# Chemotherapy-induced toxicity and its association with clinical factors and gastrointestinal microbiome among breast cancer patients in Vietnam

By

Sang Minh Nguyen

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Epidemiology

May 31, 2022

Nashville, Tennessee

Approved:

Xiao-Ou Shu, M.D., Ph.D.

Hui Cai, M.D., Ph.D.

Jirong Long, Ph.D., M.S.

Martha Shrubsole, Ph.D.

Copyright © 2022 Sang Minh Nguyen
All Rights Reserved

# ACKNOWLEDGEMENTS

Throughout the writing of this dissertation, I have received a tremendous deal of guidance, support, and assistance.

First, I would like to express my most sincere gratitude to Dr. Xiao-Ou Shu, my primary mentor and thesis supervisor, for her guidance and her highest level of mentorship, tutelage, advice, and support during the past 5.5 years of my graduate training program and Vanderbilt-Emory-Cornell-Duke (VECD) Global Health fellowship. I cannot count how much effort she had spent and tirelessly guided me through the process, from mastering relevant scientific knowledge to accumulating research experience in several projects, which have improved my skills in statistical analysis and scientific writing. Her professionalism, enthusiasm, and passion for health science have inspired and influenced me. I am proud of and incredibly grateful to have had the opportunity to work with her for the past years. My time with her at Vanderbilt University has been a treasure.

I would also like to express my indeed gratitude to other thesis committee members, Drs. Jirong Long, Martha Shrubsole, and Hui Cai. Their knowledge, insights, and feedback have been invaluable and greatly enhanced my thesis. Drs. Jirong Long and Martha Shrubsole have provided valuable suggestions for developing my thesis project and building my expertise in microbiome analysis and genetic epidemiology. Dr. Hui Cai has always been willing to discuss with me statistical analysis aspects and provided a lot of helpful advice for my thesis project and PhD studies.

I am thankful to the investigators and leaders from Vietnam National Cancer Institute, National Cancer Hospitals (K Hospitals), Hanoi Oncology hospital, and the Ministry of Health for conducting the Vietnamese Breast Cancer Study (VBCS), supported by a P20 grant (P20 CA210300) from the U.S.A. National Cancer Institute and jointly led by Drs. Xiao-Ou Shu and Thuan V. Tran. I am incredibly grateful to Drs. Huong T.T. Tran and Thuan V. Tran, my two wonderful site mentors during the VECD Global Health fellowship, for their support and help during my data collection in Vietnam. I was fortunate to have a chance to work with their hard-working staff at Vietnam National Cancer Institute, without whom it would be impossible for my thesis project to be completed.

My grateful acknowledgment is to many other faculty members and colleagues at Vanderbilt University and Vanderbilt Epidemiology Center for their support and insight. I have learned a lot from their discussion during the scientific meetings involving Drs. Yaohua Yang, Danxia Yu, Qiuyin Cai, and Wei Zheng. I also express my thanks to Dr. Mary Shannon Byers for her assistance in editing my several projects. In addition, the Vanderbilt Epidemiology PhD program has provided me with lots of assistance throughout my PhD program. I am grateful to the program leaders, Drs. Peter F. Rebeiro, Todd L. Edwards, Katherine E. Hartmann, and the program manager Melissa Krasnove for their help. I also would like to thank my academic advisor, Dr. Stephen Deppen, who has constantly guided my scientific and career development. Finally, I also appreciate other colleagues and friends from Vietnam and the U.S.A. for their support.

Finally, I wish to thank my family, especially my wife Hang, my daughters Mina and Lucy, my parents, my mother-in-law, and Dr. Mary Kamb's family for their forever love, endless support and encouragement while concentrating on my PhD studies.

TABLE OF CONTENT


LIST OF TABLES ..... vii
LIST OF FIGURES ..... ix
**CHAPTER 1** ..... 1
INTRODUCTION ..... 1
**1. Breast cancer epidemiology** ..... 1
**2. Chemotherapy in breast cancer treatment** ..... 3
3. **Chemotherapy-induced toxicities** ..... 9
**4. The gastrointestinal (GI) microbiota in breast cancer patients** ..... 13
**5. Chemotherapy changes the GI microbiota** ..... 17
6. **The GI microbiota and chemotherapy-induced toxicity** ..... 25
**7. Summary** ..... 35
**CHAPTER 2** ..... 37
**1. Method** ..... 37
1.1. Study Design ..... 37
1.2. Population and participant recruitment ..... 37
1.3. Overall Assessments ..... 38
1.4. Population Selection ..... 40
1.5. Outcome Assessment ..... 40
1.6. Covariate Assessment: ..... 44
1.7. Statistical Analyses ..... 47
1.8. Statistical Power Estimation ..... 48
**2. Results** ..... 49
**3. Discussion** ..... 69
**CHAPTER 3** ..... 76
**1. Methods** ..... 76
1.1. Parent Study ..... 76
1.2. Population Selection ..... 76
1.3. Outcome Assessment ..... 76
1.4. Covariate Assessment ..... 78
1.5. Statistical Analysis ..... 80
1.6. Statistical Power Estimation ..... 82
**2. Results** ..... 84
**3. Discussion** ..... 125
**CHAPTER 4** ..... 131
**1. Methods** ..... 131

1.1. Parent Study ........ 131
1.2. Population Selection ........ 131
1.3. Outcome Assessment ........ 131
1.4. Stool Collection Time ........ 132
1.5. Microbiome Profiling ........ 133
1.6. Covariate Assessment ........ 135
1.7. Statistical Analysis ........ 136
1.8. Statistical Power Calculation ........ 138
**2. Results** ........ 139
**3. Discussion** ........ 177
**CHAPTER 5** ........ 184
**1. Methods** ........ 184
1.1. Parent Study ........ 184
1.2. Population Selection ........ 184
1.3. Outcome Assessment ........ 184
1.4. Exposure Assessment ........ 185
1.5. Statistical Analysis ........ 186
**2. Results** ........ 188
**3. Discussion** ........ 202
**OVERALL CONCLUSION** ........ 206
**VIII. APPENDIX** ........ 210
**Supplementary Table S1:** Human studies investigating changes to the diversity of GI microbiota among breast cancer patients in comparison with healthy controls and by clinical features ........ 210
**Supplementary Table S2:** Chemotherapeutic treatment and GI microbiota changes in clinical studies ........ 214
**Supplementary Table S3:** Chemotherapeutic treatment and GI microbiota changes in pre-clinical studies ........ 218
**Supplementary Table S4:** Patient self-report chemotherapy-induced side effects ........ 222
**Supplementary Table S5:** National Cancer Institute Common Toxicity Criteria ........ 223
**REFERENCE** ........ 224

# LIST OF TABLES

Table Page

**Table 1**: Neoadjuvant/adjuvant chemotherapy regimens for HER2-negative, stage I-III breast cancer in Vietnam ....7
**Table 2**: Chemotherapy-induced toxicities and GI microbiota ....29
**Table 3:** Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 1) ....44
**Table 4:** First-line treatment for breast cancer among study participants ....46
**Table 5:** Highest grade of chemotherapy-induced toxicities among study participants (total N=396) ....49
**Table 6**: Incidence of grade ≥ 3 chemotherapy-induced toxicities by chemotherapy regimens ....51
**Table 7**: Incidence of grade ≥ 3 chemotherapy-induced toxicities by sequential anthracycline and taxane ....52
**Table 8:** Incidence of ≥ 3 grade anthracycline-induced and taxane-induced toxicities among breast cancer patients who received sequential anthracycline and taxane ....53
**Table 9:** Highest grade of chemotherapy-induced toxicity by selected demographic characteristics among study participants (total N=396) ....55
**Table 10**: Highest grade of chemotherapy-induced toxicity by selected disease characteristics among study participants (total N=396) ....56
**Table 11**: Highest grade of chemotherapy-induced toxicity by characteristics related to treatment ....58
**Table 12:** Association of demographic characteristics and clinical factors with combined hematological toxicity ....61
**Table 13**: Association of demographic characteristics and clinical factors with neutropenia ....63
**Table 14**: Association of demographic characteristics and clinical factors with combined GI toxicities ....65
**Table 15**: Association of demographic characteristics and clinical factors with nausea/vomiting ....67
**Table 16**: Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 2) ....79
**Table 17**: Demographic characteristics and clinical factors of participants by stool collection time ....84
**Table 18**: Comparison of alpha diversity indexes by stool collection time ....86
**Table 19**: Comparison of Chao1 index by demographic characteristics and clinical factors ....88
**Table 20:** Comparison of Shannon index by demographic characteristics and clinical factors ....90
**Table 21**: Comparison of inverse Simpson index by demographic characteristics and clinical factors ....92
**Table 22**: Association of selected demographic characteristics and clinical factors with alpha diversity indexes ....94
**Table 23**: Association of selected demographic characteristics and clinical factors with Chao1 index by stool collection time ....96
**Table 24**: Association of selected demographic characteristics and clinical factors with Shannon index by stool collection time ....97
**Table 25**: Association of selected demographic characteristics and clinical factors with inverse Simpson index by stool collection time ....98
**Table 26**: PERMANOVA test difference of the beta diversity between demographic characteristic and clinical factors ....100
**Table 27**: Association of stool collection time with GI microbial taxa at phylum to order levels ....103
**Table 28**: Association of TNM cancer stage with GI microbial taxa by stool collection time ....106
**Table 29**: Association of breast cancer subtypes with GI microbial taxa by stool collection time ....107
**Table 30:** Association of ER status with GI microbial taxa by stool collection time ....108
**Table 31**: Association of HER2 status with GI microbial taxa by stool collection time ....109
**Table 32**: Association of age groups with GI microbial taxa by stool collection time ....112
**Table 33**: Association of income levels with GI microbial taxa by stool collection time at phylum to order levels ....113
**Table 34**: Association of residence (rural areas vs. urban areas) with GI microbial taxa by stool collection time ....115
**Table 35**: Association of BMI levels with GI microbial taxa by stool collection time ....120

**Table 36:** Association of diagnosis delay with GI microbial taxa by stool collection time........................122
**Table 37**: Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 3) ..135
**Table 38:** Incidence of ≥ 3 grade chemotherapy-induced toxicities by stool collection time ...................139
**Table 39**: Association of grade ≥3 combined hematological toxicity with alpha diversity indexes by stool collection time ........................................................................................................................141
**Table 40**: Association of grade ≥3 neutropenia with alpha diversity indexes by stool collection time.....142
**Table 41**: Association of grade ≥3 combined GI toxicity with alpha diversity indexes by stool collection time..................................................................................................................................143
**Table 42:** Association of grade ≥3 nausea/vomiting with alpha diversity indexes by stool collection time ........................................................................................................................................144
**Table 43**: PERMANOVA test difference of pre-chemotherapy beta diversity between chemotherapy-induced toxicities (grade <3 vs. grade ≥3)..................................................................................146
**Table 44**: Association of grade ≥3 combined hematological toxicity with pre-chemotherapy GI microbial taxa by stool collection time .............................................................................................150
**Table 45:** Association of grade ≥3 neutropenia with pre-chemotherapy GI microbial taxa by stool collection time ........................................................................................................................151
**Table 46**: Association of grade ≥3 nausea/vomiting with gut microbiome taxa by stool collection time at order to genus levels...................................................................................................................155
**Table 47:** Association of grade ≥3 nausea/vomiting with common microbiome taxa by stool collection time at species levels (Top 20) ..........................................................................................158
**Table 48**: Association of grade ≥3 nausea/vomiting with rare microbiome taxa by stool collection time at species levels (Top 20) ................................................................................................160
**Table 49:** Association of grade ≥3 combined hematological toxicity with metabolic pathways by stool collection time ........................................................................................................................164
**Table 50**: Association of grade ≥3 neutropenia with metabolic pathways by stool collection time..........168
**Table 51:** Association of grade ≥3 combined GI toxicity with metabolic pathways by stool collection time ........................................................................................................................................172
**Table 52:** Association of grade ≥3 combined nausea/vomiting with metabolic pathways by stool collection time ........................................................................................................................174
**Table 53**: Highest grade of hematological toxicities by chemotherapeutic agents .................................188
**Table 54:** Highest grade of GI toxicities by chemotherapeutic agents ...................................................189
**Table 55**: Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in doxorubicin-containing regimens stratify analysis[1]........................................................................190
**Table 56:** Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in docetaxel-containing regimens stratify analysis[1].............................................................................................190
**Table 57:** Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1].........................................................................193
**Table 58**: Association of grade ≥3 neutropenia with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1] ............................................................................................................194
**Table 59**: Association of grade ≥3 combined GI toxicity with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1] .........................................................................................195
**Table 60:** Association of grade ≥3 combined hematological toxicity with GI microbial metabolic pathways in doxorubicin-containing regimens stratify analysis .........................................................................198
**Table 61:** Association of grade ≥3 combined hematological toxicity with GI microbial metabolic pathways in docetaxel-containing regimens stratify analysis ............................................................................198
**Table 62:** Association of grade ≥3 combined hematological toxicity with GI metabolic pathways in the sequential anthracycline and taxane- stratify analysis ..............................................................................199
**Table 63:** Association of grade ≥3 neutropenia with GI microbial metabolic pathways in the sequential anthracycline and taxane- stratify analysis.......................................................................................200
**Table 64:** Association of grade ≥3 combined GI toxicity with GI microbial metabolic pathways in the sequential anthracycline and taxane- stratify analysis ..............................................................................201

LIST OF FIGURES


Figure Page

**Figure 1**: Plausible links between the GI microbiota and chemotherapy ....25
**Figure 2:** Flow diagram of study participant inclusion criteria for Aim 1 ....39
**Figure 3:** Timeline of chemotherapy-induced toxicity assessment ....43
**Figure 4:** Estimated odds ratio (OR) by the prevalence of grade ≥ 3 chemotherapy-induced toxicity......48
**Figure 5:** Flow diagram of study participant inclusion criteria for Aim 2 ....77
**Figure 6:** Power curve for a test of mean Shannon index according to sample size ratios between two groups ....82
**Figure 7:** Power curve for a test of mean relative abundance according to sample size ratios between two groups....82
**Figure 8:** Alpha diversity at the species level (measured in term of the Shannon index) by stool collection time....87
**Figure 9:** Principal coordinate analysis (PCoA) of Bray-Curtis diversity by stool collection time ....99
**Figure 10:** Relative abundance of the gut microbiome at phylum levels ....101
**Figure 11:** Boxplot of relative abundance at phylum level by stool collection time....102
**Figure 12:** Flow diagram of study subject inclusion criteria for Aim 3....132
**Figure 13:** Timeline of stool collection ....134
**Figure 14:** Estimated OR per SD increase in clr-transformed absolute abundance of microbial taxa by the prevalence of grade ≥ 3 chemotherapy-induced toxicity....138

# CHAPTER 1

## INTRODUCTION

### 1. Breast cancer epidemiology

Breast cancer is among the most common cancers and the leading cause of cancer deaths among women, with approximately 2.3 million incident cases and 685,00 deaths worldwide in 2020, corresponding to 11.7% of all cancer diagnoses and 6.9% of all cancer deaths[1]. Breast cancer incidence and mortality rates vary markedly among different populations and countries. In the U.S.A, the estimated age-adjusted annual incidence rate is 126.0/100,000 women, and the estimated age-adjusted annual mortality rate was 20.1/100,000 women in 2020. High-income countries (HICs) in North America, Australia-New Zealand, and Northern and Western Europe generally have > 80 incident cases and > 12 deaths per 100,000 women.[2,3] Women in Asian countries historically have lower incidence rates but higher mortality rates than women in the U.S.A. and other HICs.[1] Breast cancer incidence rate has seen a rapid increase among Asian women over the last decades. The age-standardized incidence rate has doubled or tripled in Japan, Korea, and Singapore over the past 40 years.[4] This increment has been attributed to a transition toward the risk factors profile of Western countries brought by the globalization of economies and behaviors.[5]

Many modifiable and nonmodifiable risk factors for breast cancer have been identified and well-established in recent decades.[6,7] Besides genetic factors (e.g., inherited loss of BRCA1/2 gene[8]), family history of breast cancer[6], and aging (after the age of 50), the primary identified risk factors are hormonal and lifestyle components, including diet, alcohol intake, obesity, high-dose radiation to the chest, high and long duration of female hormone exposure such as early menarche, late menopause, use of hormonal contraceptives and hormone replacement therapy.[9] However, the precise etiology for each breast cancer case is still unknown because a vast majority of the newly diagnosed women (~70%) have no apparent risk factor.[10] High physical activity levels, successful pregnancies, and breastfeeding have been considered potential protective factors for this cancer.[11]

