and colleagues reported for the first time the dependency of IgG galactosylation and agalactosylation on age, observing that IgG agalactosyl (G0) glycoform decreased from birth, reaching the minimum level at the age of 27 years, and then increased with age. They also demonstrated a positive correlation of IgG digalactosyl (G2) with age, which then varied inversely to IgG-G0, while IgG monogalactosyl (G1) remained constant from birth [54].

Therefore, to verify the impact of the age factor on IgG galactosylation, an analysis of IgG$_1$ galactosylation traits for the combined cohorts of African and Caucasian control subjects was carried out using the Wieczorek et al. approach [82]. The degree of IgG$_1$ galactosylation was relatively stable between the ages of 18 to 48 years, albeit with minor decreases, but the biggest dip in galactosylation happened between the ages of 39 - 48 years and 49 – 54 years. Moreover, IgG$_2$ galactosylation progressively decreased with age, with the biggest dip in the degree of galactosylation also taking place between the ages of 39 - 48 years and 49 – 54 years. Consequently, the study subjects were broadly classified into two age groups (18 – 48 years and ≥49 years) for the follow-up analyses. The trend of IgG galactosylation in the current analysis agrees with the findings by Parekh et al. on the positive relationship between age and IgG galacosylation. The subsequent studies that confirmed the findings by Parekh et al. are also in agreement with the present findings on the correlation of age as well as IgG galactosylation and agalactosylation [138, 149, 275, 276]. Yamada and coworkers, for instance, not only affirmed the association of age and galactosylation, but also demonstrated the

differences in the degree of galactosylation with sex. They reported higher levels of IgG-G0 glycoform in males of the same age as females at an approximate age of 25 years [276].

In the follow-up analysis, which involved identifying IgG glycopeptide biomarker signatures of EOC, the subjects were classified as "young" (18 - 48 years) and "elderly" ($\geq$ 49 years). Incidentally, the age of 48 years, which denoted the cut-off for the young subjects in the current analysis, is what a previous study had reported as the median menopausal age in Kenya [277], and hence a relative link between the degree of galactosylation and menopausal age. Kristic et al. previously established a link between estrogen hormone and the increase of galactosylation in pregnancy as well as the decrease in menopausal age [149]. Other researchers also correlated pregnancy with galactosylation, where expression of IgG-G2 and IgG-G2S1 was shown to increase with the term of pregnancy, reaching the peak at week 30, and then decreasing in postpartum [150, 151].

The second evaluation of IgG_1 and IgG_2 glycosylation sought to find differences in the expression of the glycopeptides between the age-matched African and Caucasian control subjects. Five IgG_1 glycopeptides (G1FN, G2FN, G2SI, G1FS1 and mono G1F) and two IgG_2 glycopeptides (G1S1 and G1FN) were found to be profoundly expressed in the Caucasian population compared to their African counterparts. Conversely, mono G1F was less abundant in the Caucasian subjects compared to the Africans. The differences in the expression of some of the IgG_1 and IgG_2 glycopeptides between the two ethnicities point to the existence of inter-ethnic variations in IgG Fc glycosylations, although to the best of my knowledge, no study has outlined ethnic-influenced variations in IgG glycosylation. In a study that may give a closer perspective, Dard et al. previously reported allotypic variation of amino acid in the peptide sequence of IgG_3 glycopeptide at the N-terminal of Asn-227, which causes differences in masses that are ethnic-specific [229]. The allotypic variations of IgG_3 may cause masses that are identical to IgG_4 sequence (EEQYN-STFR), which is the predominant allotype in African and Asian ethnicities, or may also cause the IgG_2 sequence (EEQFN-STYR), which is predominant among Caucasians [229]. Therefore, to arrive at common IgG glycome-based biomarker strategies for use by multi-ethnic groups, it is important to carry out a full screen of the IgG glycome profile in as diverse populations as possible, and document their similarities and differences. The glycopeptides that are differentially expressed between ethnicities or populations should then be excluded in the final analysis that aims to arrive at the common biomarker strategy. Large studies need to be commissioned in the future to corroborate the findings of the current analysis concerning inter-ethnic variations in the expression of specific IgG glycoforms.

The third finding resulted from the analysis of the differential expression of IgG_1 and IgG_2 glycopeptides between combined African/Caucasian EOC patients and control counterparts as the basis for new diagnostic biomarkers for EOC. It was found that the IgG_1 and IgG_2 glycopeptides with

agalactosylated moieties had higher relative abundances in both young and elderly EOC patients compared to their age-matched controls. Conversely, the relative intensities of the galactosylated species of IgG_1 and IgG_2 glycopeptides were decreased in both young and elderly EOC patients compared to the controls. The current findings are in keeping with reports from several other independent research groups including our own whose study populations were of non-African ethnicity [83, 87, 212]. For instance, Wieczorek et al. reported an increase in the degree of IgG_1, IgG_2, and IgG_3 agalactosylation as well as a reduction of galactosylation in a Caucasian cohort of EOC patients compared to the age-matched healthy women. The authors also found that, of the three IgG subclasses, IgG_1 had the greatest reduction of the monogalactosylated and digalactosylated glycopeptides [82], while in another study, Theodoratou and colleagues reported decreases in galactosylation and sialylation of the fucosylated IgG glycan structures as well as increased bisecting GlcNAc in IgG of colorectal cancer patients. These changes in IgG glycosylations were associated with all-cause and colorectal cancer patient mortalities [84]. An increase in the bisecting GlcNAc is a sign of good prognosis for the cancer patient, as it is protective of the individual through suppression of the tumor. Conversely, the decrease of galactosylation and sialylation results in a pro-inflammatory IgG phenotype, which enhances the chronicity of inflammation and cancer progression leading to a poor prognosis.

In a separate study, it was reported that increased serum levels of IgG agalactosylated glycans in OC patients were partly because of depressed Gal-T activity in the plasma cells and/or underexpression of galactosyltransferases [109], while the amounts of galactosylated biantennary glycans decreased due to the downregulation of galactosyltransferases. [109]. These changes in glycosylation were associated with the chronic inflammation evident in OC which supports the invasion of the tumor and metastasis.

The present analysis also showed that the major agalactosylated IgG glycopeptide, IgG-G0F, was profoundly increased in IgG_1 compared to the IgG_2 sub-class in both young and elderly EOC patients, but significantly increased in each case compared to their matched controls. The biological implication of increased agalactosylation in malignancies has already been elucidated as far as the promotion of pro-inflammatory responses is concerned.

The high binding affinities of agalactosylated IgG glycoforms to the FcγRIII to induce a pro-inflammatory response arises from their lack of sialic acid or galactose residues on the glycan moiety [139, 158]. The role of sialic acids in immunity is to work as the molecular switch, which turns on an anti-inflammatory response (when terminal sialic acid/s are present), and off (pro-inflammatory response) when terminal sialic acid/s are absent [138, 165]. A gastric cancer study previously correlated increases in sialylation as well as bisecting GlcNAc with better patient survival. They also

reported increased fucosylation in stage II and stage III cancer patients, suggesting its association with cancer progression [142]. Increased agalactosylation is thought to be a consequence of decreased expression of galactosyltransferases [109, 159]; however, it is still not clear whether the control mechanism of galactosyltransferases happens at the transcriptional or the post-transcriptional level. Agalactosylated IgG glycoforms also trigger pro-inflammatory activities through the activation of the complement system through the alternative and lectin pathways [158, 278, 279].

IgG glycosylation change that involves decreases in the degree of galactosylation has also been reported in other malignancies, including multiple myeloma [154], prostate [155], gastric [142], lung [157], and breast cancers [146]. Decreases in the galactosylation of IgG were also reported in patients with autoimmune diseases such as rheumatoid arthritis [57, 58] and systemic lupus erythematosus [59]. van Zeben et al. associated the progression of rheumatoid arthritis characterized by joint erosion with increased expression of IgG-G0, suggesting its probable use in indicating the disease path of the patients [280]. In another rheumatoid arthritis study, the pregnancy-related increase in galactosylation was correlated with improvement in rheumatoid arthritis among the study subjects, while the subsequent deterioration of rheumatoid arthritis was linked to the decrease of galactosylation in postpartum [150]. In summary, the authors demonstrated that the changes in the IgG galactosylation trait during pregnancy are beneficial to rheumatoid arthritis patients, but the successive galactosylation change that takes place upon delivery are deleterious to the gain. It is thought that the galactosylation changes caused by pregnancy in RA patients are either cytokine or hormonal induced. Previously, it was observed that IL-6, estrogen and/or prolactin induce glycosytransferases which are responsible for galactosylation [150]. The increases in agalactosylation and the decrease in galactosylation have various biological implications. IgG Fc-glycans are necessary for the effector functions of the antibody, which include pathogen binding at the Fc receptor (FcR) and subsequent elimination [281]. However, diseases and neoplastic processes are known to upset the order of the normal glycosylation process, leading to the alteration of the effector functions of the IgG.

In the current analysis, EOC-mediated alterations of IgG_1 and IgG_2 Fc glycosylation in EOC patients were demonstrated, which culminated in the identification of age-dependent signature glycopeptide sets that had good discriminatory attributes of EOC patients from the controls. For the young mixed cohorts of African/Caucasian subjects (18 - 48 years), the glycopeptide signature set of IgG_1 comprised three glycopeptides (G0F, G2F, and G2), while IgG_2 glycopeptide signature set had five (G0, G0F, G0N, G1F, and G2F). Generally, a combination of the two glycopeptide signature sets and CA125 improved the discrimination of EOC patients from the controls (AUC, 0.93) compared to CA125 (AUC, 0.86) when used alone. Additionally, seven glycopeptides (G0F, G0N, mono G0F, G1F, G2F, G1N and

G1) formed the IgG_1 glycopeptide signature set for the elderly (≥49 years) African/Caucasian subjects, while the IgG_2 glycopeptide signature set had four (G0, G0N, G1F and G2F). When the two IgG_1 and IgG_2 glycopeptide signature sets were combined with the CA125, better discriminatory performance of EOC patients from the controls was achieved (AUC, O.94) compared to when CA125 was used alone (AUC, 0.83). Therefore, the current analysis demonstrated the potential utility of the identified age-specific IgG_1 and IgG_2 glycopeptide signatures of EOC in complementing CA125 to improve the accuracy of diagnosing EOC in the two ethnicities, African and Caucasians, considered herein.

4.1.3   The burden of HBV, HIV, and HCV in Kenyan ovarian cancer patients

The discussion on this subject is hereby presented in a summary form because, as stated earlier, the data has already been published in a peer-reviewed journal. The current analysis of HBV and HIV seroprevalence among the Kenyan EOC patients showed higher rates compared to the healthy control group, while HCV prevalence was reflective of the general population. Previously, it was reported by other studies that if infections of HIV and HBV among EOC patients are not accounted for during anti-cancer treatment, it has the negative impact of further reducing the patients' longevity [4]. The heavy burden of HBV and HIV in Kenya OC patients demonstrated in the current study, therefore, necessitates routine testing for HBV and HIV in all OC patients scheduled for chemotherapy. This will allow the selection of optimal treatment strategies to deal both with the viral load as well as preventing further growth of the tumor.

4.2 Study limitations and future perspectives

1. The present study did not have a cohort of EOC patients that were non-responders to carboplatin-paclitaxel, hence we did not have the benefit of comparing the expression patterns of the selected *N*-glycan signatures between the group of responders against non-responders. Nevertheless, the pre-treatment group was assumed as the alternative, whose glycan expression pattern, in the thinking of the present analysis, could be similar to that of non-responders. However, this view needs to be further corroborated or disputed by conducting a comparative large follow-up study that is blind to responders and possible non-responders to monitor the expression patterns of the *N*-glycans, so as to detect the point at which resistance emerges among those patients with acquired chemotherapy resistant factors as well as the inter-patient variations in sensitivities and insensitivities during treatment.

2. In the present analysis, the glycome-signature biomarkers were not assessed at the level of tumor stages to show their performance in correctly assigning EOC patients in their respective stages in comparison to the CA125 due to the limitation of sample size. However, some previous studies already described glycan changes at the level of tumor staging in

Caucasian cohorts of EOC [80, 107], although such information is not available yet for the African populations. Therefore, in the future, it will be necessary to conduct a large study that will evaluate *N*-glycosylation at the level of tumor stages among the African ethnicities to determine the differences or similarities with respect to previous Caucasian results.

3. It is necessary to carry out a large multi-ethnic study of healthy cohorts that also includes Asian ethnicity to describe similarities or variations in the expression patterns of the whole total *N*-glycome and immunoglobulin glycosylation to guide future research in the area of biomarker discovery.