Among women in Vietnam, a low- and middle-income country (LMIC), decreases in birth rates, changes in lifestyles, such as an increase in physical inactivity, body weight, and adoption of western dietary patterns, as well as improvements in breast cancer screening and awareness, have been suggested as significant contributors to the increasing rates of breast cancer in Vietnam.[1,12-14] Breast cancer incidence rate (age-adjusted with the world standard population) increased by more than 86% in women in Hanoi, the capital of Vietnam, from 1991 to 1993 to 2012 (from 18.2 to 33.9 per 100,000 women).[14,15] Similarly, in Ho Chi Minh city, the largest city, the incidence rate of breast cancer in women rose by close to 60%, from 13.6 to 21.5 per 100,000 women between the periods of 1995-1998 and 2011-2015.[16-18] In 2020, the International Agency for Research on Cancer (IARC) report shows that the estimated age-adjusted annual incidence rate for breast cancer is 34.2 per 100,000 women and the estimated age-adjusted annual mortality rate is 13.8 per 100,000 women, along with an estimated 21,555 new breast cancer incident cases and 9,345 deaths in Vietnam, making breast cancer the most common cancer and the fourth leading cause of cancer-related death among Vietnamese women.[19] According to data from Ho Chi Minh cancer registry data, the age-specific incidence rate in 2011-2015 was low for women below 30 years and it increased steeply after that, reaching a peak in the 50 to 59 age group.[16] The absolute number of incident breast cancer cases in Vietnam is expected to increase sharply over the next decades due to increases in the aged population[20] and the implementation of large-scale localized, provincial and regional screening programs in this nation.[21] The aggregate 5-year breast cancer survival rate of Vietnamese breast cancer patients treated at Vietnam National Cancer Hospital and Hue Central hospital was 86.4% and 74%, respectively [22,23], lower compared with the USA (98.9% for stage I-II and 85.2% for stage III) and other developed countries (~90%) during 2009 through 2015.[24,25] Low survival rates in Vietnam are largely attributable to late-stage presentation and lack of access to advanced treatment. Delays in diagnosis and treatment are common among Vietnamese breast cancer patients.[26] Most breast cancer patients are diagnosed three to nine months after the onset of typical cancer-related symptoms.[26] Unpublished data from the five largest oncology hospitals suggest that 9.9%, 41.0%, 16.7%, 10.1%, and 22.3% of breast cancer patients are diagnosed at stages I, II, III, IV, and unknown stage, respectively. Currently, there are no national breast cancer screening guidelines, and screening

programs for early detection and diagnosis in Vietnam. Although breast cancer screening is available in center/tertiary hospitals and some private hospitals, screening procedures are not covered by the national health insurance system.

## 2. Chemotherapy in breast cancer treatment

Breast cancer is treated with a multidisciplinary approach involving three primary modalities: surgery, radiation therapy, and systemic treatments (classically classified into three categories: (1) endocrine or hormonal therapy, (2) targeted therapy including anti-HER2 therapy, and (3) chemotherapy).[27] In recent decades, adjuvant systemic treatments have been recommended following primary breast cancer treatment by surgery with or without radiation because they have contributed to reducing cause-specific mortality from breast cancer.[28-30] Although endocrine therapy and targeted therapy have made noteworthy progress, cytotoxic chemotherapy continues to have a dominant role in the clinical treatment of breast cancer. Adjuvant chemotherapy, the administration of cytotoxic chemotherapy after breast cancer surgery, is primarily used to eradicate remaining or undetectable microscopic foci of cancer cells and keep cancer from returning. Many large randomized trials have demonstrated that adjuvant chemotherapy can prevent recurrence and prolong survival.[31] The decision to offer chemotherapy and the choice of chemotherapeutic regimen is based on the absolute benefit and risk profile and is tailored to the individual breast cancer case. The expert panel at the 2013 St Gallen International Breast Cancer Conference identified indications for adjuvant chemotherapy in early-stage breast cancer (i.e., stage I to III), including histologic grade 3 tumors, high Ki-67 status, low hormone receptor status, human epidermal growth factor receptor 2 (HER2) positive or triple-negative breast cancer (i.e., a type of breast cancer with negative expression of estrogen, progesterone and HER2), high 21-gen recurrence score, high-risk 70 gene signature, and the involvement of 3 or more lymph nodes.[32] Estrogen receptor (ER)-positive, HER2-negative breast cancer with tumor size greater than one cm or at least one lymph node involvement is also offered adjuvant chemotherapy.[33] To improve surgical outcome, neoadjuvant chemotherapy, the administration of systemic therapy before surgical removal of a breast tumor,[34] might be offered for early-stage breast cancer patients who desire breast-conserving surgery and is not surgical candidates due to a high tumor-to-breast ratio. In addition, early-stage breast cancer

patients with triple-negative breast cancer (ER-negative, progesterone receptor (PR)-negative and HER2-negative), or HER2-positive cancers may be offered neoadjuvant chemotherapy.[31]

The administration of cytotoxic chemotherapy for breast cancer has evolved significantly over the past half-century. Chemotherapy regimens for breast cancer treatment generally contain chemotherapeutic agents from four distinct classes: alkylating agents, antimetabolites, anthracyclines, and anti-microtubules (taxanes).[27] These chemotherapeutic agents are usually given in multiple-drug regimens, which proved to be superior to single agents efficacy and safety.[35] First, alkylating agents are one of the oldest classes of antineoplastic drugs which impede cellular growth and induce apoptosis through DNA crosslinking.[36] Cyclophosphamide is a member of the alkylating agents routinely used for breast cancer treatment in combination with other chemotherapeutic agents to shrink preoperatively large and advanced tumors or reduce the risk of recurrence after surgery.[31] The other class of antineoplastic drugs commonly used in breast cancer treatment are antimetabolites which impair cancer growth by acting as a substitute for the precursors of RNA and DNA, disrupting purine or pyrimidine nucleoside synthesis pathways (i.e., disturbing DNA/RNA synthesis) during the S phase of the cell cycle.[37] 5-Fluorouracil (5-FU) and methotrexate are antimetabolite medications routinely used in breast cancer treatment. 5-FU inhibits the enzyme thymidylate synthase, thereby preventing thymidine synthesis, while methotrexate interferes with the enzyme dihydrofolate reductase (DHFR), thereby limiting the production of reduced folates that are needed for purine synthesis.[37] In addition, gemcitabine (a pyrimidine antimetabolite) and capecitabine (a prodrug form of 5-FU) are also used for treating breast cancer.[37] The regimen of cyclophosphamide and methotrexate plus 5-fluorouracil (CMF) was introduced in the 1970s and was the initial standard treatment for breast cancer. The effectiveness of CMF on both relapse-free and overall survival was confirmed in a cohort study of 30 years follow-up of randomized studies.[38] This study showed significant prevention of breast cancer-related deaths by CMF in rates up to 30% during the 30-year interval.[38]

The third class of chemotherapy agents, anthracyclines, are antitumor antibiotics extracted from *Streptomyces bacterium* that prevents cell replication by intercalating between base pairs of the DNA/RNA strand and inhibiting topoisomerase II, which leads to blocking DNA transcription.[39] Moreover, they also generate iron-mediated free oxygen radicals, damaging DNA, proteins, and cell membrane.

Doxorubicin and epirubicin (as topoisomerase II inhibitors) are used for early-stage breast cancer, while irinotecan (as topoisomerase I inhibitor) can be employed in the treatment of advanced and metastatic breast cancer. Since the 1980s, doxorubicin and epirubicin have been widely used in breast cancer treatment. They form the backbone of both neoadjuvant and adjuvant chemotherapy regimens for breast cancer and are frequently used with cyclophosphamide. Anthracycline-based regimen substantially improved the disease-specific survival, which was at least equivalent to or better disease-specific survival than that of the CMF. It was recognized as standard therapy for breast cancer treatment.[40]

The last class of chemotherapy agents, taxanes, are anti-microtubule agents that exert their antineoplastic effects by blocking the synthesis of the cellular microtubules within the nucleus of cancer cells, thereby disrupting mitosis, cell division, and proliferation.[41] Taxanes often used for breast cancer treatment include paclitaxel and docetaxel.[42,43] Results of a phase III randomized study, presented at the San Antonio Breast Cancer Symposium in 2005, suggested that taxane-based regimens were defined as a chemotherapy regimen in which patients received a taxane but no anthracycline, might provide an alternative to anthracycline-based regimens.[44] Since 2006, a decline in anthracycline-based chemotherapy and a sharp increase in the use of taxane-based chemotherapy to treat breast cancer has occurred in the USA.[45] In 2008, a meta-analysis of 13 randomized trials (n = 22,903 patients) showed that the addition of taxanes into an anthracycline-based regimen led to improve disease-free survival (DFS) and overall survival (OS) when compared with the standard anthracycline-based regimen in high-risk early-stage breast cancer patients.[46] Data from the 2012 Early Breast Cancer Trialist's Collaborative Group (EBCTCG) meta-analysis involving 100,000 women in 123 randomized trials suggested that the incorporation of a taxane in an anthracycline-based regimen was associated with a significant reduction in the risk of recurrence, breast cancer mortality as well as overall mortality compared with other cytotoxic regimens.[40] The taxane incorporation was found to improve outcomes irrespective of age, nodal status, tumor size, tumor grade, or estrogen receptor expression.[40] A 2019 EBCTCG meta-analysis of 37,298 women with early breast cancer in 26 randomized trials revealed that sequencing of anthracycline and taxane was associated with reduced recurrence risk compared with concurrent regimens.[47] Moreover, this meta-analysis found that increasing the dose intensity of adjuvant chemotherapy by shortening the interval between treatment cycles (dose-dense schedule), which is typically administered on an every-

week or every-two week schedule instead of the historical every-three-week schedule, was associated with improvement in DFS and OS with similar tolerability compared with standard dosing.[47] Dose-dense chemotherapy was also not linked with an increase in treatment-related adverse events,[48] due to the use of growth factors.[49] As a result, a dose-dense regime of anthracyclines and cyclophosphamide followed by sequential taxane has been recommended for adjuvant chemotherapy in breast cancer treatment. Last but not least, capecitabine (a third-generation fluoropyrimidine), platinum compound (including carboplatin and cisplatin), gemcitabine (a pyrimidine antimetabolite), vinorelbine, and poly-ADP-Ribose Polymerase inhibitors (PARPi; including olaparib and talazoparib) have been recently introduced to treating breast cancer.[37,50] They are commonly used as single agents in advanced or recurrent breast cancer cases; however, they can be given in combination with first-line chemotherapy agents.[50]

Nevertheless, no single optimal adjuvant chemotherapy regimen is universally accepted for breast cancer treatment, and preferred regimens vary by prescribing clinician, institution, and geographic region.[51] In Vietnam, a guideline for breast cancer treatment, which is partially adopted from the National Comprehensive Cancer Network (NCCN) guidelines and European societies-like European Society for Medical Oncology (ESO-ESMO) international consensus guidelines, was released in July 2018.[52] A dose-dense doxorubicin and cyclophosphamide (AC) with sequential paclitaxel, a dose-dense AC followed by sequential weekly paclitaxel, and docetaxel plus cyclophosphamide (TC) are preferred and recommended to administer with myeloid growth factor support.[52] Recommended regimens for HER2-negative, early-stage breast cancer (stage I-III) are summarized in **Table 1**. For patients with HER2-positive breast cancer, adjuvant or neoadjuvant treatments with chemotherapy and one year of trastuzumab treatment are generally recommended. Chemotherapy is also suggested in late-stage breast cancer patients (IV) if the patients have rapid disease progression, have endocrine resistance in hormone receptor-positive breast cancer, or have a large tumor burden involving visceral organs and threatening organ function.

**Table 1**: Neoadjuvant/adjuvant chemotherapy regimens for HER2-negative, stage I-III breast cancer in Vietnam

| Regimens | Cycle | Cycle duration | Dose |
|---|---|---|---|
| *Preferred regimens* | | | |
| **Dose-dense AC followed by paclitaxel (AC-P)** ¥ | 4 cycles | 2 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 4 cycles | 2 weeks | Paclitaxel 175 mg/m$^2$/week |
| **Dose-dense AC followed by weekly paclitaxel (AC-P)** ¥ | 4 cycles | 2 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 12 cycles | weekly | Paclitaxel 80 mg/m$^2$ |
| **AC followed by paclitaxel every 3 weeks (AC-P)** | 4 cycles | 3 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 4 cycles | 3 weeks | Paclitaxel 175 mg/m$^2$/week |
| **TC**¥ | 4 cycles | 3 weeks | Docetaxel 75 mg/m$^2$<br>Cyclophosphamide 600mg/m$^2$ |
| *Other recommended regimens and regimens for certain circumstances* | | | |
| **AC (doxorubicin plus cyclophosphamide)** | 4-6 cycles | 3 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| **Dose-dense AC**¥ | 4 cycles | 2 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| **AC followed by docetaxel every 3 weeks (AC-T)** | 4 cycles | 3 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 4 cycles | 3 weeks | Docetaxel 100 mg/m$^2$ |
| **AC followed by weekly paclitaxel (AC-P)** ¥ | 4 cycles | 3 weeks | Doxorubicin 60 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 12 cycles | weekly | Paclitaxel 80 mg/m$^2$ |
| **EC chemotherapy** | 8 cycles | 3 weeks | Epirubicin 100 mg/m$^2$<br>Cyclophosphamide 830 mg/m$^2$ |
| **EC followed by docetaxel every 3 weeks (EC-T)** | 4 cycles | 3 weeks | Epirubicin 90 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 4 cycles | 3 weeks | Docetaxel 100 mg/m$^2$ |
| **FAC chemotherapy** | 6 cycles | 3 weeks | 5-fluorouracil 500 mg/m$^2$<br>Doxorubicin 50 mg/m$^2$<br>Cyclophosphamide 500 mg/m$^2$ |
| **FAC followed by paclitaxel every 3 weeks (FAC-P)** | 4 cycles | 3 weeks | 5-fluorouracil 500 mg/m$^2$<br>Doxorubicin 50 mg/m$^2$<br>Cyclophosphamide 500 mg/m$^2$ |
| | 4 cycles | 3 weeks | Paclitaxel 225 mg/m$^2$ |
| **FAC followed by weekly paclitaxel (AC-P)** | 4 cycles | 3 weeks | 5-fluorouracil 500 mg/m$^2$<br>Doxorubicin 50 mg/m$^2$ |

| | | | |
|---|---|---|---|
| | | | Cyclophosphamide 500 mg/m$^2$ |
| | 12 cycles | weekly | Paclitaxel 80 mg/m$^2$ |
| **FEC chemotherapy** | 6 cycles | 3 weeks | 5-fluorouracil 600 mg/m$^2$<br>Epirubicin 90 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| **FEC followed by docetaxel every 3 weeks (FEC-T)** | 3 cycles | 3 weeks | 5-fluorouracil 600 mg/m$^2$<br>Epirubicin 100 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 3 cycles | 3 weeks | Docetaxel 100 mg/m$^2$ |
| **FEC followed by paclitaxel every 3 weeks (FEC-P)** | 4 cycles | 3 weeks | 5-fluorouracil 500 mg/m$^2$<br>Epirubicin 75 mg/m$^2$<br>Cyclophosphamide 500 mg/m$^2$ |
| | 4 cycles | 3 weeks | Paclitaxel 225 mg/m$^2$ |
| **FEC followed by weekly paclitaxel (FEC-P) ¥** | 4 cycles | 3 weeks | 5-fluorouracil 600 mg/m$^2$<br>Epirubicin 90 mg/m$^2$<br>Cyclophosphamide 600 mg/m$^2$ |
| | 8 cycles | weekly | Paclitaxel 100 mg/m$^2$ |
| **TAC chemotherapy** | 6 cycles | 3 weeks | Docetaxel 75 mg/m$^2$<br>Doxorubicin 50 mg/m$^2$<br>Cyclophosphamide 500 mg/m$^2$ |
| **CAF chemotherapy** | 4 cycles | 4 weeks | Cyclophosphamide 600 mg/m$^2$<br>Doxorubicin 60 mg/m$^2$<br>5-fluorouracil 600 mg/m$^2$ |
| **CMF chemotherapy** | 6 cycles | 4 weeks | Cyclophosphamide 100 mg/m$^2$<br>Methotrexate 30-40 mg/m$^2$<br>5-fluorouracil 400-600 mg/m$^2$ |
| **Weekly paclitaxel plus carboplatin** | 4 cycles | 3 weeks | Paclitaxel 80 mg/m$^2$ day 1, 8 and 15<br>Carboplatin AUC 6 day |
| **Docetaxel plus carboplatin (4-6 cycles)** | 4-6 cycles | 3 weeks | Docetaxel 80 mg/m$^2$ day 1, 8 and 15<br>Carboplatin AUC 6 day |
| **Weekly Paclitaxel** | 12 cycles | Weekly | Paclitaxel 80 mg/m$^2$ |
| **Daily Capecitabine** | 6-8cycles | 3 weeks | 2000-2500 mg/m$^2$ days 1-14 |

*¥ All cycles are with myeloid growth factor support*

AC: doxorubicin plus cyclophosphamide; AC-P: AC followed by paclitaxel; AC-T: AC followed by docetaxel; CAF: cyclophosphamide, doxorubicin, plus 5-fluorouracil; CMF: cyclophosphamide, methotrexate, plus 5-fluorouracil; EC: epirubicin plus cyclophosphamide; EC-T: EC followed by docetaxel; FAC: 5-fluorouracil, doxorubicin, plus cyclophosphamide; FAC-P: FAC followed by paclitaxel; FEC: 5-fluorouracil, epirubicin, plus cyclophosphamide; FEC-P: FEC followed by paclitaxel; FEC-T: FEC followed by docetaxel; TAC: docetaxel, doxorubicin, and cyclophosphamide; TC: docetaxel plus cyclophosphamide.