5.0 Conclusion

5.1 *N*-Glycan signatures for monitoring response to chemotherapy

Seventeen *N*-glycan signatures of *m/z* <3600 were good indicators of early response (within the first three cycles) to carboplatin-paclitaxel combination therapy. They included three high-mannose (*m/z* 1579.8, 1783.9 and 1988.0), four hybrid-type and monoantennary (*m/z* 1620.8, 1416.7, 1982.0 and 2390.2) and ten complex-type multiantennary *N*-glycans (*m/z* 2285.2, 2315.2, 2489.3, 2850.4, 1661.8, 2070.0, 2519.3, 2693.4, 2880.4 and 3054.5). Then again, a further eight *N*-glycan signatures of *m/z* >3600 were good late indicators (from four cycles) of carboplatin-paclitaxel combination therapy in the African ethnicity. The eight N-glycans of high masses were all of complex-type multiantennary (triantennary or tetrantennary) with high sialylation and fucosylation (*m/z* 3415.7, 3864.9, 3951.0, 4400.2, 3776.9, 4226.1, 4761.4 and 4935.5). The 11 *N*-glycan signature index used, as described, was more robust under Caucasian ethnicities in indicating glycosylation changes in both early and late chemotherapy sessions compared to either CA125 or individual *N*-glycan signatures. Therefore, the *N*-glycan index would be suitable for EOC patient management through improved monitoring of the patients' response to the carboplatin-paclitaxel therapy among the African ethnicities.

5.2 *N*-Glycan signatures for EOC diagnosis

A total of 23 *N*-glycan signatures were found to be differentially expressed between primary EOC patients and the controls and having the potential for use as EOC diagnostic biomarkers. They included 12 *N*-glycans of *m/z* <3600 (*m/z* 1579.8, 1783.9, 1988.0, 2192.1, 1620/8, 1982.0, 2285.2, 2070.0, 2519.3, 2244.1, 2431.2 and 2693.4), which were significantly downregulated in EOC patients compared to the controls, while the remaining 11 *N*-glycan signatures of *m/z* >3600 (*m/z* 4587.3, 3776.9, 4400.2, 4413.2, 4226.1, 4761.4, 3415.7, 3864.9, 3951.0, 4052.0 and 4935.5) were upregulated in EOC patients. Comparative analysis of the 11 *N*-glycan signature index used, as described, showed better discriminatory performance under Caucasian ethnicities of EOC patients from the controls compared to CA125. The performance characteristics of accuracy and specificity improved further when the index was combined with CA125 to differentiate EOC from controls. Therefore, it can be concluded that the *N*-glycan index can supplement CA125 to improve EOC diagnosis in both European and African ethnicities.

To summarize sub-sections 5.1 and 5.2, the versatility and robustness exhibited by the *N*-glycan index has shown its potential as a cutting-edge biomarker for the diagnosis of EOC in African and Caucasian ethnicities, as well as for monitoring patients' response to chemotherapy in complementarity with CA125.

5.3 IgG$_1$ and IgG$_2$ glycopeptide signatures of EOC diagnosis

1. Analysis of IgG$_1$ and IgG$_2$ glycopeptides revealed differentially expressed glycoforms between the control group of African and Caucasian ethnicities. They included five IgG$_1$ glycopeptides, of which G1FN, G2FN, G1FS1, G2S1 had higher abundance in Caucasians, while mono G1F was higher in abundance among African ethnicity. For IgG$_2$ glycopeptides, G1S1 and G1FN were abundantly expressed in the Caucasians compared to the Africans.

2. Analysis of the IgG$_1$ and IgG$_2$ glycopeptides of combined control samples of African and Caucasians revealed that the degree of galactosylation was age-dependent and reduced as one approached old age. However, the biggest fall in galactosylation happened between the age of 48 and 49 years.

3. It was found that some age-specific glycopeptides could differentiate primary EOC patients from the BOD subjects when the two ethnicities, African and Caucasians, are investigated. The glycopeptides identified in the category of young subjects (18 – 48 years) included G0F, G2F and G2 (for IgG$_1$ glycopeptide signature set) and G0, G0F, G0N, G1F and G2F (for IgG$_2$ glycopeptide signature set), while for the elderly (from 49 years), they included G0F, G0N, mono G0F, G1F, G2F, G1N, and G1 (for IgG$_1$ glycopeptide signature set), and G0, G0N, G1F and G2F (for IgG$_2$ glycopeptide signature set). Analysis of the glycopeptide signature sets under their respective age categories showed improved discriminatory performance (accuracy) of EOC patients from the controls compared to CA125. Further improvement of accuracy in correctly assigning EOC patients and the controls into their respective groups was achieved when the glycopeptide sets were evaluated alongside CA125.

5.4   The prevalence of HBV, HCV and HIV infection in Kenyan ovarian cancer patients

1. It was concluded that Kenya has a higher prevalence of HBV and HIV seroprevalence in ovarian cancer patients. The high HBV seropositivity rate reported could be a consequence of active infection as well as from the relapse of previously resolved cases, due to chemotherapy intake following a lack of routine testing of HBV infection among the OC patients to guide on utilization of the optimal treatment options.

2. The findings of the present study have demonstrated the imperativeness of routinely testing all EOC patients scheduled for anti-cancer treatment to guide clinicians on the optimal treatment strategies that address both the virus and cancer cells to enhance patient survival.

3. The seroprevalence rate of HCV was low and mirrored that of the general population, and was hence not compelling as a routine test on OC patients.

6.0 Implications for practice and future research

The cause of early deaths in OC is in the majority due to late diagnosis following a lack of effective biomarkers for early diagnosis when the disease is still treatable as well as the emergence of resistance to chemotherapy agents. These two aspects are responsible for the description of OC as the most lethal gynecological malignancy, although it is comparably rarer in the population than cervical and breast cancers. The findings of the present study, therefore, have the overall effect of positively impacting the future management of EOC patients through improved accuracy in the early detection of EOC and monitoring of patient response to chemotherapy. The net effect of improved diagnostic accuracy is the reduction in cases of misdiagnosis, and hence a fall in the number of mortalities following increased successful cures or prolonged patient survival following effective treatment and monitoring strategies.

Infectious diseases such as HBV and HIV have also been implicated in shortening the survival of patients of gynecological malignancies, including OC patients where optimal treatment is not applied [2-4]. The findings of the present study indicated high prevalences of HBV and HIV in Kenyan OC patients, yet testing for these viruses is not a routine practice for OC patients. Therefore, to reduce early OC patient mortalities associated with HBV and HIV infections, routine testing of HBV and HIV in OC patients should be done as a standard practice prior to the administration of chemotherapy. Establishing the OC patients' HIV or HBV status, just as is the case in the United States and Hong Kong, will enable the selection of the best treatment strategies that will suppress the viral load as well as curtail tumor growth, and consequently enhance patients' longevity. Conversely, testing for HCV should be done only on the basis of necessity, given the comparatively low prevalence.

The present study provides a basis for more research into new glycome-based biomarkers that can be applied in the diagnosis and monitoring of the clinical response to anti-cancer agents by OC patients across ethnicities. However, to achieve a higher inferential power, the problem of small sample size needs to be overcome through multi-site collaborative studies.

7.0 Supplementary materials

Table S1. Differential features of the *N*-glycans expression in EOC versus BOD subjects. The 11 *N*-glycan signatures of EOC are highlighted in bold letters.

	N-glycan Formula	*m/z*	Assigned structure	BOD (n=46)		EOC (n=19)		p-value
				Median (R.I.)	Min-max (R.I.)	Median (R.I.)	Min-max (R.I.)	
1	N$_3$H$_3$	1416.7		0.002	0.001-0.005	0.001	0.0004-0.003	<0.001
2	N$_2$H$_5$	1579.8		0.031	0.008-0.059	0.013	0.003-0.020	<0.001
3	N$_3$H$_4$	1620.8		0.003	0.001-0.007	0.001	0.0003-0.002	<0.001
4	N$_4$H$_3$	1661.8		0.003	0.001-0.010	0.001	0.0003-0.004	<0.001
5	N$_2$H$_6$	1783.9		0.028	0.007-0.065	0.013	0.002-0.020	<0.001
6	N$_4$H$_3$F$_1$	1835.9		0.019	0.006-0.087	0.022	0.008-0.070	**0.604**
7	N$_3$H$_4$S$_1$	1982.0		0.008	0.004-0.015	0.004	0.001-0.008	<0.001
8	N$_2$H$_7$	1988.0		0.010	0.003-0.021	0.004	0.001-0.009	<0.001
9	N$_4$H$_4$F$_1$	2040.0		0.018	0.006-0.052	0.010	0.002-0.024	<0.001
10	N$_4$H$_5$	2070.0		0.009	0.003-0.023	0.003	0.001-0.007	<0.001
11	N$_5$H$_4$	2111.1		0.004	0.002-0.009	0.002	0.001-0.007	0.003
12	N$_2$H$_8$	2192.1		0.018	0.006-0.037	0.008	0.001-0.019	<0.001
13	N$_4$H$_4$S$_1$	2227.1		0.012	0.008-0.022	0.007	0.002-0.015	<0.001
14	N$_4$H$_5$F$_1$	2244.1		0.007	0.002-0.017	0.002	0.0003-0.006	<0.001
15	N$_5$H$_4$F$_1$	2285.2		0.002	0.001-0.003	0.001	0.0002-0.002	<0.001
16	N$_5$H$_5$	2315.2		0.002	0.001-0.004	0.001	0.0002-0.002	<0.001
17	N$_3$H$_6$S$_1$	2390.2		0.005	0.002-0.012	0.003	0.001-0.007	<0.001
18	N$_4$H$_5$S$_1$	2431.2		0.144	0.085-0.189	0.090	0.053-0.152	<0.001
19	N$_5$H$_5$F$_1$	2489.3		0.003	0.001-0.005	0.001	0.0002-0.003	<0.001
20	N$_5$H$_6$	2519.3		0.002	0.001-0.004	0.001	0.0002-0.002	<0.001
21	N$_4$H$_5$S$_1$F$_1$	2605.3		0.010	0.005-0.021	0.006	0.001-0.014	<0.001
22	N$_5$H$_5$F$_2$	2663.3		0.004	0.002-0.011	0.004	0.001-0.008	**0.462**

#	Composition	Mass	Structure	R.I. EOC median	R.I. EOC min-max	R.I. BOD median	R.I. BOD min-max	p
23	$N_5H_5S_1$	2676.3		0.007	0.004-0.016	0.003	0.001-0.010	<0.001
24	$N_5H_6F_1$	2693.4		0.002	0.001-0.002	0.001	0.0002-0.002	<0.001
25	N_6H_6	2764.4		0.005	0.002-0.010	0.003	0.001-0.006	<0.001
25	$N_4H_5S_2$	2792.4		0.447	0.302-0.550	0.467	0.292-0.722	**0.273**
27	$N_5H_5S_1F_1$	2850.4		0.005	0.003-0.013	0.002	0.0004-0.011	<0.001
28	$N_5H_6S_1$	2880.4		0.016	0.007-0.029	0.009	0.003-0.022	<0.001
29	$N_4H_5S_2F_1$	2966.5		0.011	0.006-0.023	0.015	0.004-0.023	**0.106**
30	$N_5H_5S_2$	3037.5		0.002	0.002-0.006	0.002	0.001-0.004	0.002
31	$N_5H_6S_1F_1$	3054.5		0.002	0.001-0.005	0.001	0.0004-0.003	0.012
32	$N_5H_5S_2F_1$	3211.6		0.001	0.001-0.004	0.001	0.002-0.002	**0.071**
33	$N_5H_6S_2$	3241.6		0.016	0.001-0.029	0.016	0.008-0.023	**0.604**
34	$N_5H_6S_2F_1$	3415.7		0.002	0.001-0.016	0.006	0.001-0.011	0.001
35	$N_5H_6S_3$	3602.8		0.052	0.025-0.111	0.067	0.032-0.091	0.016
36	$N_6H_7S_2$	3690.8		0.006	0.003-0.017	0.006	0.003-0.015	**0.708**
37	$N_5H_6S_3F_1$	3776.9		0.011	0.001-0.076	0.056	0.011-0.095	<0.001
38	$N_6H_7S_2F_1$	3864.9		0.001	0.0004-0.013	0.003	0.001-0.008	0.001
39	$N_5H_6S_3F_2$	3951.0		0.001	0.0003-0.006	0.002	0.0004-0.008	0.002
40	$N_6H_7S_3$	4052.0		0.008	0.004-0.019	0.013	0.005-0.024	<0.001
41	$N_6H_7S_3F_1$	4226.1		0.001	0.0003-0.015	0.009	0.002-0.017	<0.001
42	$N_6H_7S_3F_2$	4400.2		0.001	0.0004-0.006	0.003	0.0004-0.014	<0.001
43	$N_6H_7S_4$	4413.2		0.010	0.004-0.056	0.026	0.010-0.042	<0.001
44	$N_6H_7S_4F_1$	4587.3		0.002	0.0003-0.025	0.020	0.003-0.051	<0.001
45	$N_6H_7S_4F_2$	4761.4		0.001	0.0002-0.015	0.008	0.001-0.047	<0.001
46	$N_6H_7S_4F_3$	4935.5		0.0001	0.0002.005	0.002	0.0003-0.023	<0.001

The 11 *N*-glycan signatures are in bold letters. Medians were calculated from the relative intensities of each *N*-glycan of EOC and BOD as judged by MALDI-TOF-MS. Min-max, minimum and maximum value of the relative intensities of the *N*-glycan peaks of EOC and BOD. R.I., relative intensity of the *N*-glycan peaks, significant level was set at p <0.05.