3. **Chemotherapy-induced toxicities**

Chemotherapy has well-established benefits on recurrence and overall survival.[31] However, chemotherapy can lead to varied acute side effects and long-term toxicities, which may affect treatment compliance and efficacy and long-term outcomes.[53] Although recent advances in chemotherapy have achieved more tolerable and safer outcomes, up to 87% of people experienced at least one side effect/adverse event during and after treatment.[54,55] The most common acute toxicities associated with almost all chemotherapeutic agents include myelosuppression (e.g. neutropenia nadir, anemia, leukopenia, thrombocytopenia), febrile neutropenia, alopecia, and gastrointestinal (GI) toxicities such as nausea, vomiting, diarrhea, stomatitis, and constipation. The hematological and GI toxicities may be seen in 40% to 80% of breast cancer patients receiving neoadjuvant or adjuvant chemotherapy.[31,56,57] In addition, admission of anthracycline might induce rarely, but serious cardiotoxicity[58,59] that might occur as acute toxicity manifested by arrhythmias or depressed ejection fraction, particularly in the left ventricle (LVEF) or might be chronic that develop years after the administration. Meanwhile, taxanes may cause neuropathy (e.g., neuropathic pain, paresthesia)[60], hepatotoxicity[61], pulmonary toxicity (e.g., interstitial pneumonitis)[62] and musculoskeletal side effects (e.g., myalgias and arthralgias).[63] If persistent, chemotherapy-induced toxicities may adversely affect the physical health, quality of life,[64] and emotional state[65] of breast cancer patients. Severe chemotherapy-induced toxicities may put cancer patients at risk of dose delay or dose reduction, treatment discontinuation,[66] and costly health care service use, [67] some may result in premature death.[68-70] Much evidence has shown that patients who received low dose chemotherapy have reduced survival rates.[71]

The occurrence of reported toxicities varies greatly because treatment-related toxicity studies were conducted in a wide range of settings, including clinical trials, population-based setting, health care claims, and single-site patient registries, each with its own limitations, including but not limited to generalizability, selection bias, data quality, or discordant reporting between clinicians and cancer patients.[72,73] Most reports of chemotherapy-induced side effects and their frequency come from clinical trials of new treatments, where patients with major comorbidity or at risk of complications are frequently excluded and safety monitoring may be more insensitive than routine clinical care.[55,74-76] Moreover,

clinicians often underreported the number and severity of toxicities experienced by patients, particularly symptomatic toxicities such as nausea, vomiting, diarrhea, constipation, anorexia, and hair loss.[54,77-79] Therefore, observed clinician-reported toxicity ratings in clinical trials may not reflect the frequency, severity, and burden of chemotherapy-induced toxicities in breast cancer patients. For example, many clinical trials of adjuvant chemotherapy for breast cancer indicated that serious adverse effects were uncommon, with 2% or less for incidences of fever, neutropenia (absolute neutrophil count (ANC) of less than $1.5x10^9$/L), and life-threatening infection, and with less than 5% for severe nausea and vomiting.[80] However, the results from a population-based data among 12,239 American women under 64 years of age with newly diagnosed breast cancer between 1998-2002 reported that women who received chemotherapy during the 12 months after breast cancer diagnosis were more likely to be hospitalized or to visit the emergency room for chemotherapy-related severe adverse effects (16%).[81] The major serious adverse effects in the 4,075 chemotherapy recipients resulting in an emergency room visit or hospitalization during the year after their breast cancer diagnosis were fever or infection (8.4%), neutropenia or thrombocytopenia (5.5%), dehydration or electrolyte disorders (2.5%), nausea, emesis, or diarrhea (2.4%), anemia (2.2%), constitutional symptoms (2%), deep venous thrombosis or pulmonary embolus (1.2%) and malnutrition (0.9%).[81] Moreover, data from the Surveillance, Epidemiology, and End-Results (SEER) - Medicare linked database of 35,060 women aged ≥65, and diagnosed with stages I-IV breast cancer from 1991 to 1996 revealed that over 9.0% of women who received chemotherapy were admitted to hospital for serious adverse effects of neutropenia, fever, thrombocytopenia and others (e.g. dehydration, infection, anemia, delirium) compared with 0.5% of women with breast cancer who did not receive chemotherapy.[82] The rates of respective serious adverse effects were 6.3%, 8.1%, 12.3%, and 13.2% in those treated with chemotherapy during seven months after diagnosis of breast cancer,[82] and these rates were significantly associated with comorbidity score and use of anthracycline-based regimens, but did not vary by age. [82] The percentage of chemotherapy-induced serious adverse effects among the two population-based studies, including a study 12,239 American women aged < 64 and a study from the SEER-Medicare linked database of 35,060 women aged ≥65, was more common than previous clinical trials reported.[80] The results of population-based data and clinical trials reported were from periods when the most frequently used regimens for adjuvant chemotherapy were AC (doxorubicin plus

cyclophosphamide) and CMF (cyclophosphamide, methotrexate, and 5-FU). The taxane-based regimens and Granulocyte colony-stimulating factor (G-CSF) administration were not considered in these studies.[83] Currently, there have been very few observational studies on chemotherapy-induced toxicity in breast cancer patients during the era when chemotherapy includes anthracycline-, taxane-, and non-anthracycline-based regimens.

In breast cancer patients receiving chemotherapy, neutropenia and anemia are commonly reported as dose-limiting hematological toxicities. Neutropenia is defined as a decrease in absolute neutrophil count (ANC) of less than $1.5x10^9$/L; ANC of less than $0.5x10^9$/L is considered as severe neutropenia (as is grade 4 neutropenia according to NCI CTCAE classification). The majority of women with breast cancer who received combination chemotherapy experienced neutropenia during 10-14 days after each cycle, which typically resolved before the next course of chemotherapy.[80] The hematological toxicity, especially in patients with previous comorbidities, inclines toward severe development, resulting in fever or potentially lethal infections. It may lead to delays in treatment administration, dose reduction and discontinuance. The percent risk of grade 4 neutropenia widely ranged from 2% to 90%.[80,84-86] Febrile neutropenia (i.e., a fever during a period of significant neutropenia) is a severe complication of chemotherapy. For most chemotherapy regimens used in breast cancer treatment, the risk of febrile neutropenia is less than 2%. Patients receiving docetaxel-containing regimens (e.g., the regimens of TAC or TC) or dose-dense chemotherapy (dose-dense AC followed by docetaxel) have at least a 20% risk of neutropenia.[87,88] The clinical guidelines recommend using G-CSF to reduce the duration of neutropenia and the risk of neutropenic fever.[89] Second, anemia is caused by the deficiency of red blood cells or hemoglobin, and is a common complication of myelosuppressive chemotherapy, which results in functional impairment and reduction of quality of life.[90,91] However, anemia is a common consequence of cancer itself.[90-92] Adjuvant chemotherapy leads to an increase in the incidence and severity of anemia.[91,93] For example, in a series of 310 patients with stage II and III breast cancer treated with adjuvant AC chemotherapy at eight U.S. centers, 40.0% of patients experienced moderate to severe anemia (<10g/dl), and 31.3% were anemic prechemotherapy. Among the patients with mild anemia prechemotherapy, 61.9% developed moderate to severe anemia during chemotherapy; Among the patients with normal prechemotherapy hemoglobin levels (≥12g/dl), 88.3% developed some degree of

anemia during chemotherapy and 27.7% developed moderate to severe anemia.[94] However, chemotherapy-induced anemia is often underreported. It is only documented when severe, necessitating transfusion[95] because physicians treating patients with adjuvant chemotherapy for breast cancer often focus on potentially life-threatening toxicities such as febrile neutropenia or toxicities that require immediate symptomatic intervention or dose reductions, such as mucositis, diarrhea, or neuropathy.[95] The incidence and severity of anemia depend on the type, schedule, and intensity of chemotherapy.[91] Among 702 patients who received adjuvant breast cancer chemotherapy at four Breast Cancer Agency centers in 2002 and 2003, the reported high-grade (<90g/dl) anemia was seen in 26.5% (FEC), 9.7% (CMF), 7.3% (AC followed by paclitaxel or docetaxel), 6.5% (CAF) and 1.3% (AC) treated patients, respectively.[95] Type of chemotherapy, BMI, age, and hemoglobin at baseline significantly impacted the risk of anemia among breast cancer patients treated by neoadjuvant or adjuvant chemotherapy.[96] For the treatment of chemotherapy-induced anemia, erythropoietin-stimulating agents (ESAs) or concentrated red blood cell transfusions are used. The transfusions of concentrated red blood cells are often recommended for patients with symptomatic anemia to increase baseline hemoglobin ≥ 10 mg/dl. The administration of ESAs is not recommended in patients with anemia outside the period of administration of chemotherapy due to an increased risk of thromboembolic events (TEE) and an increased risk of hypertension and headache.

Among GI toxicities, nausea and vomiting remain as one of the most unpleasant and feared adverse effects of chemotherapy. They are significantly associated with poor adherence to chemotherapy and negatively impact the quality of life.[50,97] More than 70% of patients receiving chemotherapy may experience nausea and vomiting[98]; however, the incidence and severity of these side effects vary according to the chemotherapy regimen, dosage, duration, and risk factors.[99-107] Emetogenicity refers to the likelihood of a chemotherapy agent causing emesis in the absence of antiemetic prophylaxis.[108] The dose-dense AC is considered highly emetogenic (experienced by >90% of patients), whereas regimens containing either taxane (docetaxel or paclitaxel) or carboplatin are moderately emetogenic (experienced by 30 to 90 % of patients).[109] Patients with younger age, Asian race, having anxiety, high pretreatment expectancy of severe nausea/vomiting, history of nausea/vomiting, history of morning sickness, and low alcohol intake have an increased risk of more

frequent and severe chemotherapy-induced nausea and vomiting.[110] However, all patients receiving neoadjuvant/adjuvant chemotherapy require antiemetic therapy tailored to the specific treatment regimens.[108] Diarrhea is also a common side effect of chemotherapy with varying degrees across different chemotherapeutic agents.[111] Diarrhea is more prevalent in patients who received 5-fluorouracil, capecitabine, irinotecan, methotrexate, or cisplatin and is less prevalent and severe in patients treated with taxane and anthracycline.[98,111] Diarrhea can occur in 50-80% ( ≥ 30% severe grades) of patients who received some combination therapies of irinotecan and fluoropyrimidines.[112] For breast cancer treatment, the risk of grade 3 or 4 diarrhea occurred in about 1% to 9% of patients receiving standard chemotherapy.[113] Patients aged 70 years or older treated with irinotecan have an increased risk of grade 3 or 4 diarrhea. However, the treatment with the every three-week schedule was associated with a lower rate of severe diarrhea.[114] The primary mechanism of diarrhea associated with chemotherapy is multifactorial, resulting from acute damage to the intestinal mucosa, including loss of proliferating intestinal epithelium, disruption of the mucosal barrier, and impaired water and electrolyte absorption from the luminal wall.[111] Chemotherapy-induced diarrhea results in dose reduction, delay, or even treatment discontinuance. Therefore, the American Society of Clinical Oncology (ASCO) guidelines recommend using loperamide or loperamide along with oral antibiotics (as prophylaxis for infection) for mild to moderate diarrhea. Treatment with immediate octreotide therapy and antibiotics when oral antidiarrheals are ineffective in patients with grade 3 or 4 diarrhea.[56] Recently, the Multinational Association of Supportive Care in Cancer and International Society of Oral Oncology (MASCC/ISOO) Clinical Practice Guidelines for managing mucositis suggests that probiotics containing *Lactobacillus* species may be beneficial for the prevention of diarrhea in cancer patients.[115]

## 4. The gastrointestinal (GI) microbiota in breast cancer patients

The human microbiota is composed of commensal bacteria, archaea, viruses, fungi, protists, and other microorganisms that inhabit the epithelial barrier surfaces of the human body.[116] The microbiota communities are shaped by colonization at birth time[117], type of birth delivery[118], human genetics[119], individual lifestyles[120], the incidence of certain diseases, and exposure to antibiotics.[121] The microbial composition evolves during the first few years of human life.[122] It remains ecologically and functionally

constant throughout adult life, although it can be affected by diet, changes in lifestyle, disease, and disease treatments.[123,124] In humans, the GI microbiota plays a critical role in disease prevention through maintaining barrier homeostasis,[125,126] providing protection against pathogen overgrowth,[127] maturing, and continuously educating immune response,[128,28] influencing epithelial hyperproliferation[129] and supporting vascularization in the GI tract.[130] The GI microbiota also plays a vital role in regulating intestinal endocrine functions,[131] providing a source of energy biogenesis,[132] biosynthesizing vitamins,[133] regulating neurologic signaling and neurotransmitters,[134] metabolizing bile salts,[135] reacting to or modifying specific drugs and eliminating exogenous toxins.[136]

The GI microbial profile is like a fingerprint because each person's composition of the GI microbiota community is unique. The GI microbiota in a healthy adult has a high diversity and a large number of beneficial microbes, which are dominated by two bacterial divisions, *Bacteroidetes* and *Firmicutes,*[137] and include smaller proportions of *Actinobacteria*, *Proteobacteria*, *Verrucomicrobia, Tenericutes,* and *Lentisphaerae* at the phylum level.[138] Most common genera of the GI microbiota are *Bacteroides, Clostridium, Faecalibacterium, Eubacterium, Ruminococcus, Peptococcus, Peptostreptococcus, Lactobacillus, Streptococcus, Streptomyces,* and *Bifidobacterium.*[138] The GI microbiota can become compositionally unstable and less diverse in some situations, such as age-related declines in immunocompetence[139] and specific health conditions (e.g., inflammatory bowel disease and colorectal cancer[140], obesity[140], and extra-digestive diseases such as allergies and autism[141,142]). The GI microbiota might potentially be harmful due to the change of composition.[143] So far, the human GI microbiome in stool samples has been the most common human GI microbiota and serves as a model for understanding host-microbiota interactions and diseases.[144,145]

The characteristics of GI microbiota among breast cancer patients are unknown at present due to some limitations in studies targeting the relationship between breast cancer and GI microbiota. Most studies simply focused on associations and changes in GI microbiota characterized by pathological and clinical features of breast cancer and established risk factors such as BMI and menopausal status rather than cause-and-effect. Human studies of GI microbiota characteristics among breast cancer patients are summarized in **Appendix 1**.