References

1.   Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018, 68(6):394-424.

2.   Coghill AE, Newcomb PA, Madeleine MM, Richardson BA, Mutyaba I, Okuku F, Phipps W, Wabinga H, Orem J, Casper C: Contribution of HIV infection to mortality among cancer patients in Uganda. *AIDS* 2013, 27(18):2933-2942.

3.   Oliver NT, Chiao EY: Malignancies in women with HIV infection. *Curr Opin HIV AIDS* 2017, 12(1):69-76.

4.   Wong L, Cheung TH, Yim SF, Lao TT: Prevalence and impact of hepatitis B virus infection in ovarian cancer patients in an endemic area-A retrospective cohort study. *J Viral Hepat* 2020, 27(5):520-525.

5.   Wanyama FM, Tauber R, Mokomba A, Nyongesa C, Blanchard V: The Burden of Hepatitis B, Hepatitis C, and Human Immunodeficiency Viruses in Ovarian Cancer Patients in Nairobi, Kenya. *Infect Dis Rep* 2022, 14(3):433-445.

6.   Pokhriyal R, Hariprasad R, Kumar L, Hariprasad G: Chemotherapy Resistance in Advanced Ovarian Cancer Patients. *Biomark Cancer* 2019, 11:1179299x19860815.

7.   Bast RC, Jr., Feeney M, Lazarus H, Nadler LM, Colvin RB, Knapp RC: Reactivity of a monoclonal antibody with human ovarian carcinoma. *J Clin Invest* 1981, 68(5):1331-1337.

8.   Wanyama FM, Blanchard V: Glycomic-Based Biomarkers for Ovarian Cancer: Advances and Challenges. *Diagnostics (Basel)* 2021, 11(4).

9.   Zhang L, Chen Y, Wang K: Comparison of CA125, HE4, and ROMA index for ovarian cancer diagnosis. *Curr Probl Cancer* 2019, 43(2):135-144.

10.   Bandiera E, Romani C, Specchia C, Zanotti L, Galli C, Ruggeri G, Tognon G, Bignotti E, Tassi RA, Odicino F, Caimi L, Satori E, Santin A.D, Percorelli S, Ravaggi A: Serum human epididymis protein 4 and risk for ovarian malignancy algorithm as new diagnostic and prognostic tools for epithelial ovarian cancer management. *Cancer Epidemiol Biomarkers Prev* 2011, 20(12):2496-2506.

11.   Santotoribio JD, Garcia-de la Torre A, Cañavate-Solano C, Arce-Matute F, Sanchez-del Pino MJ, Perez-Ramos S: Cancer antigens 19.9 and 125 as tumor markers in patients with mucinous ovarian tumors. *Eur J Gynaecol Oncol* 2016, 37(1):26-29.

12.   Lertkhachonsuk AA, Buranawongtrakoon S, Lekskul N, Rermluk N, Wee-Stekly WW, Charakorn C: Serum CA19-9, CA-125 and CEA as tumor markers for mucinous ovarian tumors. *J Obstet Gynaecol Res* 2020, 46(11):2287-2291.

13.   Tuxen MK, Sölétormos G, Dombernowsky P: Serum tumour marker CA 125 in monitoring of ovarian cancer during first-line chemotherapy. *Br J Cancer* 2001, 84(10):1301-1307.

14.   An HJ, Kronewitter SR, de Leoz ML, Lebrilla CB: Glycomics and disease markers. *Curr Opin Chem Biol* 2009, 13(5-6):601-607.

15.   Stanley P, Moremen KW, Lewis NE, Taniguchi N, Aebi M: N-Glycans. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Mohnen D, Kinoshita T, Packer NH, Prestegard JH, Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright © 2022 The Consortium of Glycobiology Editors, La Jolla, California; published by Cold Spring Harbor Laboratory Press; doi:10.1101/glycobiology.4e.9. All rights reserved.; 2022: 103-116.

16.   Ohtsubo K, Marth JD: Glycosylation in cellular mechanisms of health and disease. *Cell* 2006, 126(5):855-867.

17.   Reily C, Stewart TJ, Renfrow MB, Novak J: Glycosylation in health and disease. *Nature Reviews Nephrology* 2019, 15(6):346-366.

18.   Spiro RG: Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 2002, 12(4):43r-56r.

19.   Moremen KW, Tiemeyer M, Nairn AV: Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* 2012, 13(7):448-462.

20.   Varki A, Sharon N: Historical Background and Overview. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.; 2009.

21.   Kinoshita T, Ohishi K, Takeda J: GPI-Anchor Synthesis in Mammalian Cells: Genes, Their Products, and a Deficiency1. *The Journal of Biochemistry* 1997, 122(2):251-257.

22.   Cummings RD: The repertoire of glycan determinants in the human glycome. *Mol Biosyst* 2009, 5(10):1087-1104.

23.   Werz DB, Ranzinger R, Herget S, Adibekian A, von der Lieth CW, Seeberger PH: Exploring the structural diversity of mammalian carbohydrates ("glycospace") by statistical databank analysis. *ACS Chem Biol* 2007, 2(10):685-691.

24.   Nairn AV, York WS, Harris K, Hall EM, Pierce JM, Moremen KW: Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem* 2008, 283(25):17298-17313.

25.   Seeberger PH: Monosaccharide Diversity. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 19-30.

26.   Chao Q, Ding Y, Chen Z-H, Xiang M-H, Wang N, Gao X-D: Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans. *Frontiers in Chemistry* 2020, 8(513).

27.   Chao Q, Ding Y, Chen Z-H, Xiang M-H, Wang N, Gao X-D: Recent Progress in Chemo-Enzymatic Methods for the Synthesis of N-Glycans. *Frontiers in Chemistry* 2020, 8.

28.   Ruhaak LR, Miyamoto S, Lebrilla CB: - Developments in the Identification of Glycan Biomarkers for the Detection of Cancer. - *Mol Cell Proteomics2013 Apr 01; 12*:- 846--855.

29.   Varki A: Essentials of glycobiology, 2. ed. edn. Cold Spring Harbor, NY: Cold Spring Harbor, NY : Cold Spring Harbor Laboratory Press; 2009.

30.   Yan A, Lennarz WJ: Unraveling the mechanism of protein N-glycosylation. *J Biol Chem* 2005, 280(5):3121-3124.

31.   Pinho SS, Reis CA: Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* 2015, 15(9):540-555.

32.   Stanley P, Taniguchi N, Aebi M: N-Glycans. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard, Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 99-111.

33.   Stanley P, Schachter H, Taniguchi N: N-Glycans. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.; 2009.

34.   Varki A: Biological roles of glycans. *Glycobiology* 2017, 27(1):3-49.

35.   Lauc G, Pezer M, Rudan I, Campbell H: Mechanisms of disease: The human N-glycome. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2016, 1860(8):1574-1582.

36.   Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH: Cell surface protein glycosylation in cancer. *Proteomics* 2014, 14(4-5):525-546.

37.   Welti M: Regulation of dolichol-linked glycosylation. *Glycoconj J* 2013, 30(1):51-56.

38.   Aebi M: N-linked protein glycosylation in the ER. *Biochim Biophys Acta* 2013, 1833(11):2430-2437.

39.   Rush JS: Role of Flippases in Protein Glycosylation in the Endoplasmic Reticulum. *Lipid Insights* 2015, 8(Suppl 1):45-53.

40.   Helenius A, Aebi M: Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004, 73:1019-1049.

41.   Hammond C, Braakman I, Helenius A: Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc Natl Acad Sci U S A* 1994, 91(3):913-917.

42.   Hammond C, Helenius A: Quality control in the secretory pathway: retention of a misfolded viral membrane glycoprotein involves cycling between the ER, intermediate compartment, and Golgi apparatus. *J Cell Biol* 1994, 126(1):41-52.

43.   Bieberich E: Synthesis, Processing, and Function of N-glycans in N-glycoproteins. *Adv Neurobiol* 2014, 9:47-70.

44.   Lederkremer GZ: Glycoprotein folding, quality control and ER-associated degradation. *Curr Opin Struct Biol* 2009, 19(5):515-523.

45.   Helenius A, Aebi M: Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004, 73:1019-1049.

46.   Stanley P: Golgi glycosylation. *Cold Spring Harb Perspect Biol* 2011, 3(4).

47.   Arnold JN, Saldova R, Hamid UM, Rudd PM: Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics* 2008, 8:3284.

48.   Chang IJ, He M, Lam CT: Congenital disorders of glycosylation. *Ann Transl Med* 2018, 6(24):477.

49.   Grunewald S, Matthijs G, Jaeken J: Congenital disorders of glycosylation: a review. *Pediatr Res* 2002, 52(5):618-624.

50.   Freeze HH, Schachter H, Kinoshita T: Genetic Disorders of Glycosylation. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 569-582.

51.   Ganetzky RD, Reynoso FJ, He M: Chapter 15 – Congenital disorders of glycosylation. In*: 2017*.

52.   Dercksen M, Crutchley AC, Honey EM, Lippert MM, Matthijs G, Mienie LJ, Schuman HC, Vorster BC, Jaeken J: ALG6-CDG in South Africa: Genotype-Phenotype Description of Five Novel Patients. *JIMD Rep* 2013, 8:17-23.

53.   Dall'Olio F, Vanhooren V, Chen CC, Slagboom PE, Wuhrer M, Franceschi C: N-glycomic biomarkers of biological aging and longevity: a link with inflammaging. *Ageing Res Rev* 2013, 12(2):685-698.

54.   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, 167:1731.

55.   Ding N, Nie H, Sun X, Sun W, Qu Y, Liu X, Yao Y, Liang X, Chen CC, Li Y: Human serum N-glycan profiles are age and sex dependent. *Age and Ageing* 2011, 40(5):568-575.

56.   Gornik O, Wagner J, Pučić M, Knežević A, Redžić I, Lauc G: Stability of N-glycan profiles in human plasma. *Glycobiology* 2009, 19(12):1547-1553.

57.   Wuhrer M, Stam JC, van de Geijn FE, Koeleman CA, Verrips CT, Dolhain RJ, Hokke CH, Deelder AM: Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* 2007, 7(22):4070-4081.

58.   Ercan A, Cui J, Chatterton DE, Deane KD, Hazen MM, Brintnell W, O'Donnell CI, Derber LA, Weinblatt ME, Shadick NA, Bell DA, Cairns E, Solomon DH, Holers VM, Rudd PM, Lee DM: Aberrant IgG galactosylation precedes disease onset, correlates with disease activity, and is prevalent in autoantibodies in rheumatoid arthritis. *Arthritis Rheum* 2010, 62(8):2239-2248.

59.   Tomana M, Schrohenloher R, Reveille J, Arnett F, Koopman W: Abnormal galactosylation of serum IgG in patients with systemic lupus erythematosus and members of families with high frequency of autoimmune diseases. *Rheumatol Int* 1992, 12(5):191-194.

60.   Ho JJ, Siddiki B, Kim YS: Association of sialyl-Lewis(a) and sialyl-Lewis(x) with MUC-1 apomucin ina pancreatic cancer cell line. *Cancer Res* 1995, 55(16):3659-3663.

61.   Nakamori S, Kameyama M, Imaoka S, Furukawa H, Ishikawa O, Sasaki Y, Kabuto T, Iwanaga T, Matsushita Y, Irimura T: Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. *Cancer Res* 1993, 53(15):3632-3637.