Notably, a series of published studies used the biobank set-up by the research group of Goedert from a population-based case-control study among 48 postmenopausal breast cancer patients with pretreatment and non-antibiotic exposure and 48 similar women with normal-mammography.[146] In this pilot study published in 2015, Goedert and colleagues found that breast cancer patients had a high fecal beta-diversity but an estrogen-independent lower alpha-diversity than healthy controls. Particularly at the family level, breast cancer patients had higher *Clostridiaceae*, *Faecalibacterium,* and *Ruminococcaceae* and lower *Dorea* and *Lachnospiraceae*.[146] In 2018, Goedert and co-workers investigated differences in the composition of immune-recognized gut microbiota in the same population-based case-control study.[147] Case and controls differed significantly the composition of the IgA-positive microbiota and the IgA-negative microbiota fractions. Postmenopausal breast cancer patients with IgA-positive showed an estrogen-independent lower richness and alpha diversity, significantly more marked than patients with the IgA-negative microbiota, suggesting that the relationship between breast cancer and the gut microbiota through estrogen-independent pathways might be related to the immune pathway.[147] In recent studies applying real-time quantitative polymerase chain reaction (qPCR) assays with specifically designed primers, the abundance of the deoxyribonucleic acid (DNA) coding for baiH (baiH ORF codes for 7-HSDH, a key enzyme in lithocholic acid biosynthesis), LdcC (constitutive lysine decarboxylase) and CadA (acid-inducible lysine decarboxylase) in several bacterial species was measured and significantly decreased in breast cancer patients, particularly early-stage breast cancer patients (stage 0-I).[148,149] These results suggested that changes in GI microbiome in early-stage breast cancer might alter bacterial metabolism, resulting in a decreased production of anti-cancer bacterial metabolites like lithocholic acid or cadaverine.[148,149]

The fecal microbiome also varies according to different clinical stages and histological grades of breast cancer among breast cancer patients. A case-only study on 31 patients with early-stage breast cancer showed that the percentage and the absolute number of *Bacteroidetes*, *Clostridium leptum cluster*, *Clostridium coccoides cluster*, *Facecalibacterium prausnitzii,* and *Blautia* species were significantly higher in clinical stage II and III breast cancer patients than in clinical stage 0 and I patients. Luu and colleagues indicated that *Blautia* species was also significantly associated with more severe histological grades. These results suggested that the GI microbiota may play a role in breast cancer

progression.[150] Additionally, some recent studies have demonstrated that microbial profile was associated with clinicopathological factors in breast cancer. Yang, Wang and co-workers reported that patients with malignant breast tumors had a distinct enrichment of gut microbiome by different clinicopathological characteristics, including ER, PR, Ki-67 levels, HER2 status, and tumor compared with patients with benign tumors.[151] Enrichment of *Megasphaera* was observed in patients with ER+ and HER2-positive tumors. Members of the family *Prevotellaceae* were abundant in patients with PR+ or ER+, while some microbes, including *Hydrogenophilus, Lactobacillus,* and *Acinetobacter,* were highly abundant in breast cancer patients with PR- and ER- tumors.[151] Wu and colleagues also found that HER2 status and age at menarche had significant associations with GI alpha diversity indexes and specific microbial composition.[152] These findings suggested that each breast cancer type might have type-specific gut microbiome communities, which need to confirm in studies with larger sample sizes.[151]

According to an analysis of established breast cancer risk factors, Luu and colleagues reported that overweight and obese patients had a significant decline in the number of total *Firmicutes*, *Facecalibacterium prausnitzii*, *Eggerthella lenta* and *Blautia* species when compared to patients with normal BMI.[150] Moreover, this inverse association of BMI and GI microbial diversity and composition was supported by the finding in a presurgical weight-loss trial among 32 overweight or obese women diagnosed with early-stage breast cancer (stage 0 to II).[153] Breast cancer patients with a high relative abundance of *Akkermansia muciniphila* had lower fat mass when comparing with low *Akkermansia muciniphila* relative abundance patients at baseline.[153] A small, case-control study of 18 breast cancer patients and 30 healthy women, in which the investigators performed Gram-stain, morphological, and biochemical analysis, showed that premenopausal breast cancer patients had a significant increase in the number of *Enterobacteriaceae*, aerobic *Streptococci*, *Lactobacilli,* and anaerobic bacteria including *Clostridia*, *Bacteroides* and anaerobic *Lactobacilli* in stools when compared with healthy controls.[154] However, Zhu and colleagues applied Illumina sequencing to detect microbiome in fecal samples and revealed that the diversity and composition of gut microbiome differed significantly between postmenopausal breast cancer patients and healthy controls while were the diversity and composition of gut microbiome similar between premenopausal breast cancer patients and healthy controls. They found that 38 species were enriched in postmenopausal breast cancer patients, including *Escherichia coli,*

*Citrobacter koseri, Acinetobacter radioresistens, Enterococcus gallinarum, Shewanella putrefaciens, Erwinia amylovora, Actinomyces sp. HPA0247, Salmonella enterica*, and *Fusobacterium nucleatum.*[155] In addition, a study conducted by Hou and colleagues also showed that the gut microbiota in premenopausal patients differed from that in postmenopausal breast cancer patients.[156] Hou and colleagues identified 14 microbial makers that varied by the menopausal status of the individual with breast cancer. Premenopausal breast cancer patients had significantly higher levels of *Anaerostipes* and *Bacteroides fragilis*, whereas postmenopausal breast cancer patients had significantly higher *Proteobacteria* and *Klebsiella pneumoniae.* The higher levels of four bacterial taxa were not affected by age. The alpha diversity was significantly reduced in premenopausal breast cancer patients, and the beta diversity only differed significantly between breast cancer patients and controls.[156] Moreover, He and colleagues found the composition and symbiosis of gut microbiota in premenopausal women with breast cancer changed significantly, with a reduced abundance of short-chain fatty acids (SCFA)-producing bacteria and significantly lower levels of intestinal SCFA-producing enzymes, in comparison with that in premenopausal healthy women.[156] They emphasized the potential reference value of specific gut microbiota such as *Pediococcus* and *Desulfovibrio* to diagnose premenopausal breast cancer.

Overall, despite the limited sample size and inability to disentangle the causality of association, these above-mentioned studies illustrated that the diversity and composition of the GI microbiome in breast cancer patients were varied by clinical features and risk factors as well as significantly different from that of healthy controls, indicating that certain bacteria might be associated with cancer development, progress or with different responses to therapy.[145]

## 5. Chemotherapy changes the GI microbiota

*Clinical studies*

The structure of the GI microbiota during and after chemotherapy treatment may exhibit marked changes in diversity and composition.[157] To date, over ten (11) clinical studies (four with acute myeloid leukemia (AML), three with colorectal cancer (CRC), two with Non-Hodgkin's lymphoma, one with ovarian cancer patients, and one with a group of various cancers) assessed the impact of chemotherapy during treatment. They found that chemotherapy modulates the gut microbiome of patients with cancer,

which is summarized in **Supplementary Table S2**. Although most studies were conducted on patients with other types of cancers, the findings might apply to breast cancer patients.

A small and earliest study of acute leukemia cases used a standard microbiological culture technique to detect fecal microbiota and reported no significant changes in the numbers of bacteria or *Candida* species during chemotherapy.[158] However, a tremendous decrease in the number of anaerobic bacteria (i.e., *Bacteroides* species, *Clostridium cluster XIVa, Faecalibacterium prausnitzii*, and *Bifidobacterium* species) and *Streptococci* species, as well as a disturbed balance between aerobic and anaerobic bacteria were observed in patients with AML following chemotherapy in the other study.[159] Many recent clinical studies performed 16S rRNA gene amplicon sequencing to analyze chemotherapy-induced changes in fecal microbiota. Galloway-Peña and colleagues conducted three sequential studies in which oral swabs and stool samples were collected biweekly from baseline until neutrophil recovery following induction chemotherapy among patients with AML.[160-162] They observed a consistent decrease in alpha diversity of the gut microbiome throughout induction chemotherapy in both the oral and stool samples.[160,162] The study with a small sample size (n = 34) found statistically significant increases in *Lactobacillus* in both oral and stool samples, while significant decreases were primarily observed in anaerobic genera, such as *Blautia, Prevotella*, and *Leptotrichi.*[160] Another study, which enrolled a large sample size (n = 97), reported that *Clostridiales* and *Blautia* were significantly higher at baseline than at the end-of-study stool samples.[162]

Montassier and colleagues examined the impact of chemotherapy on the GI microbiome among patients with non-Hodgkin lymphoma in two studies and found that patients' fecal microbiota exhibited a rapid decrease in overall diversity and a distinct disruption in bacteria composition.[163,164] The study that collected pre- and post-chemotherapy fecal samples from eight adult patients found that overall diversity decreased significantly in evenness as measured by the Shannon index and richness measured by phylogenetic diversity.[163] At the phylum level, the fecal microbiota of patients after chemotherapy showed a decrease in phylum *Firmicutes* and an increase in phylum *Bacteroidetes*. Compared to before chemotherapy, there was a profound decrease in *Blautia*, *Faecalibacterium* and *Roseburia*, and *Bifidobacterium*, and increases in *Bacteroides* and *Escherichia* at the genus level after chemotherapy.[163] A subsequent study showed the alpha diversity in stool samples collected after chemotherapy was

significantly lower than that collected before chemotherapy among 28 patients with non-Hodgkin lymphoma undergoing a consecutive 5-day, myeloablative chemotherapy regimens before human hematopoietic stem cell transplantation (HSCT). Significant decreases in abundances of phylum *Firmicutes* and *Actinobacteria* and marked increases in abundance of phylum *Proteobacteria* were observed in fecal samples collected after chemotherapy. At the genus level, post-chemotherapy fecal samples exhibited decreased abundances of *Ruminococcus*, *Oscillospira*, *Blautia*, *Lachnospira*, *Roseburia*, *Dorea*, *Coprococus*, *Anaerostipes*, *Clostridium*, *Collinsella*, *Adlercreutzia,* and *Bifidobacterium*, but increased abundances of *Citrobacte*r, *Klebsiella*, *Enterococcus*, *Megasphaera,* and *Parabacteroides*.[164]

Three studies conducted on CRC patients reported changes in the gut microbiome pre- and post-different chemotherapy regimens.[165-167] One study collected fecal samples multiple times after each treatment cycle of capecitabine plus oxaliplatin (CapeOx) and found an increased ratio of *Bacteroidetes* to *Firmicutes,* an increase in the abundance of *Collinsella, Anaerostipes, Bilophila, Comamonas, Weissella, Bacteroides,* and *Eggerthella*, and a decrease in the abundance of *Escherichia-Shigella, Morganella, Pyramidobacter,* and *Proteus* after CapeOx therapy.[165] Additionally, this study observed the "rebound effect" of chemotherapy-adapted bacteria. The abundance of *Dorea, Ruminococcaceae UCG-010, Streptococcus, Prevotella 9, Mogibacterium,* and *Roseburia* fluctuated after one or two cycles but recovered after five cycles.[165] One study examined the effects of three chemotherapy regimens on the gut microbiome, namely XELOX (oxaliplatin plus capecitabine), FOLFIRI (Irinotecan, leucovorin, plus 5-fluorouracil), and FOLFIRI regimen plus cetuximab. They discovered that the community structure of gut bacteria and fungi changed in chemotherapy and varied according to the different regimens of chemotherapy.[167] In postoperative CRC patients treated with the XELOX regimen, the abundances of *Veillonella, Humicola, Tremellomycetes,* and *Malassezia* were increased. The abundances of *Faecalibacterium, Clostridiales, Phascolarctobacterium, Humicola,* and *Rhodotorula* were decreased, while *Candida, Magnusiomyces, Tremellomycetes, Dipodascaceae, Saccharomycetales, Malassezia* and *Lentinula* were increased in advanced CRC patients treated with the FOLFIRI regimen. Moreover, in comparison with those treated with the FOLFIRI regimen alone, the abundances of *Humicola, Rhodotorula*, and *Magnusiomyces* were decreased in advanced CRC patients treated with the FOLFIRI

regimen combined with cetuximab, whereas those of *Candida, Tremellomycetes, Dipodascaceae, Saccharomycetales, Malassezia,* and *Lentinula* were increased.[167] The third study compared the composition of gut microbiota in four groups, including healthy individuals (n = 33), patients diagnosed with CRC before treatment (n = 17), surgically treated CRC patients (n = 5), and chemotherapy-treated CRC patients (n = 14).[166] This study reported that, at the genus level, *Sutterella* species and *Veillonella dispar* were significantly associated with CRC patients who undertook 6-8 cycles of the chemotherapeutic cocktail of oxaliplatin and tegafur, but not in the other three groups. In addition, two species, *Prevotella copri* and *Bacteroides plebeius,* were only enriched in patients treated with chemotherapy whereas alpha diversity of the gut microbiome was lower in CRC patients who received surgery, compared with the other three groups.[166]

One study assessed changes in the gut microbiome of ovarian cancer patients pre- and post-surgery and per- and post-chemotherapy with TC (paclitaxel plus carboplatin) and TP (Cisplatin plus paclitaxel).[168] A comparison of pre- and post-chemotherapy found an increase in the abundance of phyla *Bacteroidetes* and *Firmicutes* and a decrease in the abundance of phylum *Proteobacteria* after chemotherapy. Furthermore, the abundance of some forms of anaerobic bacteria such as *Bacteroides, Collinsella*, and *Blautia* species increased after multiple cycles of chemotherapy, whereas the abundance of *Veillonella, Lachnospiraceae unclassified, Roseburia, Akkermansia*, and *Bifidobacterium* increased during the first to third cycles and decreased during subsequent cycles. Interestingly, this study found significant decreases in the abundance of phyla *Bacteroidetes* and *Firmicutes* and increases in the phylum *Proteobacteria* after surgery.[168]

A study of 17 ambulant patients diagnosed with various cancers receiving different chemotherapy regimens, with or without concomitant antibiotics, compared with 17 healthy controls matched for gender, age, and lifestyle, showed a lower total bacteria in cancer patients.[169] The patients' GI microbiota was severely disrupted following chemotherapy, characterized by decreases of *Bifidobacteria* and *Clostridium cluster XIVa* and increases of *Bacteroides* and *Clostridium cluster IV*. Patients who received antibiotics had higher bacterial abundances than those without concomitant antibiotics. The authors reported that a coincidental development of *Clostridium difficile* in patients immediately after chemotherapy was

accompanied by a reduction of *Bifidobacterium*, *Lactobacillus*, *Veillonella, Escherichia coli* or *Shigella* species, and particularly *Faecalibacterium prausnitzii* (from 9% to undetected).[169]

*Mouse model studies*

Similar to human studies, much evidence has been published concerning the chemotherapy-induced modifications in intestinal microbiota from mouse model studies. Although over 85% of sequencing representing genera in mice was not detected in humans, the results from mouse model studies still provide insights into the direct effect of chemotherapeutic agents on the structure of human GI microbiota.[157] These mouse model studies used chemotherapeutic agents including 5-fluorouracil (5-FU)[170-175], cyclophosphamide (CTX) [176-178], gemcitabine[179], methotrexate (MTX)[180], and irinotecan (CTP-11)[181-184] are summarized in **Supplementary Table S3**.

The influence of 5-FU treatment on mice intestinal microbiota was reported in a study conducted in 2003 using culture-based techniques. 5-FU treatment caused an increase in the number of Gram-negative anaerobes and an increased translocation to mesenteric lymph nodes.[170] Using microbiological culture techniques for colon samples and real-time PCR for fecal samples, significant changes to intestinal flora after 5-FU administration in rats were detected. 5-FU treatment resulted in decreased abundance of *Bacteroides* species, *Lactobacillus* species, and *Enterococcus* species, whereas induced increases in *Clostridium* species, *Staphylococcus* species, and *Escherichia coli.* The study also suggested that *Enterococcus faecalis*, *Streptococcus pneumoniae*, *Lactobacillus acidophilus*, *Bifidobacterium lactis*, *Clostridium botulinum,* and *Staphylococcus epidermidis* were susceptible to 5-FU.[171] With the wide use of the sequencing of a specific region within the bacterial 16S rRNA gene to detect microbiome in fecal samples, several mouse model studies found that 5-FU administration greatly diminished the community richness and diversity and altered microbial composition in fecal microbiota.[172-174] Hamouda and colleagues observed that 5-FU treatment decreased the abundance of intestinal phylum *Firmicutes* but increased the abundance of phyla *Bacteroidetes* and *Verrucomicrobia.* They further showed that these responses could be blocked by co-administered ampicillin.[172] Similarly, Li and co-workers found a significantly decreased ratio of *Firmicutes* to *Bacteroides* in both cecum contents and feces due to a reduction in phylum *Firmicutes* after 5-FU treatment.[173] In addition, 5-FU treatment

significantly decreased *Actinobacteria*, *Alistipes, Lactobacillus* and increased the relative abundance of *Enterobacter*, *Lachnospiraceae Nk4 A136 group*, *Escherichia-Shigella*, *Alloprevotella*, *Bacteroides*, *Rikenella*, *Blautia*, *Mucispirillum*, and *Mycoplasma*. 5-FU treatment induced profound losses of several bacterial species such as *lactobacillus animalis* and *Helicobacter hepaticus* and dominance of *Lachnospiraceae bacterium COE1*, *Bacteroides vulgatus*, *Mycoplasma sualvi*, and *Escherichia coli*.[174] In contrast to these mentioned changes in intestinal and fecal microbiota, Vanlancker et al. collected mucosal and luminal samples from an *in vitro* mucosal simulator of the human intestinal microbial system and found that 5-FU only displayed a minor impact on colon microbial functionality and composition.[175]