62.   Nakamori S, Arai I, Okamura S, Imaoka S, Furukawa H, Kabuto T, Ishikawa O, Sasaki Y, Kameyama M, Iwanaga T: [Clinical value of carbohydrate antigens, sialyl Lewis-x and sialyl Lewis-a in gastrointestinal cancer]. *Nihon Geka Gakkai Zasshi* 1996, 97(2):165-171.

63.   Yu CJ, Shih JY, Lee YC, Shun CT, Yuan A, Yang PC: Sialyl Lewis antigens: association with MUC5AC protein and correlation with post-operative recurrence of non-small cell lung cancer. *Lung Cancer* 2005, 47(1):59-67.

64.   Biskup K, Braicu EI, Sehouli J, Fotopoulou C, Tauber R, Berger M, Blanchard V: Serum glycome profiling: a biomarker for diagnosis of ovarian cancer. *J Proteome Res* 2013, 12(9):4056-4063.

65.   Fujita K, Hatano K, Hashimoto M, Tomiyama E, Miyoshi E, Nonomura N, Uemura H: Fucosylation in Urological Cancers. *Int J Mol Sci* 2021, 22(24).

66.   Fuster MM, Esko JD: The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer* 2005, 5(7):526-542.

67.   Arnold JN, Saldova R, Galligan MC, Murphy TB, Mimura-Kimura Y, Telford JE, Godwin AK, Rudd PM: Novel glycan biomarkers for the detection of lung cancer. *J Proteome Res* 2011, 10:1755.

68.   Ozcan S, Barkauskas DA, Renee Ruhaak L, Torres J, Cooke CL, An HJ, Hua S, Williams CC, Dimapasoc LM, Han Kim J, Camolinga-Ponce M, Rocke D, Lebrilla CB, Solnick JV: Serum glycan signatures of gastric cancer. *Cancer Prev Res (Phila)* 2014, 7(2):226-235.

69.   Hakomori S: Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci U S A* 2002, 99(16):10231-10233.

70.   Hakomori S, Kannagi R: Glycosphingolipids as tumor-associated and differentiation markers. *J Natl Cancer Inst* 1983, 71(2):231-251.

71.   Pinho SS, Reis CA: Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer* 2015, 15(9):540-555.

72.   Julien S, Adriaenssens E, Ottenberg K, Furlan A, Courtand G, Vercoutter-Edouart AS, Hanisch FG, Delannoy P, Le Bourhis X: ST6GalNAc I expression in MDA-MB-231 breast cancer cells greatly modifies their O-glycosylation pattern and enhances their tumourigenicity. *Glycobiology* 2006, 16(1):54-64.

73.   Marcos NT, Bennett EP, Gomes J, Magalhaes A, Gomes C, David L, Dar I, Jeanneau C, DeFrees S, Krustrup D: ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues. 2011.

74.   Kannagi R, Yin J, Miyazaki K, Izawa M: Current relevance of incomplete synthesis and neo-synthesis for cancer-associated alteration of carbohydrate determinants--Hakomori's concepts revisited. *Biochim Biophys Acta* 2008, 1780(3):525-531.

75.   Oliveira-Ferrer L, Legler K, Milde-Langosch K: Role of protein glycosylation in cancer metastasis. *Semin Cancer Biol* 2017, 44:141-152.

76.   Stowell SR, Ju T, Cummings RD: Protein glycosylation in cancer. *Annu Rev Pathol* 2015, 10:473-510.

77.   Büll C, Stoel MA, den Brok MH, Adema GJ: Sialic Acids Sweeten a Tumor's Life. *Cancer Res* 2014, 74(12):3199-3204.

78.   Kim K, Ruhaak LR, Nguyen UT, Taylor SL, Dimapasoc L, Williams C, Stroble C, Ozcan S, Miyamoto S, Lebrilla CB *et al*: Evaluation of glycomic profiling as a diagnostic biomarker for epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2014, 23(4):611-621.

79.   Alley WR, Jr., Vasseur JA, Goetz JA, Svoboda M, Mann BF, Matei DE, Menning N, Hussein A, Mechref Y, Novotny MV: N-linked glycan structures and their expressions change in the blood sera of ovarian cancer patients. *J Proteome Res* 2012, 11(4):2282-2300.

80.	Dědová T, Braicu EI, Sehouli J, Blanchard V: Sialic Acid Linkage Analysis Refines the Diagnosis of Ovarian Cancer. *Front Oncol* 2019, 9:261-261.

81.	Saldova R, Fan Y, Fitzpatrick JM, Watson RW, Rudd PM: Core-fucosylation and alpha2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology* 2011, 21(2):195-205.

82.	Wieczorek M, Braicu EI, Oliveira-Ferrer L, Sehouli J, Blanchard V: Immunoglobulin G Subclass-Specific Glycosylation Changes in Primary Epithelial Ovarian Cancer. *Front Immunol* 2020, 11(654).

83.	Ruhaak LR, Kim K, Stroble C, Taylor SL, Hong Q, Miyamoto S, Lebrilla CB, Leiserowitz G: Protein-Specific Differential Glycosylation of Immunoglobulins in Serum of Ovarian Cancer Patients. *J Proteome Res* 2016, 15(3):1002-1010.

84.	Theodoratou E, Thaçi K, Agakov F, Timofeeva MN, Štambuk J, Pučić-Baković M, Vučković F, Orchard P, Agakova A, Din FVN, Brown E, Rudd PM, Farrington SM, Dunlop MG, Campbell H, Lauc G: Glycosylation of plasma IgG in colorectal cancer prognosis. *Sci Rep* 2016, 6:28098-28098.

85.	Ruhaak LR, Barkauskas DA, Torres J, Cooke CL, Wu LD, Stroble C, Ozcan S, Williams CC, Camorlinga M, Rocke DM *et al*: The Serum Immunoglobulin G Glycosylation Signature of Gastric Cancer. *EuPA Open Proteom* 2015, 6:1-9.

86.	Weiz S, Wieczorek M, Schwedler C, Kaup M, Braicu EI, Sehouli J, Tauber R, Blanchard V: Acute-phase glycoprotein N-glycome of ovarian cancer patients analyzed by CE-LIF. *Electrophoresis* 2016, 37(11):1461-1467.

87.	Saldova R, Royle L, Radcliffe CM, Hamid UMA, Evans R, Arnold JN, Banks RE, Hutson R, Harvey DJ, Antrobus R, Petrescu SM, Dwek RA, Rudd PM: Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* 2007, 17:1344.

88.	Nakano M, Nakagawa T, Ito T, Kitada T, Hijioka T, Kasahara A, Tajiri M, Wada Y, Taniguchi N, Miyoshi E: Site-specific analysis of N-glycans on haptoglobin in sera of patients with pancreatic cancer: a novel approach for the development of tumor markers. *Int J Cancer* 2008, 122(10):2301-2309.

89.	Carlsson MC, Cederfur C, Schaar V, Balog CI, Lepur A, Touret F, Salomonsson E, Deelder AM, Fernö M, Olsson H, Wuhrer M, Leffler H: Galectin-1-binding glycoforms of haptoglobin with altered intracellular trafficking, and increase in metastatic breast cancer patients. *PLoS One* 2011, 6(10):e26560.

90.	Varki A: Sialic acids in human health and disease. *Trends Mol Med* 2008, 14(8):351-360.

91.	Wyss DF, Choi JS, Li J, Knoppers MH, Willis KJ, Arulanandam AR, Smolyar A, Reinherz EL, Wagner G: Conformation and function of the N-linked glycan in the adhesion domain of human CD2. *Science* 1995, 269(5228):1273-1278.

92.	Mengerink KJ, Vacquier VD: Glycobiology of sperm–egg interactions in deuterostomes. *Glycobiology* 2001, 11(4):37R-43R.

93.	Varki A, Schnaar RL, Schauer R: Sialic Acids and Other Nonulosonic Acids. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 179-195.

94.	Bergfeld AK, Varki A: Cytidine Monophospho-N-Acetylneuraminic Acid Hydroxylase (CMAH). In: *Handbook of Glycosyltransferases and Related Genes.* Edited by Taniguchi N, Honke K, Fukuda M, Narimatsu H, Yamaguchi Y, Angata T. Tokyo: Springer Japan; 2014: 1559-1580.

95.	Altman MO, Gagneux P: Absence of Neu5Gc and Presence of Anti-Neu5Gc Antibodies in Humans-An Evolutionary Perspective. *Front Immunol* 2019, 10:789.

96.	Severi E, Müller A, Potts JR, Leech A, Williamson D, Wilson KS, Thomas GH: Sialic acid mutarotation is catalyzed by the Escherichia coli beta-propeller protein YjhT. *J Biol Chem* 2008, 283(8):4841-4849.

97.   Paul A, Padler-Karavani V: Evolution of sialic acids: Implications in xenotransplant biology. *Xenotransplantation* 2018, 25(6):e12424.

98.   D'Addio M, Frey J, Otto VI: The manifold roles of sialic acid for the biological functions of endothelial glycoproteins. *Glycobiology* 2020, 30(8):490-499.

99.   Hugonnet M, Singh P, Haas Q, von Gunten S: The Distinct Roles of Sialyltransferases in Cancer Biology and Onco-Immunology. *Front Immunol* 2021, 12.

100.   Läubli H, Borsig L: Selectins promote tumor metastasis. *Semin Cancer Biol* 2010, 20(3):169-177.

101.   Boligan KF, Mesa C, Fernandez LE, von Gunten S: Cancer intelligence acquired (CIA): tumor glycosylation and sialylation codes dismantling antitumor defense. *Cell Mol Life Sci* 2015, 72(7):1231-1248.

102.   Bendas G, Borsig L: Cancer cell adhesion and metastasis: selectins, integrins, and the inhibitory potential of heparins. *Int J Cell Biol* 2012, 2012:676731.

103.   Kim YJ, Varki A: Perspectives on the significance of altered glycosylation of glycoproteins in cancer. *Glycoconj J* 1997, 14(5):569-576.

104.   Park JJ, Lee M: Increasing the α 2, 6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer. *Gut Liver* 2013, 7(6):629-641.

105.   Suzuki O, Abe M, Hashimoto Y: Sialylation by β-galactoside α-2,6-sialyltransferase and N-glycans regulate cell adhesion and invasion in human anaplastic large cell lymphoma. *Int J Oncol* 2015, 46(3):973-980.

106.   Saldova R, Piccard H, Pérez-Garay M, Harvey DJ, Struwe WB, Galligan MC, Berghmans N, Madden SF, Peracaula R, Opdenakker G, Rudd PM: Increase in sialylation and branching in the mouse serum N-glycome correlates with inflammation and ovarian tumour progression. *PLoS One* 2013, 8(8):e71159-e71159.

107.   Biskup K, Braicu EI, Sehouli J, Tauber R, Blanchard V: The serum glycome to discriminate between early-stage epithelial ovarian cancer and benign ovarian diseases. *Dis Markers* 2014, 2014:238197.

108.   Wang PH, Lee WL, Juang CM, Yang YH, Lo WH, Lai CR, Hsieh SL, Yuan CC: Altered mRNA expressions of sialyltransferases in ovarian cancers. *Gynecol Oncol* 2005, 99(3):631-639.

109.   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):219-232.

110.   Aubert M, Panicot L, Crotte C, Gibier P, Lombardo D, Sadoulet MO, Mas E: Restoration of alpha(1,2) fucosyltransferase activity decreases adhesive and metastatic properties of human pancreatic cancer cells. *Cancer Res* 2000, 60(5):1449-1456.

111.   Zhou X, Yang G, Guan F: Biological Functions and Analytical Strategies of Sialic Acids in Tumor. *Cells* 2020, 9(2).

112.   Zahradnikova M, Ihnatova I, Lattova E, Uhrik L, Stuchlikova E, Nenutil R, Valik D, Nalezinska M, Chovanec J, Zdrahal Z, Vojtesek B, Hernychova L, Novotny MV: N-Glycome changes reflecting resistance to platinum-based chemotherapy in ovarian cancer. *J Proteomics* 2021, 230:103964.

113.   Lange T, Ullrich S, Müller I, Nentwich MF, Stübke K, Feldhaus S, Knies C, Hellwinkel OJ, Vessella RL, Abramjuk C, Anders M, Schröder-Schwarz J, Schlomm J, Hauland H, Sauter G, Schumaker U: Human prostate cancer in a clinically relevant xenograft mouse model: identification of β(1,6)-branched oligosaccharides as a marker of tumor progression. *Clin Cancer Res* 2012, 18(5):1364-1373.

114.   Liu X, Nie H, Zhang Y, Yao Y, Maitikabili A, Qu Y, Shi S, Chen C, Li Y: Cell surface-specific N-glycan profiling in breast cancer. *PLoS One* 2013, 8(8):e72704.

115.   Yoshimura M, Nishikawa A, Ihara Y, Taniguchi S, Taniguchi N: Suppression of lung metastasis of B16 mouse melanoma by N-acetylglucosaminyltransferase III gene transfection. *Proc Natl Acad Sci U S A* 1995, 92(19):8754-8758.

116. Hua S, Williams CC, Dimapasoc LM, Ro GS, Ozcan S, Miyamoto S, Lebrilla CB, An HJ, Leiserowitz GS: Isomer-specific chromatographic profiling yields highly sensitive and specific potential N-glycan biomarkers for epithelial ovarian cancer. *J Chromatogr A* 2013, 1279:58-67.

117. Mok SC, Chao J, Skates S, Wong K, Yiu GK, Muto MG, Berkowitz RS, Cramer DW: Prostasin, a potential serum marker for ovarian cancer: identification through microarray technology. *J Natl Cancer Inst* 2001, 93(19):1458-1464.

118. Miyoshi E, Moriwaki K, Nakagawa T: Biological function of fucosylation in cancer biology. *J Biochem* 2008, 143(6):725-729.

119. . In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.; 2009.

120. Blanas A, Sahasrabudhe NM, Rodríguez E, van Kooyk Y, van Vliet SJ: Fucosylated Antigens in Cancer: An Alliance toward Tumor Progression, Metastasis, and Resistance to Chemotherapy. *Front Oncol* 2018, 8:39.

121. Narita T, Funahashi H, Satoh Y, Watanabe T, Sakamoto J, Takagi H: Association of expression of blood group-related carbohydrate antigens with prognosis in breast cancer. *Cancer* 1993, 71(10):3044-3053.

122. Ura Y, Dion AS, Williams CJ, Olsen BD, Redfield ES, Ishida M, Herlyn M, Major PP: Quantitative dot blot analyses of blood-group-related antigens in paired normal and malignant human breast tissues. *Int J Cancer* 1992, 50(1):57-63.

123. Shan M, Yang D, Dou H, Zhang L: Fucosylation in cancer biology and its clinical applications. *Prog Mol Biol Transl Sci* 2019, 162:93-119.

124. Chen CY, Jan YH, Juan YH, Yang CJ, Huang MS, Yu CJ, Yang PC, Hsiao M, Hsu TL, Wong CH: Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. *Proc Natl Acad Sci U S A* 2013, 110(2):630-635.

125. Yue L, Han C, Li Z, Li X, Liu D, Liu S, Yu H: Fucosyltransferase 8 expression in breast cancer patients: A high throughput tissue microarray analysis. *Histol Histopathol* 2016, 31(5):547-555.

126. Saldova R, Struwe WB, Wynne K, Elia G, Duffy MJ, Rudd PM: Exploring the glycosylation of serum CA125. *Int J Mol Sci* 2013, 14(8):15636-15654.

127. Liu YC, Yen HY, Chen CY, Chen CH, Cheng PF, Juan YH, Chen CH, Khoo KH, Yu CJ, Yang PC, Tsu T.L, Wong CH: Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc Natl Acad Sci U S A* 2011, 108(28):11332-11337.

128. Osumi D, Takahashi M, Miyoshi E, Yokoe S, Lee SH, Noda K, Nakamori S, Gu J, Ikeda Y, Kuroki Y, Senyoku K, Ishikawa M: Core-fucosylation of E-cadherin enhances cell-cell adhesion in human colon carcinoma WiDr cells. *Cancer Sci* 2009, 100(5):888-895.

129. Zhao YP, Xu XY, Fang M, Wang H, You Q, Yi CH, Ji J, Gu X, Zhou PT, Cheng C, Gao CF: Decreased core-fucosylation contributes to malignancy in gastric cancer. *PLoS One* 2014, 9(4):e94536.

130. Hua S, An HJ, Ozcan S, Ro GS, Soares S, DeVere-White R, Lebrilla CB: Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst* 2011, 136(18):3663-3671.

131. Nakagoe T, Fukushima K, Nanashima A, Sawai T, Tsuji T, Jibiki M, Yamaguchi H, Yasutake T, Ayabe H, Matuo T, Tagawa Y, Arisawa K: Expression of Lewis(a), sialyl Lewis(a), Lewis(x) and sialyl Lewis(x) antigens as prognostic factors in patients with colorectal cancer. *Can J Gastroenterol* 2000, 14(9):753-760.

132. Sato Y, Nakata K, Kato Y, Shima M, Ishii N, Koji T, Taketa K, Endo Y, Nagataki S: Early recognition of hepatocellular carcinoma based on altered profiles of alpha-fetoprotein. *N Engl J Med* 1993, 328(25):1802-1806.

133. Yan X, Lin Y, Liu S, Yan Q: Fucosyltransferase IV (FUT4) as an effective biomarker for the diagnosis of breast cancer. *Biomed Pharmacother* 2015, 70:299-304.

134.   Nakayama T, Watanabe M, Katsumata T, Teramoto T, Kitajima M: Expression of sialyl Lewis(a) as a new prognostic factor for patients with advanced colorectal carcinoma. *Cancer* 1995, 75(8):2051-2056.

135.   Chiu ML, Goulet DR, Teplyakov A, Gilliland GL: Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* 2019, 8(4).

136.   Bakovic MP, Selman MH, Hoffmann M, Rudan I, Campbell H, Deelder AM, Lauc G, Wuhrer M: High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides. *J Proteome Res* 2013, 12:821.

137.   Mimura Y, Church S, Ghirlando R, Ashton PR, Dong S, Goodall M, Lund J, Jefferis R: The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* 2000, 37(12-13):697-706.

138.   Gudelj I, Lauc G, Pezer M: Immunoglobulin G glycosylation in aging and diseases. *Cell Immunol* 2018, 333:65-79.

139.   Plomp R, Bondt A, de Haan N, Rombouts Y, Wuhrer M: Recent Advances in Clinical Glycoproteomics of Immunoglobulins (Igs). *Mol Cell Proteomics* 2016, 15(7):2217-2228.

140.   Selman MH, McDonnell LA, Palmblad M, Ruhaak LR, Deelder AM, Wuhrer M: Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 2010, 82:1073.

141.   Gebrehiwot AG, Melka DS, Kassaye YM, Gemechu T, Lako W, Hinou H, Nishimura SI: Exploring serum and immunoglobulin G N-glycome as diagnostic biomarkers for early detection of breast cancer in Ethiopian women. *BMC Cancer* 2019, 19(1):588.

142.   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, 29(1):57-66.

143.   Zauner G, Selman MH, Bondt A, Rombouts Y, Blank D, Deelder AM, Wuhrer M: Glycoproteomic analysis of antibodies. *Mol Cell Proteomics* 2013, 12(4):856-865.

144.   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.

145.   Kanoh Y, Mashiko T, Danbara M, Takayama Y, Ohtani S, Egawa S, Baba S, Akahoshi T: Changes in serum IgG oligosaccharide chains with prostate cancer progression. *Anticancer Res* 2004, 24:3135.

146.   Kawaguchi-Sakita N, Kaneshiro-Nakagawa K, Kawashima M, Sugimoto M, Tokiwa M, Suzuki E, Kajihara S, Fujita Y, Iwamoto S, Tanaka K: Serum immunoglobulin G Fc region N-glycosylation profiling by matrix-assisted laser desorption/ionization mass spectrometry can distinguish breast cancer patients from cancer-free controls. *Biochemical and biophysical research communications* 2016, 469(4):1140-1145.

147.   Pilkington C, Taylor PV, Silverman E, Isenberg DA, Costello AMdL, Rook GAW: Agalactosyl IgG and materno-fetal transmission of autoimmune neonatal lupus. *Rheumatol Int* 1996, 16(3):89-94.

148.   Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, Stanworth D, Rademacher TW, Mizuochi T, Taniguchi T, Matsuta K, Takeuchi F, Nagano Y, Miyamoto T, Kobata A: Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 1985, 316(6027):452-457.

149.   Krištić J, Vučković F, Menni C, Klarić L, Keser T, Beceheli I, Pučić-Baković M, Novokmet M, Mangino M, Thaqi K: Glycans are a novel biomarker of chronological and biological ages. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences* 2014, 69(7):779-789.

150.   van de Geijn FE, Wuhrer M, Selman MH, Willemsen SP, de Man YA, Deelder AM, Hazes JM, Dolhain RJ: Immunoglobulin G galactosylation and sialylation are associated with pregnancy-

induced improvement of rheumatoid arthritis and the postpartum flare: results from a large prospective cohort study. *Arthritis Res Ther* 2009, 11(6):R193.

151.   Rook GA, Steele J, Brealey R, Whyte A, Isenberg D, Sumar N, Nelson JL, Bodman KB, Young A, Roitt IM, Williams M, Rudd P, Redman CWA, Dwek R.A, Rademacher TW: Changes in IgG glycoform levels are associated with remission of arthritis during pregnancy. *J Autoimmun* 1991, 4(5):779-794.

152.   Chen G, Wang Y, Qiu L, Qin X, Liu H, Wang X, Wang Y, Song G, Li F, Guo Y *et al*: Human IgG Fc-glycosylation profiling reveals associations with age, sex, female sex hormones and thyroid cancer. *J Proteomics* 2012, 75(10):2824-2834.

153.   Dube R, Rook G, Steele J, Brealey R, Dwek R, Rademacher T, Lennard-Jones J: Agalactosyl IgG in inflammatory bowel disease: correlation with C-reactive protein. *Gut* 1990, 31(4):431-434.

154.   Aurer I, Lauc G, Dumić J, Rendić D, Matišić D, Miloš M, Heffer-Lauc M, Flogel M, Labar B: Aberrant glycosylation of Igg heavy chain in multiple myeloma. *Coll Antropol* 2007, 31(1):247-251.

155.   Kazuno S, Furukawa J-i, Shinohara Y, Murayama K, Fujime M, Ueno T, Fujimura T: Glycosylation status of serum immunoglobulin G in patients with prostate diseases. *Cancer Medicine* 2016, 5(6):1137-1146.

156.   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, 29(1):57-66.

157.   Chen G, Wang Y, Qin X, Li H, Guo Y, Wang Y, Liu H, Wang X, Song G, Li F, Li F, Guo S, Qiu L, Li Z: Change in IgG1 Fc N-linked glycosylation in human lung cancer: Age- and sex-related diagnostic potential. *Electrophoresis* 2013, 34(16):2407-2416.

158.   Karsten CM, Pandey MK, Figge J, Kilchenstein R, Taylor PR, Rosas M, McDonald JU, Orr SJ, Berger M, Petzold D, Blanchard V, Winkler A, Hess C, Reid DM, Majoul IV; Strait RI, Harris NL; Köhl, G, Wex E, Ludwig R, Zillikens D, Nimmerjahn F, Finkelman FD, Brown GD, Ehlers M, Jöng K: Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. *Nat Med* 2012, 18(9):1401-1406.

159.   Arnold JN, Saldova R, Hamid UM, Rudd PM: Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics* 2008, 8(16):3284-3293.

160.   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, 2(2):101-114.

161.   Kanoh Y, Mashiko T, Danbara M, Takayama Y, Ohtani S, Egawa S, Baba S, Akahoshi T: Changes in serum IgG oligosaccharide chains with prostate cancer progression. *Anticancer Res* 2004, 24(5b):3135-3139.

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. *Glycoconjugate J* 2011, 29:57.

163.   Bond A, Alavi A, Axford JS, Bourke BE, Bruckner FE, Kerr MA, Maxwell JD, Tweed KJ, Weldon MJ, Youinou P, Hay FC: A detailed lectin analysis of IgG glycosylation, demonstrating disease specific changes in terminal galactose and N-acetylglucosamine. *J Autoimmun* 1997, 10(1):77-85.

164.   Le NP, Bowden TA, Struwe WB, Crispin M: Immune recruitment or suppression by glycan engineering of endogenous and therapeutic antibodies. *Biochim Biophys Acta* 2016, 1860(8):1655-1668.

165.   Anthony RM, Nimmerjahn F: The role of differential IgG glycosylation in the interaction of antibodies with FcγRs in vivo. *Curr Opin Organ Transplant* 2011, 16(1):7-14.