Similar to 5-FU, treatment with cyclophosphamide-induced changes to bacterial diversity and composition, significantly increased intestinal permeability, and led to the translocation of some bacterial species into secondary lymphoid organs.[176-178] Viaud and colleagues reported a reduction of species of phylum *Firmicutes* distributed within four genera and groups (*Clostridium* cluster XIVa, *Roseburia*, *Lachnospiraceae*, *Corprococcus)* and a declined abundance of *Lactobacilli* species and *Enterococci* species in the small intestine mucosa of cyclophosphamide-treated mice. In addition, cyclophosphamide treatment caused translocation of several Gram-positive bacterial species, including *Lactobacillus johnsonii*, *Lactobacillus murinus,* and *Enterococcus hirae,* into mesenteric lymph nodes and spleens due to increased intestinal permeability. [176] In line with that, Yang and colleagues reported that administration of cyclophosphamide altered mucosal barrier and colonization resistance, increased intestinal permeability, and the bacterial counts of potentially pathogenic bacteria (*Escherichia coli*, *Enterobacteriaceae*, *Pseudomonas,* and *Enterococci)* after treatment, especially at high dose. Xu et al. supported that cyclophosphamide treatment led to a low bacterial community alpha diversity and increased *Firmicutes* to *Bacteroides* ratio due to a decreased relative abundance of *Bacteroides* in fecal samples. The authors also reported that administration of cyclophosphamide significantly increased the relative abundance of classes *Bacilli*, *Clostridia*, *Coriobacteriia*, and *Mollicutes,* and the family *Lacnospiraceae, Coriobacteriaceae*, *Lactobacillaceae*, and *Staphylococcaceae* as well as decreased the class *Bacteroidia* and *AlphaProteobacteria,* and the family *Prevotellaceae*, S24-7, *Alcaligenaceae*, and

*Rhodospirillaceae,* and disappeared *Verrucomicrobia* and *Streptococcaceae* in cyclophosphamide-treated mice.[178]

Only one recent study has reported the influence of gemcitabine (a pyrimidine antimetabolite like 5-FU) treatment on the intestinal microbiota profile of pancreatic cancer xenografted mice.[179] Administration of gemcitabine-induced a considerable reduction in the proportion of two of the most dominant phyla (i.e., Gram-positive *Firmicutes* and Gram-negative *Bacteroidetes)* and shifted to a significant increase of phylum *Proteobacteria* (mainly *Escherichia. coli* and *Aeromonas hydrophila*) and *Verrucomicrobia* (mainly *Akkermansia muciniphila*), while almost disappeared the genus of *Erysipelatoclostrium*, *Alistipes,* and *Anaerotruncus*.[179] The gemcitabine induced-alternation in intestinal microbiota suggested a shift towards an inflammation-related bacterial profile.[179] Similar to gemcitabine, only one study has been published investigating the influence of methotrexate (which belongs to a class of antimetabolites) on mice's fecal microbiota. Methotrexate treatment induced substantial decreases in most intestinal bacteria and increases in the relative number of enteropathogenic bacteria such as *Bacteroides*, *Lactobacilli*, *Enterococci*, and *Enterobacteriaceae.*[180]

Using the standard culture method, in a study (2007), Stringer and colleagues investigated microflora changes at 6, 12, and 24 hours in mouse colon after treatment with Irinotecan (CTP-11; belongs to a class of topoisomerase I inhibitors). They reported that CTP-11 treatment caused increases in genera *Escherichia*, *Clostridium, Enterococcus*, *Serratia*, *Staphylococcus*, *Bacillus*, *Peptostreptococcus*, and *Lactobacillus* in mice's colon. In addition, the authors also found that several bacteria at genus levels, including *Bacillus*, *Bifidobacterium*, *Clostridium*, *Veillonella*, and *Actinobacillus*, were not detected during the follow-up period, while four genera, including *Prosteus*, *Peptostreptococcus*, *Clostridium,* and *Enterobacter*, increased in fecal samples.[181] In a subsequent report ( 2008), Stringer and co-workers reported an increase in the level of the β-glucuronidase-producing bacteria such as *Staphylococcus* species, *Clostridium* species, and *Escherichia coli*, whereas the levels of two beneficial bacteria were decreased for *Lactobacillus* species, and *Bifidobacterium* species.[182] In a colon cancer-rat model, CTP-11 treatment changed the intestinal microbiota composition, with increases in the abundance of *Clostridium* clusters XI, and *Enterobacteriaceae* after dose-intensive treatment. Following a low-dose

regimen with a combination of irinotecan and 5-FU, *Clostridium* cluster XI, *Clostridium* cluster XIVa, and *Enterobacteriaceae* increased, while *Clostridium* cluster IV declined in cecum samples. In particular, the authors found that CTP-11-induced changes in fecal microbiota were less pronounced than those in cecum microbiota.[183] In a recent study of 48 male Sprague-Dawley rats receiving administration of CTP-11, 5-FU, or Oxalipan, Forsgård et al. found that administration of CTP-11 led to a significantly decreased relative abundance of Actinobacteria, *Bacteroides*, and Synergistetes, and significant increases in *Fusobacteria* and *Proteobacteria* at the end of the experiment. 5-FU treatment induced an increase in *Verrucomicrobia*, while oxaliplatin caused a rise of *Proteobacteria* when compared with the control group at the end of the experiment.[184]

In addition to chemotherapy, the diversity and composition of the GI microbiota may exhibit marked changes as a result of surgery, administration of antibiotics, whole-body irradiation, cachexia, and systemic tumor-promoting inflammation.[185,186] In general, early-stage breast cancer patients usually undergo breast-conserving surgery (lumpectomy) or mastectomy to the breast and regional nodes. Then postoperative radiation therapy and adjuvant treatment may be offered to eradicate residual tumor or undetectable microscopic foci of cancer cells, reduce the risk of locoregional recurrence and improve breast cancer-specific and overall survival. Notably, breast cancer patients are regularly prescribed prophylactic antibiotics to combat infections before, during and after cancer treatments.[187,188] Many studies have confirmed gut microbiota changes before and after surgery and prophylaxis antibiotics before and during chemotherapy, but no investigation was conducted in breast cancer patients.[160-166,168,169,189] The gut microbiota dysbiosis increases harmful microbe populations and subsequently contributes to systemic side effects related to chemotherapy.

6. **The GI microbiota and chemotherapy-induced toxicity**

As mentioned previously, most pre-clinical and clinical studies attempted to elucidate gut microbiota changes caused by chemotherapy but associations between chemotherapy-induced toxicity and the gut microbiome prior to, during, and after treatment were less investigated. A framework called TIMER (Translocation, Immunomodulation, Metabolism, Enzymatic degradation, and Reduced diversity and ecological variation) was suggested by Alexander and colleagues to describe the interactions between the microbiome and cancer drugs.[190] The GI microbiota modulate chemotherapeutic agents through key mechanisms including direct cytotoxicity, bacterial translocation, immune response, and drug metabolism that may facilitate drug efficacy, abrogate and compromise anticancer effects[191], but may also mediate or exacerbate the systemic toxicity effects.[192]

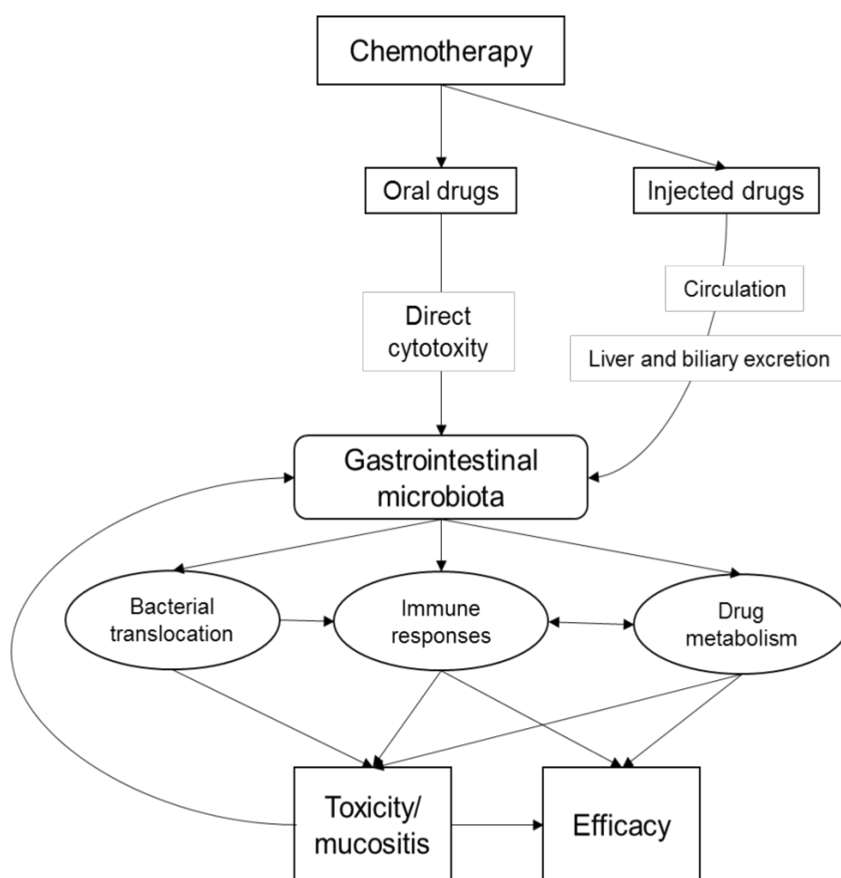



**Figure 1**: Plausible links between the GI microbiota and chemotherapy adapted the TIMER framework suggested by Alexander et al.[190]

Plausible links between the GI microbiota and chemotherapy-induced toxicity are illustrated in **Figure 1**. Since the GI microbiota plays a critical role in the human body including maintaining barrier

homeostasis, and protecting against pathogen overgrowth, the microbiota might reduce chemotherapy-induced toxicity. However, some chemotherapeutic agents such as doxorubicin and cyclophosphamide may induce translocation of commensal microbes from barrier surfaces due to the cause of shortening of villi mucosal, directly damaging the integrity of the mucosal epithelium and increasing intestinal permeability. This permits commensal microorganisms and pathogens to cross the intestinal barrier and enter the lymphatic system and the circulation. The translocated bacteria activate innate immune cells and initiate local and systemic inflammation.[192] On the other hand, the direct interactions between the GI microbiota and oral drugs can lead to bacterial modification of pharmaceuticals. Such changes may liberate secondary metabolites and toxic compounds secreted into the circulation and excreted by the kidney, which may in turn cause toxicity.[190] In addition, injected drugs may induce changes in the diversity of the mucosal and fecal microbiota via biliary excretion and secondary metabolism. As a result, the predominance of pathogens could lead to harmful effects such as diarrhea and colitis.[190] Lastly, chemotherapeutic agents may result in severe impacts on the gut microbiota which disrupts commensal homeostasis (dysbiosis) which is a pathological alteration in the microbiota composition disturbing the physiological interaction between epithelial cells and the microbiota in the intestine among cancer patients.[185] Dysbiosis could participate in the development of mucositis as well as linked with the development of neoplastic and autoimmune disorder that may exacerbate systemic toxicity effects and lead to life-threatening systemic infections.[157,193]

To date, eight studies collected fecal samples multiple times (before, during and after chemotherapy) and assessed the impact of chemotherapy on the gut microbiome and the association between the GI microbiota and chemotherapy-induced toxicities, including infection, febrile neutropenia, diarrhea, weight gain and neurological side effects (summarized in **Table 2**).

Galloway- Peña and colleagues conducted three studies to determine the predictive value of the GI microbiome and its relationship to infection risk in patients with AML.[160-162] These studies found consistent associations between the risk of infection and the low alpha diversity of the gut microbiome before chemotherapy. The earlier study (n = 34) also found that the enriched genus *Stenotrophomonas* before chemotherapy was associated with the risk of infection in patients with AML.[160] A subsequent study conducted with AML patients (n = 55) reported high intra-patient temporal variability of alpha

diversity and increased relative abundance of *Staphylococcus, Streptococcus, Akkermansia, Subdilogranulum*, and *Pseudobutyrivibrio* were associated with an increased risk of infection.[161] The authors proposed that increased microbiome interpatient temporal variability, driven by prolonged antibiotic exposure (prophylactic antibiotic prior to chemotherapy), would be associated with infection during induction chemotherapy. The third study conducted with a larger sample size (n = 97) reported that, at baseline, higher Shannon index (alpha diversity) and higher relative abundance of *Porphyromonadaceae* were associated with an increased probability of remaining infection-free during neutropenia.[162] Patients who received antibiotics (carbapenem >72 hours) during chemotherapy had significantly lower alpha diversity and lower relative low abundance of *Clostridiales, Ruminococcaceae,* and *Porphyromonadaceae* at neutrophil recovery and were approximately four times more likely to have an infection in the 90 days following neutrophil recovery.[162] Rattanathammethee and colleagues showed a significant decrease in gut microbiota diversity (Shannon and Simpson indexes) in patients with AML who developed first febrile neutropenia after induction chemotherapy, which remained constant despite the recovery.[194] *Enterococcus* was more abundant, while *Escherichia* notably declined during the febrile neutropenia period compared with pretreatment. Additionally, the study supported that the unscrupulous use of prophylactic antibiotics may contribute to dysbiosis and the consequent increased risk of translocation and systemic infections.[194] Viaud and collogues noted in the study mentioned above in mice that the antineoplastic mechanism of cyclophosphamide may be linked to gut microbiota characteristics.[176] Daillere and colleagues confirmed these findings, demonstrating that the antineoplastic effect of cyclophosphamide was mediated by two gram-positive *Enterococcus hirae* and *Barnesiella intestinohominis* species.[195]

Some studies investigated the predictive value of microbiome composition on the development of diarrhea during chemotherapy and proposed possible mechanisms. A small human study evaluating 16 cancer patients receiving different chemotherapy regimens for various cancer diagnoses (the majority being colorectal cancer, breast cancer, melanoma, and others) and two healthy controls - has investigated the intestinal microbiota alterations, methanogenic archaea, matrix metalloproteinase, and level of NF-κB, IL-1β, and TNF associated with chemotherapy-induced diarrhea.[189] Using conventional culture techniques and quantitative real-time PCR (qRT-PCR), that study found that cancer patients

exhibited differences in fecal microbiota composition compared with healthy controls. Decreases in species members of genera *Lactobacillus*, *Bifidobacterium*, *Bacteroides* were observed while *Escherichia coli* and *Staphylococcus* species tended to be increased in cancer patients with diarrhea. Cancer patients with diarrhea also tended to have more methanogenic archaea, fecal calprotectin, matrix metalloproteinases (MMP) 3, and MMP-9. These changes might diminish GI microbial functions, initiate intestinal damage, and then result in the onset of diarrhea.[189] However, that study had a very small sample size and did not report any statistical test results.