166.   Pearce OMT, Läubli H: Sialic acids in cancer biology and immunity. *Glycobiology* 2016, 26(2):111-128.

167.   Kemna MJ, Plomp R, van Paassen P, Koeleman CAM, Jansen BC, Damoiseaux J, Cohen Tervaert JW, Wuhrer M: Galactosylation and Sialylation Levels of IgG Predict Relapse in Patients With PR3-ANCA Associated Vasculitis. *EBioMedicine* 2017, 17:108-118.

168.   Ogata S, Shimizu C, Franco A, Touma R, Kanegaye JT, Choudhury BP, Naidu NN, Kanda Y, Hoang LT, Hibberd ML, Tremoulet AH, Varki A, Burns JC: Treatment response in Kawasaki disease is associated with sialylation levels of endogenous but not therapeutic intravenous immunoglobulin G. *PLoS One* 2013, 8(12):e81448.

169.   Fokkink WJ, Selman MH, Dortland JR, Durmuş B, Kuitwaard K, Huizinga R, van Rijs W, Tio-Gillen AP, van Doorn PA, Deelder AM, Wuhrer M, Jacobs BC: IgG Fc N-glycosylation in Guillain-Barré syndrome treated with immunoglobulins. *J Proteome Res* 2014, 13(3):1722-1730.

170.   Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, Uchida K, Anazawa H, Satoh M, Yamasaki M: 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, 278(5):3466-3473.

171.   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, 74(4):288-294.

172.   Qiu Y, Patwa TH, Xu L, Shedden K, Misek DE, Tuck M, Jin G, Ruffin MT, Turgeon DK, Synal S, Bresalier R, Marcon N, Brenner DE, Lubman DM: Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. *J Proteome Res* 2008, 7(4):1693-1703.

173.   Carrascal MA, Silva M, Ramalho JS, Pen C, Martins M, Pascoal C, Amaral C, Serrano I, Oliveira MJ, Sackstein R: Inhibition of fucosylation in human invasive ductal carcinoma reduces E-selectin ligand expression, cell proliferation, and ERK 1/2 and p38 MAPK activation. *Mol Oncol* 2018, 12(5):579-593.

174.   Niwa R, Hatanaka S, Shoji-Hosaka E, Sakurada M, Kobayashi Y, Uehara A, Yokoi H, Nakamura K, Shitara K: Enhancement of the antibody-dependent cellular cytotoxicity of low-fucose IgG1 is independent of FcγRIIIa functional polymorphism. *Clinical Cancer Research* 2004, 10(18):6248-6255.

175.   Knezevic A, Polasek O, Gornik O, Rudan I, Campbell H, Hayward C, Wright A, Kolcic I, O'Donoghue N, Bones J: Variability, heritability and environmental determinants of human plasma N-glycome. *J Proteome Res* 2009, 8(2):694-701.

176.   Scanlan CN, Burton DR, Dwek RA: Making autoantibodies safe. *Proc Natl Acad Sci U S A* 2008, 105(11):4081-4082.

177.   Masuda K, Kubota T, Kaneko E, Iida S, Wakitani M, Kobayashi-Natsume Y, Kubota A, Shitara K, Nakamura K: Enhanced binding affinity for FcγRIIIa of fucose-negative antibody is sufficient to induce maximal antibody-dependent cellular cytotoxicity. *Mol Immunol* 2007, 44(12):3122-3131.

178.   Ferrara C, Stuart F, Sondermann P, Brünker P, Umaña P: The carbohydrate at FcγRIIIa Asn-162: an element required for high affinity binding to non-fucosylated IgG glycoforms. *J Biol Chem* 2006, 281(8):5032-5036.

179.   Torre LA, Trabert B, DeSantis CE, Miller KD, Samimi G, Runowicz CD, Gaudet MM, Jemal A, Siegel RL: Ovarian cancer statistics, 2018. *CA Cancer J Clin* 2018, 68(4):284-296.

180.   Briggs MT, Condina MR, Klingler-Hoffmann M, Arentz G, Everest-Dass AV, Kaur G, Oehler MK, Packer NH, Hoffmann P: Translating N-Glycan Analytical Applications into Clinical Strategies for Ovarian Cancer. *Proteomics Clin Appl* 2019, 13(3):e1800099.

181.   Meinhold-Heerlein I, Fotopoulou C, Harter P, Kurzeder C, Mustea A, Wimberger P, Hauptmann S, Sehouli J: Statement by the Kommission Ovar of the AGO: The New FIGO and WHO Classifications of Ovarian, Fallopian Tube and Primary Peritoneal Cancer. *Geburtshilfe Frauenheilkd* 2015, 75(10):1021-1027.

182. Reid BM, Permuth JB, Sellers TA: Epidemiology of ovarian cancer: a review. *Cancer biology & medicine* 2017, 14(1):9-32.

183. Cheserem EJ, Kihara A-B, Kosgei RJ, Gathara D, Gichuhi S: Ovarian cancer in Kenyatta National Hospital in Kenya: Characteristics and management. *Open Journal of Obstetrics and Gynecology* 2013, Vol.03No.01:7.

184. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018, 68(6):394-424.

185. Doubeni CA, Doubeni AR, Myers AE: Diagnosis and Management of Ovarian Cancer. *Am Fam Physician* 2016, 93(11):937-944.

186. Biskup K, Braicu EI, Sehouli J, Tauber R, Blanchard V: The ascites N-glycome of epithelial ovarian cancer patients. *J Proteomics* 2017, 157:33-39.

187. Burges A, Schmalfeldt B: Ovarian cancer: diagnosis and treatment. *Dtsch Arztebl Int* 2011, 108(38):635-641.

188. Hippisley-Cox J, Coupland C: Identifying women with suspected ovarian cancer in primary care: derivation and validation of algorithm. *BMJ* 2011, 344:d8009.

189. Bell R, Petticrew M, Sheldon T: The performance of screening tests for ovarian cancer: results of a systematic review. *Br J Obstet Gynaecol* 1998, 105(11):1136-1147.

190. Grewal K, Hamilton W, Sharp D: Ovarian cancer prediction: development of a scoring system for primary care. *BJOG* 2013, 120(8):1016-1019.

191. Bast RC, Jr., Hennessy B, Mills GB: The biology of ovarian cancer: new opportunities for translation. *Nat Rev Cancer* 2009, 9(6):415-428.

192. Prat J: Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet* 2014, 124(1):1-5.

193. Alsop K, Fereday S, Meldrum C, deFazio A, Emmanuel C, George J, Dobrovic A, Birrer MJ, Webb PM, Stewart C, Friedlander M, Fox S, Bowtell D, Mitchell G: BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women with ovarian cancer: a report from the Australian Ovarian Cancer Study Group. *J Clin Oncol* 2012, 30(21):2654-2663.

194. Holschneider CH, Berek JS: Ovarian cancer: epidemiology, biology, and prognostic factors. *Semin Surg Oncol* 2000, 19(1):3-10.

195. Fathalla MF: Incessant ovulation--a factor in ovarian neoplasia? *Lancet* 1971, 2(7716):163.

196. Ness RB, Cottreau C: Possible role of ovarian epithelial inflammation in ovarian cancer. *J Natl Cancer Inst* 1999, 91(17):1459-1467.

197. Gong TT, Wu QJ, Vogtmann E, Lin B, Wang YL: Age at menarche and risk of ovarian cancer: a meta-analysis of epidemiological studies. *Int J Cancer* 2013, 132(12):2894-2900.

198. Cramer DW, Welch WR: Determinants of ovarian cancer risk. II. Inferences regarding pathogenesis. *J Natl Cancer Inst* 1983, 71(4):717-721.

199. Riman T, Nilsson S, Persson IR: Review of epidemiological evidence for reproductive and hormonal factors in relation to the risk of epithelial ovarian malignancies. *Acta Obstet Gynecol Scand* 2004, 83(9):783-795.

200. Dreher D, Junod AF: Role of oxygen free radicals in cancer development. *Eur J Cancer* 1996, 32a(1):30-38.

201. Scholler N, Urban N: CA125 in ovarian cancer. *Biomark Med* 2007, 1(4):513-523.

202. Badgwell D, Bast RC, Jr.: Early detection of ovarian cancer. *Dis Markers* 2007, 23(5-6):397-410.

203. Goff BA, Mandel L, Muntz HG, Melancon CH: Ovarian carcinoma diagnosis. *Cancer* 2000, 89(10):2068-2075.

204. Liu JH, Zanotti KM: Management of the adnexal mass. *Obstet Gynecol* 2011, 117(6):1413-1428.

205. Vergote I, Tropé CG, Amant F, Kristensen GB, Ehlen T, Johnson N, Verheijen RH, van der Burg ME, Lacave AJ, Panici PB, Kennter GG, Casado A, Mendiola C, Coens C, Verleye L, Stuart GCE,

Percorelli S, Reed NS: Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N Engl J Med* 2010, 363(10):943-953.

206. Chandra A, Pius C, Nabeel M, Nair M, Vishwanatha JK, Ahmad S, Basha R: Ovarian cancer: Current status and strategies for improving therapeutic outcomes. *Cancer Med* 2019, 8(16):7018-7031.

207. Abu Hassan SO, Nielsen DL, Tuxen MK, Petersen PH, Sölétormos G: Performance of seven criteria to assess CA125 increments among ovarian cancer patients monitored during first-line chemotherapy and the post-therapy follow-up period. *Future Sci OA* 2017, 3(3):Fso216.

208. Taniguchi N, Kizuka Y: Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Adv Cancer Res* 2015, 126:11-51.

209. Kyselova Z, Mechref Y, Kang P, Goetz JA, Dobrolecki LE, Sledge GW, Schnaper L, Hickey RJ, Malkas LH, Novotny MV: Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem* 2008, 54:1166.

210. Zhao R, Lin G, Wang Y, Qin W, Gao T, Han J, Qin R, Pan Y, Sun J, Ren C, Ren S, Xu C: Use of the serum glycan state to predict ovarian cancer patients' clinical response to chemotherapy treatment. *J Proteomics* 2020, 223:103752.

211. Saldova R, Haakensen VD, Rødland E, Walsh I, Stöckmann H, Engebraaten O, Børresen-Dale AL, Rudd PM: Serum N-glycome alterations in breast cancer during multimodal treatment and follow-up. *Mol Oncol* 2017, 11(10):1361-1379.

212. Qian Y, Wang Y, Zhang X, Zhou L, Zhang Z, Xu J, Ruan Y, Ren S, Xu C, Gu J: Quantitative Analysis of Serum IgG Galactosylation Assists Differential Diagnosis of Ovarian Cancer. *J Proteome Res* 2013, 12(9):4046-4055.

213. Peracaula R, Sarrats A, Rudd PM: Liver proteins as sensor of human malignancies and inflammation. *Proteomics Clin Appl* 2010, 4(4):426-431.

214. Mulloy B, Dell A, Stanley P, J HP: Structural Analysis of Glycans. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH , Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 639-652.

215. Plummer TH, Jr., Elder JH, Alexander S, Phelan AW, Tarentino AL: Demonstration of peptide:N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. *J Biol Chem* 1984, 259(17):10700-10704.

216. Grunow D, Blanchard V: Enzymatic Release of Glycoprotein N-Glycans and Fluorescent Labeling. *Methods Mol Biol* 2019, 1934:43-49.

217. Harvey DJ: Proteomic analysis of glycosylation: structural determination of N- and O-linked glycans by mass spectrometry. *Expert Rev Proteomics* 2005, 2(1):87-101.

218. Biskup K: Serum-based N-glycan biomarker for diagnosis of epithelial ovarian cancer. The ascites N-glycome of epithelial ovarian cancer. 2017.

219. Wedepohl S, Kaup M, Riese SB, Berger M, Dernedde J, Tauber R, Blanchard V: N-glycan analysis of recombinant L-Selectin reveals sulfated GalNAc and GalNAc-GalNAc motifs. *J Proteome Res* 2010, 9(7):3403-3411.

220. Mechref Y, Novotny MV: Structural investigations of glycoconjugates at high sensitivity. *Chem Rev* 2002, 102(2):321-369.

221. Selman MH, Hemayatkar M, Deelder AM, Wuhrer M: Cotton HILIC SPE microtips for microscale purification and enrichment of glycans and glycopeptides. *Anal Chem* 2011, 83(7):2492-2499.

222. Rudd P, Karlsson NG, Khoo KH, Packer NH: Glycomics and Glycoproteomics. In: *Essentials of Glycobiology.* Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH Schnaar R, Seeberger P.H. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.; 2015: 653-666.

223.   Zhou S, Veillon L, Dong X, Huang Y, Mechref Y: Direct comparison of derivatization strategies for LC-MS/MS analysis of N-glycans. *Analyst* 2017, 142(23):4446-4455.

224.   Dalpathado DS, Desaire H: Glycopeptide analysis by mass spectrometry. *Analyst* 2008, 133(6):731-738.

225.   Fishman JB, Berg EA: Protein A and Protein G Purification of Antibodies. *Cold Spring Harb Protoc* 2019, 2019(1).

226.   Powell LD: Preparation of glycopeptides. *Curr Protoc Mol Biol* 2001, Chapter 17:Unit17.14A.

227.   Huhn C, Selman MH, Ruhaak LR, Deelder AM, Wuhrer M: IgG glycosylation analysis. *Proteomics* 2009, 9(4):882-913.

228.   Balbín M, Grubb A, de Lange GG, Grubb R: DNA sequences specific for Caucasian G3m(b) and (g) allotypes: allotyping at the genomic level. *Immunogenetics* 1994, 39(3):187-193.

229.   Dard P, Lefranc M-P, Osipova L, Sanchez-Mazas A: DNA sequence variability of IGHG3 alleles associated to the main G3m haplotypes in human populations. *European Journal of Human Genetics* 2001, 9(10):765-772.

230.   Kavallaris M, Marshall GM: Proteomics and disease: opportunities and challenges. *Med J Aust* 2005, 182(11):575-579.

231.   Han L, Costello CE: Mass spectrometry of glycans. *Biochemistry Biokhimiia* 2013, 78(7):710-720.

232.   Gustafsson JO, Oehler MK, Ruszkiewicz A, McColl SR, Hoffmann P: MALDI Imaging Mass Spectrometry (MALDI-IMS)-application of spatial proteomics for ovarian cancer classification and diagnosis. *Int J Mol Sci* 2011, 12(1):773-794.

233.   Kuzmanov U, Kosanam H, Diamandis EP: The sweet and sour of serological glycoprotein tumor biomarker quantification. *BMC Med* 2013, 11:31.

234.   Reiding KR, Blank D, Kuijper DM, Deelder AM, Wuhrer M: High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal Chem* 2014, 86(12):5784-5793.

235.   Powell AK, Harvey DJ: Stabilization of sialic acids in N-linked oligosaccharides and gangliosides for analysis by positive ion matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* 1996, 10(9):1027-1032.

236.   Reddy G, Dalmasso EA: SELDI ProteinChip(R) Array Technology: Protein-Based Predictive Medicine and Drug Discovery Applications. *J Biomed Biotechnol* 2003, 2003(4):237-241.

237.   Wulfkuhle JD, Liotta LA, Petricoin EF: Proteomic applications for the early detection of cancer. *Nat Rev Cancer* 2003, 3(4):267-275.

238.   Wilm M: Principles of electrospray ionization. *Mol Cell Proteomics* 2011, 10(7):M111.009407.

239.   Jansen BC, Reiding KR, Bondt A, Hipgrave Ederveen AL, Palmblad M, Falck D, Wuhrer M: MassyTools: A High-Throughput Targeted Data Processing Tool for Relative Quantitation and Quality Control Developed for Glycomic and Glycoproteomic MALDI-MS. *J Proteome Res* 2015, 14(12):5088-5098.

240.   Wanyama FM, Sekadde-Kigondu C, Midigo G, Mathaiya J, Muthuri J, Mogi D, Maghanga J, Kamau S, Wanyonyi I, Maina F: Analytical evaluation of thermo scientific Indiko clinical chemistry analyzer. *Global Health* 2019.

241.   Wanyama F, Sekadde-Kigondu C, Maturi P: Analytical Evaluation of Hemochroma POC Haemoglobin Reader. *J Med Diagn Meth* 2016, 5(219):2.

242.   Abbott KL, Nairn AV, Hall EM, Horton MB, McDonald JF, Moremen KW, Dinulescu DM, Pierce M: Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. *Proteomics* 2008, 8(16):3210-3220.

243.   An HJ, Miyamoto S, Lancaster KS, Kirmiz C, Li B, Lam KS, Leiserowitz GS, Lebrilla CB: Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *J Proteome Res* 2006, 5:1626.

244.   An HJ, Miyamoto S, Lancaster KS, Kirmiz C, Li B, Lam KS, Leiserowitz GS, Lebrilla CB: Profiling of glycans in serum for the discovery of potential biomarkers for ovarian cancer. *Journal of proteome research* 2006, 5(7):1626-1635.

245.   Kulasingam V, Pavlou MP, Diamandis EP: Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. *Nat Rev Cancer* 2010, 10(5):371-378.

246.   Gebrehiwot AG, Melka DS, Kassaye YM, Rehan IF, Rangappa S, Hinou H, Kamiyama T, Nishimura SI: Healthy human serum N-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers. *PLoS One* 2018, 13.

247.   Tuxen MK, Sölétormos G, Dombernowsky P: Serum tumor marker CA 125 for monitoring ovarian cancer during follow-up. *Scand J Clin Lab Invest* 2002, 62(3):177-188.

248.   Soletormos G, Duffy MJ, Othman Abu Hassan S, Verheijen RH, Tholander B, Bast RC, Jr., Gaarenstroom KN, Sturgeon CM, Bonfrer JM, Petersen PH, Troonen H, Torre GC, Kulpa JK, Tuxen MK, Molina R: Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *Int J Gynecol Cancer* 2016, 26(1):43-51.

249.   Rubin SC, Randall TC, Armstrong KA, Chi DS, Hoskins WJ: Ten-year follow-up of ovarian cancer patients after second-look laparotomy with negative findings. *Obstet Gynecol* 1999, 93(1):21-24.

250.   Armstrong DK: Relapsed ovarian cancer: challenges and management strategies for a chronic disease. *Oncologist* 2002, 7 Suppl 5:20-28.

251.   Higai K, Aoki Y, Azuma Y, Matsumoto K: Glycosylation of site-specific glycans of alpha1-acid glycoprotein and alterations in acute and chronic inflammation. *Biochim Biophys Acta* 2005, 1725(1):128-135.

252.   Gornik O, Royle L, Harvey DJ, Radcliffe CM, Saldova R, Dwek RA, Rudd P, Lauc G: Changes of serum glycans during sepsis and acute pancreatitis. *Glycobiology* 2007, 17(12):1321-1332.

253.   Hashimoto S, Asao T, Takahashi J, Yagihashi Y, Nishimura T, Saniabadi AR, Poland DC, van Dijk W, Kuwano H, Kochibe N, Yazawa S: alpha1-acid glycoprotein fucosylation as a marker of carcinoma progression and prognosis. *Cancer* 2004, 101(12):2825-2836.

254.   Lau KS, Dennis JW: N-Glycans in cancer progression. *Glycobiology* 2008, 18(10):750-760.

255.   Miyahara K, Nouso K, Morimoto Y, Kinugasa H, Kato H, Yamamoto N, Tsutsumi K, Kuwaki K, Onishi H, Ikeda F, Nakamura S, Shiraha H, Takaki A, Nakahara T, Miura Y, Asada H, Amano M, Nishimura SI, Yamamoto K: Prognostic value of altered N-glycosylation of circulating glycoproteins in patients with unresectable pancreatic cancer treated with gemcitabine. *Pancreas* 2015, 44(4):551-556.

256.   Burris HA, 3rd, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, Cripps MC, Portenoy RK, Storniolo AM, Tarassoff P, Nelson R, Dorr FA, Stephens CD, Von Hoff AD: Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: a randomized trial. *J Clin Oncol* 1997, 15(6):2403-2413.

257.   Schultz MJ, Swindall AF, Wright JW, Sztul ES, Landen CN, Bellis SL: ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J Ovarian Res* 2013, 6(1):25.

258.   Miyahara K, Nouso K, Miyake Y, Nakamura S, Obi S, Amano M, Hirose K, Nishimura S, Yamamoto K: Serum glycan as a prognostic marker in patients with advanced hepatocellular carcinoma treated with sorafenib. *Hepatology* 2014, 59(1):355-356.

259.   Kim JH, Park CW, Um D, Baek KH, Jo Y, An H, Kim Y, Kim TJ: Mass spectrometric screening of ovarian cancer with serum glycans. *Dis Markers* 2014, 2014:634289.

260.   Dedova T, Grunow D, Kappert K, Flach D, Tauber R, Blanchard V: The effect of blood sampling and preanalytical processing on human N-glycome. *PLoS One* 2018, 13(7):e0200507.

261.   Sarojini S, Tamir A, Lim H, Li S, Zhang S, Goy A, Pecora A, Suh KS: Early detection biomarkers for ovarian cancer. *J Oncol* 2012, 2012:709049.

262.   Moore RG, McMeekin DS, Brown AK, DiSilvestro P, Miller MC, Allard WJ, Gajewski W, Kurman R, Bast RC, Jr., Skates SJ: A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 2009, 112(1):40-46.

263. Knezević A, Polasek O, Gornik O, Rudan I, Campbell H, Hayward C, Wright A, Kolcic I, O'Donoghue N, Bones J, Rudd PM, Lauc G: Variability, heritability and environmental determinants of human plasma N-glycome. *J Proteome Res* 2009, 8(2):694-701.

264. Balog CI, Stavenhagen K, Fung WL, Koeleman CA, McDonnell LA, Verhoeven A, Mesker WE, Tollenaar RA, Deelder AM, Wuhrer M: N-glycosylation of colorectal cancer tissues: a liquid chromatography and mass spectrometry-based investigation. *Mol Cell Proteomics* 2012, 11(9):571-585.

265. De Leoz MLA, Young LJ, An HJ, Kronewitter SR, Kim J, Miyamoto S, Borowsky AD, Chew HK, Lebrilla CB: High-mannose glycans are elevated during breast cancer progression. *Mol Cell Proteomics* 2011, 10(1).

266. Zhao YY, Takahashi M, Gu JG, Miyoshi E, Matsumoto A, Kitazume S, Taniguchi N: Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer Sci* 2008, 99(7):1304-1310.

267. Haakensen VD, Steinfeld I, Saldova R, Shehni AA, Kifer I, Naume B, Rudd PM, Børresen-Dale AL, Yakhini Z: Serum N-glycan analysis in breast cancer patients--Relation to tumour biology and clinical outcome. *Mol Oncol* 2016, 10(1):59-72.

268. Giménez E, Balmaña M, Figueras J, Fort E, Bolós C, Sanz-Nebot V, Peracaula R, Rizzi A: Quantitative analysis of N-glycans from human alfa-acid-glycoprotein using stable isotope labeling and zwitterionic hydrophilic interaction capillary liquid chromatography electrospray mass spectrometry as tool for pancreatic disease diagnosis. *Anal Chim Acta* 2015, 866:59-68.

269. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW: Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nat Med* 2000, 6(3):306-312.

270. Dennis JW, Laferté S, Waghorne C, Breitman ML, Kerbel RS: Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science* 1987, 236(4801):582-585.

271. Miyagi T, Takahashi K, Hata K, Shiozaki K, Yamaguchi K: Sialidase significance for cancer progression. *Glycoconj J* 2012, 29(8-9):567-577.

272. Almaraz RT, Tian Y, Bhattarcharya R, Tan E, Chen SH, Dallas MR, Chen L, Zhang Z, Zhang H, Konstantopoulos K, Yarema KJ: Metabolic flux increases glycoprotein sialylation: implications for cell adhesion and cancer metastasis. *Mol Cell Proteomics* 2012, 11(7):M112.017558.

273. Ihara S, Miyoshi E, Ko JH, Murata K, Nakahara S, Honke K, Dickson RB, Lin CY, Taniguchi N: Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding beta 1-6 GlcNAc branching. *J Biol Chem* 2002, 277(19):16960-16967.

274. Tang DG, Honn KV: Adhesion molecules and tumor metastasis: an update. *Invasion Metastasis* 1994, 14(1-6):109-122.

275. Pucic M, Muzinic A, Novokmet M, Skledar M, Pivac N, Lauc G, Gornik O: Changes in plasma and IgG N-glycome during childhood and adolescence. *Glycobiology* 2012, 22(7):975-982.

276. Yamada E, Tsukamoto Y, Sasaki R, Yagyu K, Takahashi N: Structural changes of immunoglobulin G oligosaccharides with age in healthy human serum. *Glycoconjugate J* 1997, 14:401.

277. Noreh J, Sekadde-Kigondu C, Karanja JG, Thagana NG: Median age at menopause in a rural population of western Kenya. *East Afr Med J* 1997, 74(10):634-638.