**Table 2**: Chemotherapy-induced toxicities and GI microbiota

| Study/Subjects | Treatment | Toxicities | GI microbiota | Proposed mechanism |
|---|---|---|---|---|
| **Clinical study** | | | | |
| Stringer et al.,2013 [189]<br>• 16 cancer patients (11 patients with CRC) | Different chemotherapy regimens include capecitabine, cisplatin/5-FU, FOLFOX, 5-FU/folinic acid, COFF and paclitaxel, carboplatin, and gemcitabine. | Diarrhea | Decreases in species member of *Lactobacillus*, *Bifidobacterium*, and *Bacteroides* and increases in *Escherichia coli* and *Staphylococcus* species. | Changes in intestinal microflora, methanogenic archaea, matrix metalloproteinase, and level of NF-κB, IL-1β, and TNF might alter GI microbial functions and initiate intestinal damage and then result in the onset of diarrhea. |
| Galloway- Peña et al., 2016 [160]<br>• 34 patients with AML | Different Chemotherapy regimens were subdivided into fludarabine-containing regimens (Fludarabine/ idarubicin/ cytarabine or Fludarabine/ idarubicin/ cytarabine with G-CSF), high intensity non-fludarabine-containing regimens (Clofarabine/ idarubicin/ cytarabine), hypo-methylators (Cladribine/ low-dose cytarabine/ decitabine) | Infection | Low alpha-diversity and an increase in *Stenotrophomonas* (genus) of the baseline stool sample | The oral and gut microbiome serve as portals of infection in immunocompromised patients, especially cancer patients undergoing induction chemotherapy. |
| Galloway- Peña et al., 2017 [161]<br>• 59 patients with AML | | Infection | High intra-patient temporal variability of alpha diversity (Shannon index) and high relative abundance of *Staphylococcus, Streptococcus, Akkermansia, Subdiloqranulum*, and *Pseudobutyrivibrio* | Higher microbiome interpatient temporal variability, driven by prolonged antibiotic exposure (prophylactic antibiotic prior to chemotherapy), would be associated with infection during induction chemotherapy |
| Guthrie et al., 2017 [196]<br>• 20 health individuals | Irinotecan | Diarrhea | β-glucuronidase producing GI microbiota: *Facecabacterium prausnitzii*, *uncultured Clostridium* species, and *Bacteroides* species. | Supports for proposed mechanism based on the greater or lesser ability in intestinal microbiota to reactivate SN-38G into SN-38. The high metabotype individuals contained a significantly higher abundance of β-glucuronidases producing GI microbiota. |
| Fei et al.,<br>2019 [197]<br>• 17 stage III CRC patients | CapeOX regimen (capecitabine twice daily combined with oxaliplatin every three weeks) | Diarrhea | The community richness (ACE and Chao estimator) and community diversity (smaller Shannon index and larger Simpson index) of the gut microbiome were lower in the patients with chemotherapy-induced diarrhea.<br>*Klebsiella pneumoniae* was the most predominant species among the gut | The differentiated microorganisms, their metabolic products, and the relevant pathways make up the intestinal microecosystem that causes diarrhea. |

| Study/Subjects | Treatment | Toxicities | GI microbiota | Proposed mechanism |
|---|---|---|---|---|
| | | | microbiome in CRC patients with CID. There were 75 microorganisms with statistically significant differences between the CRC patients with and without CID at the species level including *Proteobacteria, Enterobacteriales, Gammaproteobacteria, Enterobacteriaceae, Klebsiella, Clostridiales, Clostridia, Ruminococcaceae, Bacteroidetes, Bacteroidia, Bacteroidales, Bacteroides,* and *Bacteroidaceae.* | |
| Galloway- Peña et al., 2020 [162]<br>• 97 patients with AML | Different Chemotherapy regimens were subdivided into fludarabine-containing regimens, high-intensity non-fludarabine-containing regimens, hypo-methylators, or others. | Infection before neutrophil recovery and infection post-neutrophil recovery | Low stool alpha diversity (Shannon index) and low relative abundance of *Porphyromonadaceae* and *Lachnospiraceae* for Infection before neutrophil recovery<br>A sharper decrease in Shannon index and low *Clostridiales*, *Ruminococcaceae*, and *Porphyromonadaceae* abundance at time neutrophil recover for infection postneutrophil recovery | Infections during chemotherapy result from the translocation of pathogenic bacteria across the damaged intestinal mucosa, leading to subsequent bacteremia. |
| Uzan-Yulzari et al., 2020 [198]<br>• 33 breast, ovarian, or endometrial cancer | Adjuvant chemotherapy including AC-P regimen, paclitaxel/Docetaxel plus carboplatin, paclitaxel/docetaxel alone | Weight gain | An increase in alpha diversity and higher relative abundance of member of the family *Erysipelotrichaceae* | The gut microbiome mediates adverse metabolic effects of chemotherapy, such as results in glucose intolerance, adverse lipid changes, and inflammatory changes, which lead to weight gain. |
| Rattanathammethee et al., 2020 [194]<br>• 10 patients with AML | Chemotherapy " 7+3 regimen" (seven-days of Cytarabine 100 mg/m$^2$ intravenous continuous infusion over 24 hours combined with three-day of Idarubicin 12 mg/m$^2$ bolus intravenously) | Febrile neutropenia | A significant decrease in gut microbiota diversity (Shannon and Simpson indexes) | The gut microbiota composition may be altered in patients with AML who developed a first episode of neutropenic fever during the first cycle of intensive chemotherapy. |
| Terrasse et al. 2021 [199] | Eight cycles of Adjuvant chemotherapy (almost | Neurological, gastrointestinal, rheumatological | The alpha diversity pre-chemotherapy was not significantly different between patients with or without side | Neurological side effects were associated with intestinal functional pathways involved mainly in energy |

| Study/Subjects | Treatment | Toxicities | GI microbiota | Proposed mechanism |
|---|---|---|---|---|
| • 76 patients with breast cancer | anthracycline-taxane based regimen) | side effects and weight gain | effects/toxicities. The alpha diversity post-chemotherapy varied according to BMI, diarrhea, and constipation.<br>Post-chemotherapy's beta diversity was significantly associated with neurological side effects, overt weight gain, constipation, diarrhea, or hot flashes.<br>The bacteria that were associated with neurological side effects belonged to the *Clostridiaceae* family (i.e., *C. symbosium, C. bolteae, C. spiriforme, C. aldenense, C. citroniae, C. asparagiforme* and *E. ramosum*) | production with an enrichment in the glycolysis pathways, L-histidine degradation, fatty acid biosynthesis and beta-oxidation. In contrast, neuroprotection was associated with microbial genes coding for enzymes involved in ribonucleotide de novo biosynthesis, polyamine biosynthesis and the GABA shunt. |
| **<u>Mouse model study</u>** | | | | |
| Wallace et al., 2010 [200]<br>Lin et al., 2012 [183] | Irinotecan<br>(CPT-11) | Diarrhea | β-glucuronidase producing GI microbiota: *Escherichia coli*, *Clostridium* cluster XI, and *Enterobacteriaceae* | Intestinal bacterial β-glucuronidase converts SN-38G into SN-38, inducing dose-limiting diarrhea. |
| Kurita et al., 2011 [201] | Irinotecan<br>(CPT-11) | GI toxicities | Unknown GI microbiota change | Inhibiting the absorption of CPT-11 from the intestinal lumen, reducing the CES activity, and increasing the UGT in the intestinal epithelium, which might alleviate GI toxicities |
| Fijlstra et al., 2015 [180] | Methotrexate (MTX) | Diarrhea | Increase almost of anaerobes, *Streptococci*, and *Bacteroides* species | Substantial decreases in the absolute number of bacteria and a shift in the relative species composition, which induce diarrhea. |
| Frank et al. 2015 [202] | Methotrexate (MTX) | MTX induced toxicities | Unknown GI microbiota change | TLR2 stimulation in the intestinal microbiota increases the expression and activity of the multidrug resistance pump ABCB1/MDR1 which might allow the innate immune system to protect the host from genotoxicity of high-dose chemotherapy and its toxicity. |

| Study/Subjects | Treatment | Toxicities | GI microbiota | Proposed mechanism |
|---|---|---|---|---|
| Forsgård et al., 2017 [184] | 5-FU<br>Oxaliplatin<br>Irinotecan | GI toxicities | A decrease in microbial diversity, increases in *Proteobacteria* and *Fusobacteria* | Alternations in the composition of fecal microbiota with increases in the relative abundance of bacteria known to produce lipopolysaccharide (LPS) that activate inflammatory processes and increase intestinal permeability may increase LPS leakage into circulation and further exacerbate chemotherapy-induced GI toxicities. |
| Shen et al., 2017 [203] | Oxaliplatin | Peripheral neuropathy | Unknown GI microbiota change | Gut microbiota plays a crucial role in mechanical hyperalgesia induced by oxaliplatin regarding the interrelationship between microbial LPS-TLR4 on macrophage cells. |
| Ramakrishma et al., 2019 [204] | Paclitaxel | Peripheral neuropathy (Neuropathic pain) | A decrease in the abundance of *Akkermansia muciniphila* | Paclitaxel could compromise barrier integrity, decreasing the number and function of beneficial gut bacteria resulting in systemic exposure to bacterial metabolites and products – that act via the gut-immune-brain axis – which could result in altered brain function. |

*CES: carboxylesterase; GABA: γ-aminobutyric acid; LPS: lipopolysaccharide; UGT: UDP-glucuronosyltransferase.*

Another study explored the association between the gut microbiome and chemotherapy-induced diarrhea in patients diagnosed with stage III CRC who completed eight cycles of the CapeOX regimen (capecitabine twice daily plus oxaliplatin every three weeks). This study found that patients with chemotherapy-induced diarrhea had lower gut microbial community richness (ACE and Chao estimator) and diversity (smaller Shannon index and larger Simpson index) than a control group.[197] It also reported that *Klebsiella pneumoniae* was the most predominant species among the gut microbiome in patients with chemotherapy-induced diarrhea. There were significant differences in 75 micro-organisms between the CRC patients with and without chemotherapy-induced diarrhea at the species level, including *Proteobacteria, Enterobacteriales, Gammaproteobacteria, Enterobacteriaceae, Klebsiella, Clostridiales, Clostridia, Ruminococcaceae, Bacteroidetes, Bacteroidia, Bacteroidales, Bacteroides* and *Bacteroidaceae.* [197] Irinotecan (CPT-11) is mainly used to treat advanced colorectal cancer but can also be applied to treat advanced and metastatic breast cancer. CPT-11 is converted into SN-38 metabolite by a carboxylesterase and then into its inactive form SN-38G by UDP-glucuronosyltransferases (UGP) in the liver and eliminated through biliary excretion. Once in the intestine, intestinal bacterial β-glucuronidase reconverts SN-38G into SN-38, which subsequently can induce cellular toxicity in the intestine and then increase CPT-11-induced diarrhea. In experiments with mice receiving CPT-11 treatment, Wallace et al.[200] and Lin et al.[183] demonstrated that increase in CPT-11-induced diarrhea was linked with the increased abundance of β-glucuronidase producing GI microbiota, including *Escherichia coli*, *Clostridium cluster XI*, and *Enterobacteriaceae.* Wallace and colleagues also found that co-administration of CPT-11 with a selective inhibitor of bacterial β-glucuronidase can prevent the GI production of toxic CPT-11 metabolites and the appearance of diarrhea.[200] Alleviation of irinotecan-induced diarrhea can also be achieved by the other mechanism such as inhibiting the absorption of CPT-11 from the intestinal lumen, reducing the carboxylesterase activity, and increasing the UDP-glucuronosyltransferase in the intestinal epithelium.[201] In line with the mechanism based on the reactivated ability of SN-38G into SN-38 of intestinal microbiota, Guthrie et al. collected stool samples from 20 healthy individuals and divided them into two groups, including low and high turnover microbiota metabotype.[196] The microbiome of the high metabotypes had a significantly high abundance of three β-glucuronidase producing GI microbiota including *Facecabacterium prausnitzii*, uncultured *Clostridium*

species and *Bacteroides* species when compared with the microbiome of low metabotypes.[196] An overlapping set of β-glucuronidases producing GI microbiota are carried by the high metabotypes, and age-matched, advanced adenoma, carcinoma patients were found.[196]

In mice experiments that received methotrexate treatment, a reduction of *anaerobes* and *Streptococci* and an increase in *Bacteroides* species were related to the presence of methotrexate induced-diarrhea.[180] Frank and colleagues proposed that functional TLR2 signaling in the small intestinal mucosa might be against methotrexate-induced toxicities. TLR2 stimulation in the intestinal microbiota increases the expression and activity of the multidrug resistance pump ABCB1/MDR1 which might allow the innate immune system to protect the host from genotoxicity of high-dose chemotherapy and its toxicity.[202] Evidence of the link between GI microbiome, immune response, and chemotherapy toxicity was supported by some mouse model studies.[184,203] Forsgård and colleagues demonstrated changes in fecal microbiota composition with increases in the relative abundance of phylum *Proteobacteria* known to produce lipopolysaccharide (LPS) that activate inflammatory processes and increase intestinal permeability, which may increase LPS leakage into circulation and further exacerbate chemotherapy-induced GI toxicities.[184] Moreover, the interrelationship between microbial LPS-TLR4 on macrophage cells might play a key role in mechanical hyperalgesia induced by oxaliplatin.[203]

A study of 35 patients with breast cancer and gynecological cancers assessed the relationship between the gut microbiome and weight gain in those treated with adjuvant chemotherapy, including AC-P regimen, paclitaxel/docetaxel alone or paclitaxel/docetaxel plus carboplatin.[198] Uzan-Yulzari and colleagues reported that higher alpha diversity and enriched composition of the microbiome and higher relative abundance of members of the family *Erysipelotrichaceae* in pre-chemotherapy fecal samples were associated with weight gain following chemotherapy.[198] Furthermore, in their experiments with germ-free Swiss Webster mice, they found that fecal microbiota transplantation from pre-chemotherapy samples of those patients' who gained weight post-treatment induced glucose intolerance, adverse lipid changes, and inflammatory changes.[198] These findings suggest that the gut microbiome mediates adverse metabolic effects of chemotherapy in women who received adjuvant treatment,[198] but additional research in a larger patient cohort is warranted. A study conducted by Terrasse, and colleagues reported that the alpha diversity in pre-chemotherapy fecal samples was not significantly different between early-

stage I/II breast cancer patients (n = 75) with or without side effects/toxicities.[199] However, the alpha diversity post-chemotherapy varied according to BMI, diarrhea, and constipation, and the beta diversity post-chemotherapy was significantly associated with neurological side effects (comprising paresthesia, peripheral sensory, neuropathy, memory disorders, concentration defects) and overt weight gain, constipation, diarrhea, or hot flashes. The bacteria associated with neurological side effects belongs to the family *Clostridiaceae* (i.e., *C. symbosium, C. bolteae, C. spiriforme, C. aldenense, C. citroniae, C. asparagiforme* and *E. ramosum*). Neurological side effects were associated with intestinal functional pathways involved mainly in energy production with an enrichment in the glycolysis pathways, L-histidine degradation, fatty acid biosynthesis, and beta-oxidation. In contrast, neuroprotection was associated with microbial genes coding for enzymes involved in ribonucleotide de novo biosynthesis, polyamine biosynthesis, and the GABA shunt.[199] A recent preclinical study supported the hypothesis that paclitaxel could compromise barrier integrity, decreasing the number and function of beneficial gut bacteria (e.g., *Akkermansia muciniphila*), resulting in systemic exposure to bacterial metabolites and products- which via the gut-immune-brain axis-could result in altered brain function.[204]

In general, growing evidence suggests that relationships between dysbiosis of gut microbiota and chemotherapy-induced toxicities, imply that the gut microbiome has the potential to be applied as a biomarker to predict chemotherapy outcomes and associated toxicities or identify potential microbial targets for improving treatment tolerance and efficacy.

## 7. Summary

Studies investigating the microbiome and human health are in their infancy.[205] Currently, evidence has started to appear that the GI microbiota may influence the efficacy of cancer therapy, particularly chemotherapy.[206] It was suggested that GI microbiota might modulate chemotherapeutic agents through fundamental mechanisms, including direct cytotoxicity, bacterial translocation, immune response, drug metabolism, drug efficacy and, abrogating and compromising anticancer effects,[191] through which to mediate or exacerbate the systemic toxicity effects.[192] However, the relationship between GI microbiome and chemotherapy-induced toxicity is almost uncharted research territory. Therefore, I focus my dissertation research on investigating the influence of the pre-chemotherapy GI

microbiota on chemotherapy-induced toxicity among breast cancer patients, including assessing potential drug-microbiome interaction effects with chemotherapy-induced toxicity. This is an important step towards understanding how the GI microbiome influences chemotherapy efficacy.

The proposed study has the following specific aims:

**Specific aim 1**: To describe the incidence of chemotherapy-induced toxicity and evaluate the associations between the chemotherapy-induced toxicity and sociodemographic and clinic factors.

**Specific aim 2**: To evaluate the associations between pre-chemotherapy GI microbiome and sociodemographic and clinical features among breast cancer patients.

**Specific aim 3**: To evaluate the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients.

**Specific aim 4**: To explore drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients.

In this proposed study, we describe the occurrence of chemotherapy-induced toxicities and explore its association with patients' demographics and clinical features among Vietnamese women with breast cancer, which still is limited in current literature. There is a general lack of information about the relationship between the GI microbiome and clinical features and risk factors of breast cancer patients, particularly Vietnamese population. We characterized the fecal microbiome and evaluated the associations between the fecal microbiome and sociodemographic and clinic features among those women with newly diagnosed breast cancer before receiving any chemotherapy. Our study would be expected to contribute the knowledge on the role of the GI microbiome in the development of chemotherapy-induced toxicity, which would be necessary for future development of targeted therapy to modify or restore the GI microbiome as a preventive measure improve the efficacy of cancer treatment. Understanding the role of GI microbiota in chemotherapy-induced toxicity may also lead to the identification of new options for supportive care to improve long-term cancer outcomes and quality of life for survivors.

# CHAPTER 2

SPECIFIC AIM 1

**Chemotherapy-induced toxicities and their associations with clinical and non-clinical factors**

To describe the incidence of chemotherapy-induced toxicity and evaluate the associations between the chemotherapy-induced toxicity and sociodemographic and clinic factors.