278. Banda NK, Wood AK, Takahashi K, Levitt B, Rudd PM, Royle L, Abrahams JL, Stahl GL, Holers VM, Arend WP: Initiation of the alternative pathway of murine complement by immune complexes is dependent on N-glycans in IgG antibodies. *Arthritis Rheum* 2008, 58(10):3081-3089.

279. 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, 1(3):237-243.

280. van Zeben D, Rook GA, Hazes JM, Zwinderman AH, Zhang Y, Ghelani S, Rademacher TW, Breedveld FC: Early agalactosylation of IgG is associated with a more progressive disease

course in patients with rheumatoid arthritis: results of a follow-up study. *Br J Rheumatol* 1994, 33(1):36-43.

281.   Jefferis R, Lund J: Interaction sites on human IgG-Fc for FcgammaR: current models. *Immunol Lett* 2002, 82(1-2):57-65.

Statutory declaration

"I, Francis Mugeni Wanyama, by personally signing this document in lieu of an oath, hereby affirm that I prepared the submitted dissertation on the topic Determination of N-glycome signatures for Early Diagnosis of Epithelial Ovarian Cancer and Monitoring Response to Chemotherapy in an African Cohort (Analyse von N-Glykom-Strukturen für die Frühdiagnose des epithelialen Ovarialkarzinoms und für das Ansprechen auf Chemotherapien in einer afrikanischen Kohorte), independently and without the support of third parties, and that I used no other sources and aids than those stated.

All parts which are based on the publications or presentations of other authors, either in letter or in spirit, are specified as such in accordance with the citing guidelines. The sections on methodology (in particular regarding practical work, laboratory regulations, statistical processing) and results (in particular regarding figures, charts and tables) are exclusively my responsibility.

Furthermore, I declare that I have correctly marked all of the data, the analyses, and the conclusions generated from data obtained in collaboration with other persons, and that I have correctly marked my own contribution and the contributions of other persons (cf. declaration of contribution). I have correctly marked all texts or parts of texts that were generated in collaboration with other persons.

My contributions to any publications to this dissertation correspond to those stated in the below joint declaration made together with the supervisor. All publications created within the scope of the dissertation comply with the guidelines of the ICMJE (International Committee of Medical Journal Editors; www.icmje.org) on authorship. In addition, I declare that I shall comply with the regulations of Charité – Universitätsmedizin Berlin on ensuring good scientific practice.

I declare that I have not yet submitted this dissertation in identical or similar form to another Faculty.

The significance of this statutory declaration and the consequences of a false statutory declaration under criminal law (Sections 156, 161 of the German Criminal Code) are known to me."

_____   _____

  Date   Signature

Declaration of your own contribution to any publications

[If parts of your monograph have already been published, you must declare this in the preface following the cover sheet and complete this statement. Where applicable, the contributions to publications must be stated clearly and in detail, so that the Doctoral Commission and the scientific assessors can clearly establish which parts were contributed by you. A concrete referral to the publication is desirable, for instance: "Tables 1, 4, 47 and 60 were created on the basis of my statistical evaluation."]

I Francis Mugeni Wanyama contributed the following to the below listed publications:

Publication 1:

Francis Mugeni Wanyama and Véronique Blanchard, Glycomic-based biomarkers for ovarian cancer: Advances and challenges, Diagnostics [Basel], 2021, 11(4).

Contribution:

- literature search
- manuscript write-up
- review and editing

Publication 2:

Francis Mugeni Wanyama, Rudolf Tauber, Alfred Mokomba, Catherine Nyongesa and Véronique Blanchard. The Burden of Hepatitis B, Hepatitis C and Human Immunodeficiency Viruses in Ovarian Cancer Patients in Nairobi, Kenya, *Infect Dis Rep* 2022, 14(3):433-445

Contribution:

- Concept development and study design.
- Patient recruitment and data acquisition,
- Material transfer
- laboratory analyses
- All the work pertaining to data analysis as presented in the tables.
- literature search, and manuscript write-up
- Review and editing

_____

Signature, date and stamp of first supervising university professor / lecturer

_____

Signature of doctoral candidate

My CV will not be published in the electronic version of my thesis for data protection reasons.

Publications and conference contributions

Publications

1. Wanyama F.M and Blanchard V. (2021) Glycomic-based biomarkers for ovarian cancer: Advances and challenges. Diagnostics (Basel) 2021 Vol. 11 Issue 4 https://dx.doi.org/10.3390%2Fdiagnostics11040643.

2. F. M. Wanyama, R. Tauber, A. Mokomba, C. Nyongesa and V. Blanchard. (2021) The burden of Hepatitis B, Hepatitis C and Human immunodeficiency Viruses in ovarian cancer patients Nairobi, Kenya. Infect Dis Rep 2022 Vol. 14 Issue 3, Pages 433-445. doi: 10.3390/idr14030047.

Conference poster presentations

1. Francis Mugeni-Wanyama, Alfred Mokomba, Catherine Nyongesa, and Véronique Blanchard. IgG_1 and IgG_2 Fc glycosylation signatures for diagnosis of epithelial ovarian cancer in an African cohort. 2022, International Federation of Clinical Chemistry and Laboratory Medicine – EUROMED LAB, Munich, Germany, DOI 10.1515/cclm-2021-5003 Clin Chem Lab Med 2021; 59, Special Suppl, pp S94 – S998, Nov/Dec 2021.

2. Wanyama F.M, Mokomba A, Nyongesa C, and Blanchard V. Altered IgG_1 and IgG_2 glycosylation as a potential biomarker for epithelial ovarian cancer, 2021. 31st Joint Glycobiology Meeting, Germany.

3. Wanyama F.M, Mokomba A, Nyongesa C, and Blanchard V. Altered IgG_1 and IgG_2 glycosylation: a potential biomarker for epithelial ovarian cancer, 2021. New and emerging technologies conference, Germany.

4. Dédova T, Wanyama F.M, Krüger L, Braicu E, Sehouli J, and Blanchard V: Analysis of Sialic acid refines diagnosis of ovarian cancer, 2019. New and emerging technologies conference, Potsdam, Germany.

5. Wanyama F.M, Krüger L, and Blanchard V: Evaluation of the diagnostic utility of total serum N-glycome and IgG glycopeptide as biomarkers of Epithelial ovarian cancer in a Kenyan cohort, 2019. Global Health Conference, Berlin, Germany

Acknowledgments

Although a single name appears as the author of this very noble scientific venture, it would be an immeasurable oversight if the effort of a dedicated team of ladies and gentlemen that cumulatively in large or small parts contributed to this work is not put down to pen and paper.

My very deepest sense of gratitude goes to my advisor and mentor Prof. Dr. Véronique Blanchard who guided me throughout the research work. Her invaluable counsel, encouragement and useful criticism that led to the fulfillment of the mandate of this study cannot go unappreciated.

I am also indebted to my colleagues at Charité Universitätsmedizin Berlin for their warmth and friendship. More specifically, my fellow "Glycos" of the AG. Blanchard's research group, namely Axel Teigeler, Lynn Krugger, Marta Grzeski, Karina Biskup and Christian Schwedler for their warmth, technical support and useful laboratory interactions that kept my research work in the progression mode.

My very special thanks go to a team like no other that for two years, they ensured steady enrollment of study participants in the three enrollment centers of Nairobi. The team which comprised Nurses and Laboratory Scientists under the exceptional leadership of Mrs. Anne Alindah did a tremendous job to safeguard the study timelines. But more importantly, the excellent execution of their mandate provided a firm foundation on which the current scientific outputs stand on as a testimony that a research finding is as good as the sample from which it was generated. Just to mention a few attributes, the team constantly exhibited high standards of blood sample collection, correct sample storage and subsequent shipment of samples to Charité Universitätsmedizin Berlin.

But the real fulcrum and the levers that propelled this study to achieve its set goals were the volunteer patients who freely offered themselves as the study subjects for the benefit of the future transformation of medical care for ovarian cancer patients. These generous women are indeed the heroines of our time and they will continue to shine wherever they set their feet.

My sincere thanks to my Clinical Chemistry colleagues at the University of Nairobi, but more specifically to my mentor, Prof. Christine Sekkade-Kingondu for always believing in me and encouraging me to push on. It will be an unforgivable error of omission if I fail to acknowledge my employer, the University of Nairobi, for giving me the time off to pursue my dream in Germany and the Charité Universitaetsmedizin Berlin for actualizing my dream by offering me the position and for providing all the necessary study materials. But perhaps all this would not have been possible were it not for the DAAD who believed in me and sponsored me to work with the best in the industry and

exploit my scientific talents to the fullest. Hence my very massive gratitude is hereby penned in golden letters to the people and the Government of Germany for their outpouring generosity.

Last but not least, to my dear wife, Anne, for accepting with a lot of humility, the role of single physical parenthood to our two lovely kids, Jansen and Melanie. Indeed that sacrifice to act as a two-in-one is unquantifiable and I have no fitting word left in my vocabulary other than to say "Asante sana swahiba". As for Jansen and Melanie, during the course of this study it never made sense to you why I was only a few times physically present in your lives, until you had to rename me "baba kuja, baba toroka" (Kiswahili words for the Father who comes around then runs away), I promise I will be much more around henceforth. Importantly, my most profound gratitude to my late dad, Mwalimu P.B.O Wanyama who was a real inspiration to me, pushed me, and believed that this day will happen, well my mother Selina is still around to witness this special happening of your joint effort. Finally, to my siblings, relatives, and friends, thank you for the support you gave to my family, this made it a lot easier for us to navigate through the period.

Allen, die hier erwähnt wurden, und denen, die ich nicht erwähnt habe, danke ich für ihre großzügige Unterstützung.

In the words of the "Father of the National Parks,"

"The power of imagination makes us infinite"

-John Muir

CHARITÉ

CharitéCentrum für Human- und Gesundheitswissenschaften

Charité | Campus Charité Mitte | 10117 Berlin

Surname, Name: Mugeni Wanyama, Francis

Email: francis.wanyama@charite.de

Matriculation Number: 9182105

Thesis Advisor: Prof. Dr. Véronique Blanchard

Institution / Clinic: Institute of Laboratory Medicine

Institut für Biometrie und Klinische Epidemiologie (iBikE)

Direktor: Prof. Dr. Frank Konietschke

Postanschrift:
Charitéplatz 1 | 10117 Berlin
Besucheranschrift:
Reinhardtstr. 58 | 10117 Berlin

Tel. +49 (0)30 450 562171
frank.konietschke@charite.de
https://biometrie.charite.de/

Certification

I hereby certify that Mr. Francis Mugeni-Wanyama has received a statistical consultation on a PhD project with me, within the Service Unit Biometry of the Institute of Biometry and Clinical Epidemiology (iBikE). The following 3 consultation dates were attended:

- *Appointment 1: 20.12.2021; Appointment 2: 31.1.2022, Appointment 3: 15.2.2022*

The following key advice regarding meaningful analysis and interpretation of data was given during the consultation:

- *The focus should be on the clear descriptive analysis and presentation of the data, where measures of both central tendency (e.g. mean, median) and of dispersion (e.g. standard deviation, interquartile range) should be used. Boxplots are recommended for the presentation of continuous data, bar charts for the presentation of absolute or relative proportions.*

- *T-tests can be used for comparing two groups if the data are approximately normally distributed. If this assumption does not hold, non-parametric methods should be used, e.g., the Mann-Whitney test to compare two independent groups with respect to an ordinal or continuous characteristic.*

- *The area under the receiver operating characteristic curve (AUC) with 95% confidence intervals is a good way to describe and compare the diagnostic ability of different methods/markers. If a cut-off for a certain marker is derived, it needs to be clearly stated how it was derived, as different criteria can be applied.*

- *In general the focus should be on effect sizes, as e.g. measured with an AUC or with a difference of means, rather than significance testing. When performing multiple tests, we need to consider a multiple testing issue, i.e., that when multiple tests are performed as part of an evaluation, the probability of finding differences between groups exceeds the 5% significance level. This must be taken into account when interpreting the results.*

This certificate does not guarantee the correct implementation of the suggestions made in the consultation, the correct performance of the recommended statistical procedures, nor the correct presentation and interpretation of the results. The responsibility for this rests solely with the doctoral student. The Institute of Biometry and Clinical Epidemiology assumes no liability for this.

Date: 9.2.2023   Name of consultant: Annette Aigner, PhD

Signature consultant, institute stamp

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN
Institut für Biometrie und
Klinische Epidemiologie
Campus Charité Mitte
Charitéplatz 1 | D-10117 Berlin