## 1. Method

### 1.1. Study Design

This study was based on a prospective follow-up of 501 newly diagnosed Vietnamese breast cancer patients who were recruited into the Vietnamese Breast Cancer Study (VBCS), a case-control study of breast cancer, supported by a P20 grant (P20 CA210300, 9/2017-8/2019) from the USA National Cancer Institute and jointly led by Drs. Shu and Tran. Details of designs and methods for the VBCS have been described.[26,207]

### 1.2. Population and participant recruitment

Patients were recruited from inpatient surgical units and chemotherapy inpatient and outpatient units of two major cancer hospitals in North Vietnam, the Vietnam National Cancer Hospital and Hanoi Oncology Hospital, from July 2017 to June 2018. Study team members identified breast cancer patients through several sources, including hospital records, operative reports, history, physical examinations, clinical mammographic assessment forms completed by radiologists, or a pathological report of breast biopsy. When a breast cancer case was identified, information on the International Classification of Diseases (ICD) code, date of the diagnosis, and tumor behavior (invasive, in situ, or borderline) was collected.

Eligible patients for this study included women with newly diagnosed (i.e., first diagnosed confirmation) breast cancer at any stage who had been biopsy or surgically confirmed and aged 18 to 79 years old, who had not received any chemotherapy and were able to provide both verbal and written informed consent. Individuals with a history of long-term antibiotic treatment or concurrent life-threatening

illnesses (e.g., stroke, heart failure) were not eligible for the study. We also excluded women who had a history of cancer before breast cancer diagnosis. Approximately 538 Vietnamese women who were newly diagnosed with breast cancer in the Vietnam National Cancer Hospitals (K3 Hospital and K1 Hospital) and Hanoi Oncology Hospital and met the eligibility criteria were approached for the study, accounting for 35% of newly diagnosed breast cancer patients seeking care at these hospitals between July 2017 to June 2018. The overall study participation rate was 93% (501 participants completed a baseline survey). Written and informed consent was obtained from all VBCS participants. Approvals for human subject research were obtained from the Vietnam National Cancer Institute and Vanderbilt University Medical Center.

### 1.3. Overall Assessments

In-person interviews were conducted at enrollment by trained interviewers using a structured questionnaire built in the Research Electronic Data Capture (REDCap) mobile application. The questionnaire included information on demographics, habitual dietary intake (via a food frequency questionnaire), other lifestyle factors, reproductive and menstrual history, medical history, and family cancer history. Bodyweight (kg), height (cm), waist and hip circumference (cm), blood pressures (mmHg), and heart rate (pulse) were measured by trained interviewers following a standard protocol. A stool sample with two fecal occult blood test (FOBT) cards, a 10 ml peripheral blood sample, and a 100 ml urine sample were collected before any systemic treatments (i.e., chemotherapy, targeted therapy, and endocrine therapy) and radiation therapy, processed within 6 hours and stored in a -80$^{o}$C freezer. Participants were followed-up via interviewer-administered surveys for a current health condition (i.e., quality of life, and self-report chemotherapy side effects), diet and lifestyle habits, cancer recurrence, and vital status after study enrollment by using telephone calls or through social networks at the first follow-up (~6-11months; response rate of 77.6% after excluding deceased patients) and the second follow-up (~12- to 18-months; response rate of 61.9% after excluding deceased patients). The median time interval between the first and second follow-up was 11 months. Additionally, clinical information was collected by reviewing patients' medical records and data were entered into the REDCap data management platform hosted at Vanderbilt University.[208]

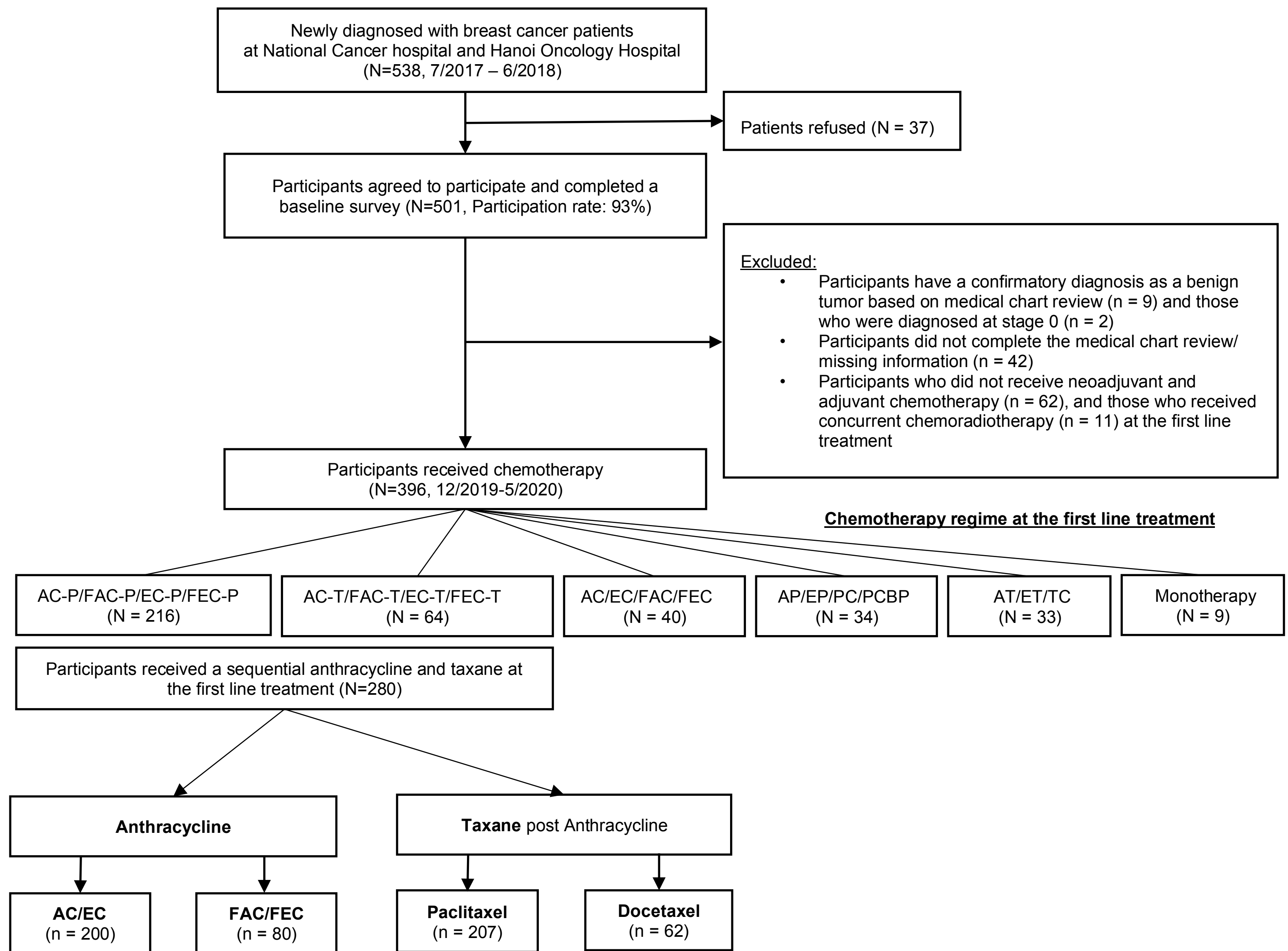



**Figure 2:** Flow diagram of study participant inclusion criteria for Aim 1

### 1.4. Population Selection

To be included in this analysis, newly diagnosed breast cancer patients must have received neoadjuvant or adjuvant chemotherapy during breast cancer treatment. Therefore, we excluded participants who were subsequently confirmed to have a benign tumor based on pathological reviews (n = 9) and those diagnosed at stage 0 (n = 2). In addition, participants with incomplete medical chart reviews or missing treatment information were excluded (n = 42). We also excluded patients who did not receive neoadjuvant and adjuvant chemotherapy (n = 62) and those who received concurrent chemoradiotherapy (n = 11) at the first-line treatment. Finally, A population of approximately 396 cases was included for Aim 1 (**Figure 2**).

### 1.5. Outcome Assessment

Breast cancer patients routinely have blood and urine tests before each cycle of chemotherapy/hospital visit to assess their health condition and chemotherapy-induced side effects to assist physicians' decision on prescribing a treatment regimen and dosing. All test results are included in the medical records. Trained study staff reviewed medical charts and abstracted information on test dates, hemoglobin (Hgb), white blood cells (WBC), absolute neutrophil count (ANC), lymphocytes, platelets (PLT), total bilirubin, serum glutamic-oxaloacetic transaminase (SGOT), serum glutamic-pyruvic transaminase (SGPT), creatinine, proteinuria, and hematuria, and directly entered it into the REDCap data management platform. Chemotherapy-induced toxicities, including neutropenia, anemia, lymphopenia, thrombocytopenia, hyperbilirubinemia, high SGOT or SGPT, evaluated creatinine, proteinuria, and hematuria were then graded according to the National Cancer Institute Common Toxicity Criteria of Adverse Events (NCI CTCAE) classification version 2.0. (**Supplementary Table S4**). The study outcomes are the highest grade of toxicities reached during the first-line chemotherapy treatment until the first day of radiotherapy for patients who received chemotherapy and sequential radiotherapy and during the first-line treatment of chemotherapy and 90 follow-up days after the treatment for patients who received only chemotherapy without radiotherapy. In terms of the sequential anthracycline and taxane regimen at the first-line treatment, anthracycline-induced and taxane-induced toxicity grades and dates that reached the highest grade of toxicities were captured (**Figure 3**). Combined hematological

toxicity refers to having any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia. Combined nephrotoxicity includes evaluated creatinine, proteinuria, or hematuria, whereas combined hepatotoxicity was identified as high levels of bilirubin, SGOT, or SGPT.

GI toxicities were identified through a combination of patient self-report side effects at the two follow-ups and the assessments recorded by treating physicians/nurses during each cycle of chemotherapy/hospital visit. Patients' self-reported side effects on non-hematological toxicities such as GI toxicities have demonstrated validity and reliability.[72,73] In terms of patient self-report side effects, participants were asked in face-to-face or telephone interviews at the two follow-ups if they had experienced nausea, vomiting, diarrhea, constipation, sore mouth, and pain or difficulty swallowing after receiving chemotherapy (**Supplementary Table S5**). These symptoms might also be recorded in medical records, but details vary. A protocol for collecting the information on these symptoms in medical charts was developed and pilot-tested (**Supplementary Table S6**). Moderate and severe symptoms that required clinical intervention were more like to be documented at each cycle of chemotherapy/hospital visit. We combined self-reported symptoms and medical chart information and graded these toxicities. The moderate and severe levels of side effects in the patient self-report form were considered as grade 2 and 3 or above NCI CTCAE classification version 2.0. Combined GI toxicity incorporated four symptoms: nausea, vomiting, diarrhea, constipation, or stomatitis.

Some chemotherapy-induced toxicities, particularly acute cardiotoxicity (ischemia) and infection, were abstracted and graded if documented in medical charts. In addition, participants were asked if they had experienced high fever, allergic reaction, itching or rash, cough, myalgia or arthralgia (muscle or joint pain), peripheral neuropathy (tingling or numbness in hands), and fatigue (feeling weak) at the two follow-ups (**Supplementary Table S5 and S6**).

In the current study, we focused on two major toxicities, including combined hematological toxicity and combined GI toxicity. Since grade 3 to 4 toxicities (according to NCI CTCAE classification) might lead to changes in a patient's management (e.g., treatment delays, dose reductions, or treatment discontinuance), toxicities were grouped as dichotomous variables (i.e., grade $\geq$3 vs. grade <3) for evaluating the associations with sociodemographic and clinic factors. Due to the rarity of severe (grade $\geq$3) hepatotoxicity and nephrotoxicity (2.0% and 0.5%) and the incidence of cardiotoxicity (no reported

ischemia case), these types of toxicities were excluded from our analysis. In addition, we did not evaluate febrile neutropenia in this study because of the lack of a reliable assessment. The incidence of documented infection and self-reported moderate/severe fever was 1.8% and 3.0%, respectively.

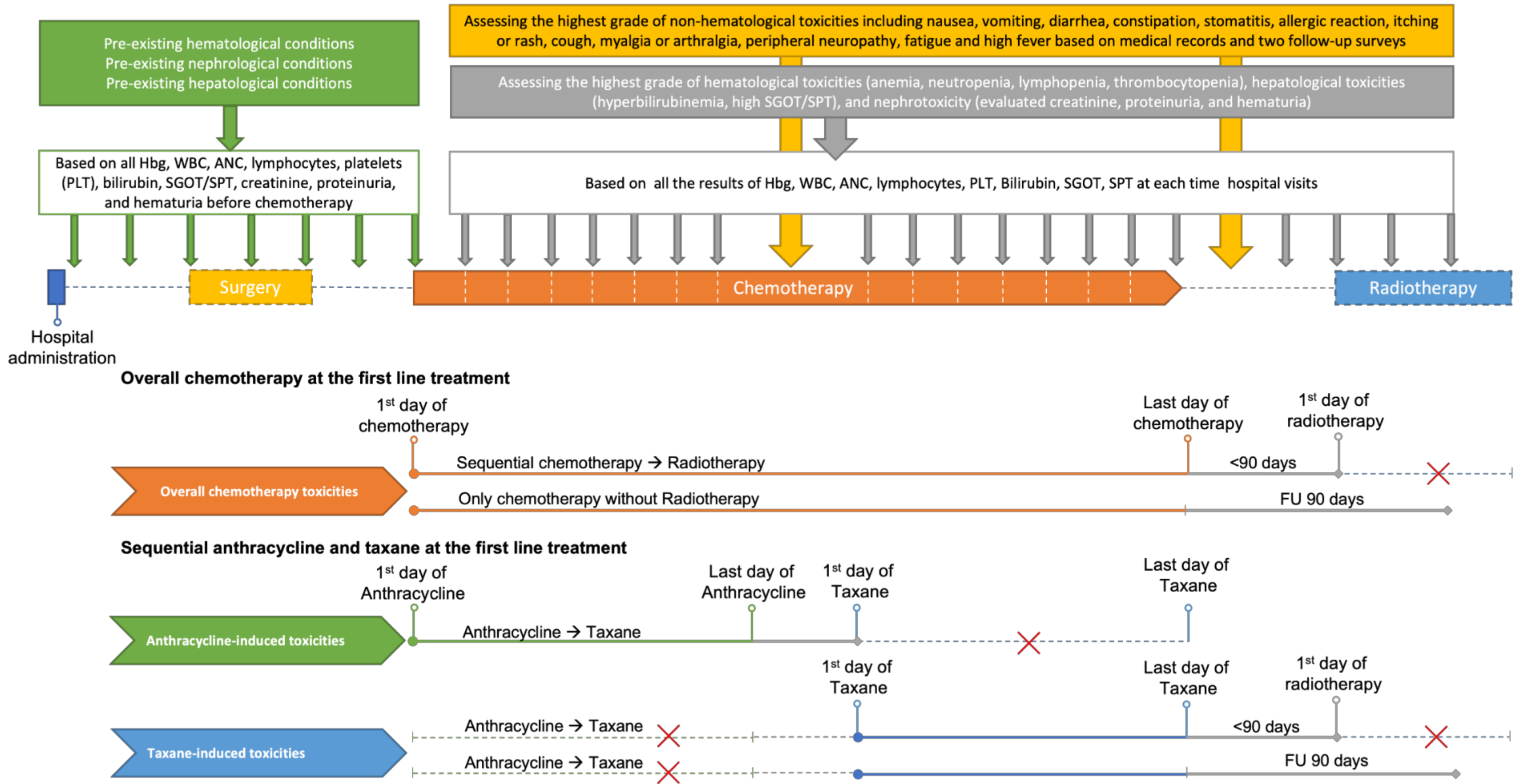



**Figure 3:** Timeline of chemotherapy-induced toxicity assessment

### 1.6. Covariate Assessment:

We collected information on important covariates, including sociodemographic characteristics and clinical features, which are summarized in **Table 3** and **Table 4**.

**Table 3:** Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 1)

| | All eligible participants (N=396) | | | All eligible participants (N=396) | |
|---|---|---|---|---|---|
| | n | % | | n | % |
| **Age (Mean±SD; years)** | 49.4±9.7 | | **Menopausal status** | | |
| **Age group** | | | Pre-menopausal | 228 | 57.6 |
| < 40 | 61 | 15.4 | Post-menopausal | 168 | 42.4 |
| 40-49 | 153 | 38.6 | **ER status** | | |
| 50-59 | 135 | 34.1 | Negative | 152 | 38.4 |
| ≥ 60 | 47 | 11.9 | Positive | 244 | 61.6 |
| **Education levels** | | | **PR status** | | |
| Primary school | 60 | 15.2 | Negative | 174 | 43.9 |
| Middle school | 168 | 42.4 | Positive | 222 | 56.1 |
| High school | 98 | 24.8 | **HER2 status** | | |
| College or higher | 70 | 17.7 | Negative | 213 | 53.8 |
| **Family annual income levels** | | | Positive | 183 | 46.2 |
| Low (T1) | 141 | 35.6 | **Ki-67 levels** | | |
| Middle (T2) | 128 | 32.3 | <20% | 132 | 33.3 |
| High (T3) | 127 | 32.1 | ≥20% | 264 | 66.7 |
| **Residence** | | | **Breast cancer subtypes** | | |
| Urban area | 150 | 37.9 | Luminal/HER2-negative | 163 | 41.2 |
| Rural area | 246 | 62.1 | Luminal/HER2-positive | 97 | 24.5 |
| **Family history of breast cancer** | | | HER2 enriched | 86 | 21.7 |
| No | 380 | 96.0 | Triple-negative/basal-like | 50 | 12.6 |
| Yes | 16 | 4.0 | **Tumor size stage** | | |
| **BMI levels** (kg/m$^2$) | | | 1 | 101 | 25.5 |
| Underweight (<18.5) | 42 | 10.6 | 2 | 230 | 58.1 |
| Normal weight (18.5-22.9) | 245 | 61.9 | 3 | 34 | 8.6 |
| Overweight (23-24.9) | 75 | 18.9 | 4 | 31 | 7.8 |
| Obese (≥25) | 34 | 8.6 | **Node stage** | | |
| **Comorbidity** | | | 0 | 211 | 53.3 |
| No | 330 | 83.3 | 1 | 104 | 26.3 |
| Yes | 66 | 16.7 | 2 | 61 | 15.4 |
| **Pre-existing hematological condition** | | | 3 | 20 | 5.1 |
| No | 280 | 70.7 | **TNM stage** | | |
| Yes | 116 | 29.3 | Stage I | 76 | 19.2 |
| **Pre-existing nephrological condition** | | | Stage II | 217 | 54.8 |
| No | 319 | 80.6 | Stage III | 84 | 21.2 |
| Yes | 77 | 19.4 | Stage IV | 19 | 4.8 |
| **Pre-existing hepatological condition** | | | **Histological subtype** | | |
| No | 330 | 83.3 | Invasive ductal carcinoma (IDC) | 303 | 76.5 |
| Yes | 66 | 16.7 | Non-IDC | 43 | 10.9 |
| | | | Unknown | 50 | 12.6 |

At the time of diagnosis, the mean age at the time of diagnosis and treatment of 396 study participants was 49.4 years. Most patients were aged between 40 to 59 years (72.7%). Approximately 62.0% of patients lived in rural areas, and 42.4% of cases had attained a high school, college, or higher education. Only 4.0% of cases reported a family history of breast cancer among first-degree relatives. In our breast cancer patients, the percentages of underweight (BMI <18.5 kg/m$^2$), overweight (BMI: 23-24.9 kg/m$^2$), and obese (BMI ≥25 kg/m$^2$) were 10.6%, 18.9%, and 8.6%, respectively. Comorbidity was defined as the existence of selected comorbid diseases such as diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease, stroke, myocardial infarction, arthritis, lupus, and other chronic diseases, as self-reported by breast cancer patients in the study enrollment. Comorbidity was reported by approximately 17% of patients in our study. In addition, the presence of pre-existing hematological, nephrological, and hepatological conditions among breast cancer patients was identified using the results of blood and urine tests within 120 days prior to chemotherapy if they had respectively at least (grade ≥ 1 according to the NCI CTCAE classification) one hematological symptoms (anemia, neutropenia, lymphopenia, and thrombocytopenia), nephrological symptoms (high creatinine, proteinuria, and hematuria) and hepatological symptoms (high bilirubin, SGOT, and SGPT). Prior to chemotherapy, 29.3%, 19.4%, and 16.7% of breast cancer patients had pre-existing hematological, nephrological, or hepatological conditions (**Table 3**).

Tumor stages T2, N0, and M0 were the most frequent among breast cancer patients. Over half (54.8%) of participants were diagnosed at stage II, while 21.2% and 4.8% were diagnosed at stage III and IV. The percentage of breast cancer patients with ER+, PR+, and HER2-positive was 61.6%, 56.1%, and 46.2%, respectively. Ki67 levels were greater than 20% at over 66.0% of participants. Breast cancer was classified into four major molecular subtypes: 1) Luminal/HER2-negative (i.e., ER and or PR positive, and HER2-negative), 2) Luminal/HER2-positive (i.e., ER and or PR positive, and either HER2-positive), 3) HER2 enriched (i.e., HER2-positive, ER-negative and PR negative), and 4) triple-negative/basal-like (i.e., ER, PR and HER2-negative). Most participants had luminal/HER2-negative subtypes (41.2%). The percentage of breast cancer patients with HER2 enriched and triple-negative/basal-like subtypes was 21.7% and 12.6%, respectively. Moreover, the majority of patients had invasive ductal carcinoma (76.5%), and 12.6% of patients had unknown histological subtypes (**Table 3**).

**Table 4:** First-line treatment for breast cancer among study participants

| | All eligible participants (N=396) | |
|---|---|---|
| | n | % |
| **Diagnosis delay [a]** | | |
| No delay (< 3 months) | 212 | 53.5 |
| Moderate delay (4-8 months) | 114 | 28.8 |
| Serious delay (≥ 9 months) | 70 | 17.7 |
| **Health system delay [b]** | | |
| No delay (< 1 month) | 378 | 95.4 |
| Delay (≥ 1 month) | 18 | 4.6 |
| **Breast cancer surgery** | | |
| No surgery | 22 | 5.6 |
| Modified radical mastectomy | 346 | 87.3 |
| Radical mastectomy | 20 | 5.1 |
| Partial/sub-total mastectomy/ lumpectomy | 8 | 2.0 |
| **Chemotherapy timing** | | |
| Neoadjuvant | 63 | 15.9 |
| Adjuvant | 333 | 84.1 |
| **Chemotherapy regimens** | | |
| AC-P/FAC-P/ EC-P/FEC-P | 216 | 54.5 |
| AC-T/FAC-T/ EC-T/FEC-T | 64 | 16.2 |
| AC/EC/ FAC/FEC | 40 | 10.1 |
| AP/EP/PC/PCBP | 34 | 8.6 |
| AT/ET/TC | 33 | 8.3 |
| Monotherapy | 9 | 2.3 |
| **Sequential anthracycline and taxane** | | |
| No | 116 | 29.3 |
| Yes | 280 | 70.7 |
| **Taxane types** | | |
| Paclitaxel | 252 | 63.6 |
| Docetaxel | 99 | 25.0 |
| No taxane | 45 | 11.4 |
| **Using 5-Fluorouracil** | | |
| No | 313 | 79.0 |
| Yes | 83 | 21.0 |
| **Dose-dense chemotherapy** | | |
| No | 349 | 88.1 |
| Yes | 47 | 11.9 |
| **Using G-CSF** | | |
| No | 290 | 73.2 |
| Yes | 106 | 26.8 |
| **Relative Dose intensity (RDI)[c]** | | |
| RDI≥85% | 325 | 82.1 |
| RDI<85% | 71 | 17.9 |
| **Chemotherapy discontinuance** | | |
| No | 363 | 91.7 |
| Yes | 33 | 8.3 |

[a] Diagnosis delay: a delay in diagnosis from the first signs/noticeable breast cancer symptoms to the diagnosis.
[b] Health system delay: a delay within the health care system from the first medical visit to the initiation of cancer treatment
[c] RDI: ratio of the dose intensity delivered to the reference standard dose intensity for a chemotherapy regimen
AC: Doxorubicin and cyclophosphamide; EC: Epirubicin and cyclophosphamide; FAC: 5-FU, doxorubicin and cyclophosphamide; FEC: 5-FU, epirubicin and cyclophosphamide; AC-P: AC followed by paclitaxel; FAC-P: FAC followed by paclitaxel; EC-P: EC followed by paclitaxel; FEC-P: FEC followed by paclitaxel; AC-T: AC followed by docetaxel; FAC-T: FAC followed by docetaxel; EC-T: EC followed by docetaxel; FEC-T: FEC followed by docetaxel; AP: Doxorubicin and paclitaxel; EP: epirubicin and paclitaxel; PC: paclitaxel and cyclophosphamide; PCBP: paclitaxel and carboplatin; AT: Doxorubicin and docetaxel; ET: epirubicin and docetaxel; TC: Docetaxel and cyclophosphamide.

In our study, diagnosis delay time (i.e., a delay in diagnosis from the first signs/noticeable breast cancer symptoms to the diagnosis) and health system delay time (i.e., typically defined as a delay of at least one month within the health care system from the first medical visit to the initiation of cancer treatment) were assessed.[26,209] The percentage of breast cancer patients who experienced moderate (4-8 months) and serious delays (≥ 9 months) in diagnosis were 31.0% and 17.5%, respectively, only 4.6% experienced health system delay (≥ 1 month). Overall, almost all patients (94.4%) had breast cancer surgery, and most (84.1%) received adjuvant chemotherapy. Sequential anthracycline and taxane was

the most common (70.7%) chemotherapy regimens, with paclitaxel being the predominant taxane used, such as AC-P/FAC-P/ EC-P/FEC-P. Paclitaxel was more commonly used than docetaxel (63.6% vs. 25.0%). In addition, approximately 21.0% of patients were also treated with 5-FU, frequently used in combination with an anthracycline. Around 12% of participants were treated with dose-dense chemotherapy, and 26% received G-CSF during chemotherapy. The proportion of regimens that met the recommended minimum goal relative dose intensity (RDI) of 85%, the ratio of the dose intensity delivered to the reference standard dose intensity for a chemotherapy regimen[210], was 82.1%. In our study, RDI <85% was considered a proxy for drug dose reduction, and 17.9% of patients had a dose reduction. Last but not least, 33 patients (8.3%) had chemotherapy discontinuance at the first-line setting (**Table 4**).

1.7. Statistical Analyses

We described the frequency and severity of combined hematological toxicity, combined GI toxicity, and their specific toxicities in the analyses overall and by chemotherapy regimens. We also described the frequency and severity of hematological toxicity and GI toxicity by selected demographic characteristics and clinical factors. The differences across subgroups were compared using chi-square tests for categorical variables. Associations for sociodemographic and clinic factors with chemotherapy-induced toxicities, including combined hematological toxicity, neutropenia, combined GI toxicity, and nausea/vomiting, were evaluated by multivariable logistic regression analysis. Adjusted odds ratios (OR) and 95% CIs were derived from logistic regression models. Potential confounders adjusted in the multivariable model 1 include age groups at diagnosis (<40, 40-49, 50-59, and 60+), income levels (tertile distribution), and residence (urban/rural). Multivariable model 2 includes all covariates included in the multivariable model 1 with additional adjustment for BMI levels (underweight, normal weight, overweight, and obese), comorbidity (yes/no), pre-existing hematological, nephrological and hepatological conditions (yes/no), TNM cancer stage (stage I, stage II, and stage III-IV), breast cancer subtype (luminal/HER2-negative, luminal/HER2-positive, HER2 enriched, and triple-negative//basal-like), sequential anthracycline and taxane (yes/no), and dose-dense chemotherapy (yes/no). Using G-CSF, relative dose intensity and treatment discontinuance were not included in the multivariable models

because they all occurred after the appearance of chemotherapy-induced toxicities. We also did not include tumor size stage, node stage, ER, PR, and HER2, and ki-67 levels into multivariable models because tumor size stage and node stage were highly correlated with TNM stage, whereas breast cancer subtype was incorporated from the status of ER, PR, and HER2, and ki-67 levels. All statistical analyses were performed at a 2-sided significance level of 0.05 using R version 3.6.3.

1.8. Statistical Power Estimation

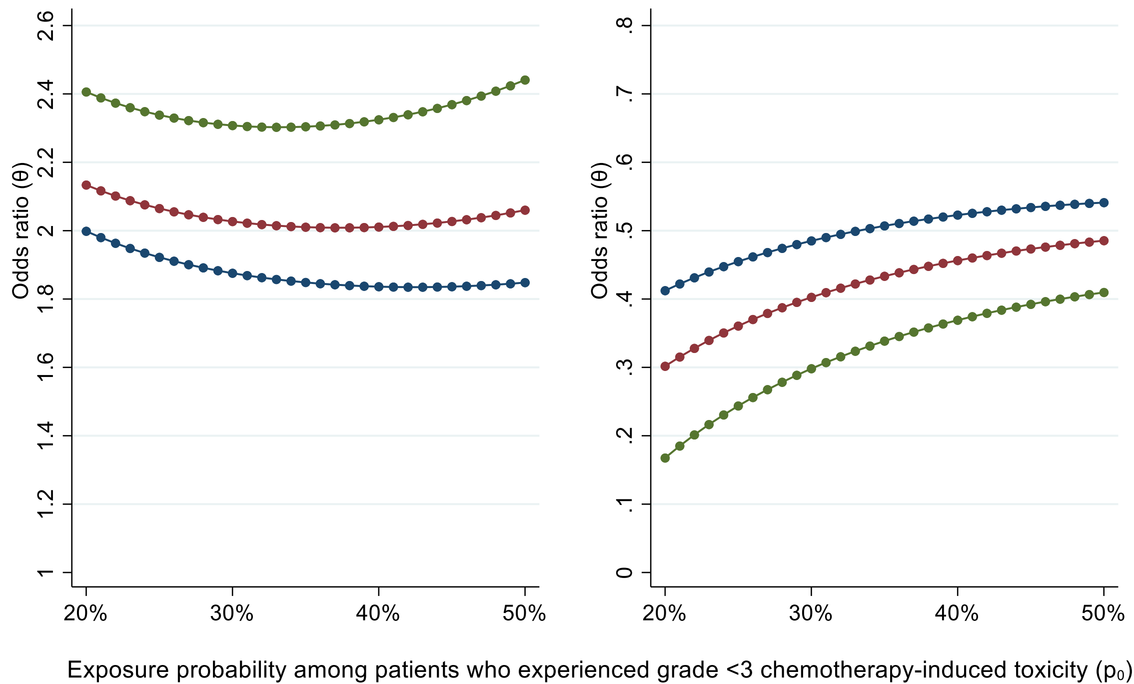



**Figure 4:** Estimated odds ratio (OR) by the prevalence of grade ≥ 3 chemotherapy-induced toxicity (Blue, red, and green curves represent15%, 25% and 50% severe toxicity at 85% power)

We assumed that 15-50% of our population have had severe (grade ≥ 3) chemotherapy-induced toxicities. We also anticipated a range of exposure prevalence (e.g., comorbidity, pre-existing hematological, nephrological, and hepatological conditions) among patients who experienced grade <3 chemotherapy-induced toxicity from 20% to 50% in a fixed sample size of 396 participants. The minimum detectable OR ranges from 1.83 to 2.30 for OR >1, and the maximum detectable OR ranges from 0.41 to 0.54 for OR<1 for 15% and 50% severe chemotherapy-induced toxicity at 85% power and a 0.05 two-sided significance level. For example, our sample size of 396 with 25% severe chemotherapy-induced toxicity was able to detectable an OR greater than 2.13 or less than 0.36 if the prevalence of exposure of interest among patients who experienced grade <3 chemotherapy-induced toxicity is 30%. (**Figure 4**).

## 2. Results

**Table 5:** Highest grade of chemotherapy-induced toxicities among study participants (total N=396)

| | **Non (Grade 0)** | **Mild (Grade 1)** | **Moderate (Grade 2)** | **Severe (Grade ≥ 3)** |
|---|---|---|---|---|
| | *n* (%) | *n* (%) | *n* (%) | *n* (%) |
| **Hematological toxicity** | | | | |
| Anemia | 179 (45.2) | 171 (43.2) | 40 (10.1) | 6 (1.5) |
| Neutropenia | 125 (31.6) | 67 (16.9) | 87 (22.0) | 117 (29.5) |
| Lymphopenia | 83 (21.0) | 180 (45.5) | 117 (29.5) | 16 (4.0) |
| Thrombocytopenia | 316 (79.8) | 38 (9.6) | 5 (1.3) | 37 (9.3) |
| Combined toxicity [a] | 32 (8.1) | 87 (22.0) | 124 (31.3) | 153 (38.6) |
| **GI toxicity** | | | | |
| Nausea/vomiting | 169 (42.7) | 100 (25.3) | 87 (22.0) | 40 (10.1) |
| Diarrhea | 316 (79.8) | 47 (11.9) | 27 (6.8) | 6 (1.5) |
| Stomatitis | 299 (75.5) | 50 (12.6) | 41 (10.4) | 6 (1.5) |
| Constipation | 292 (73.7) | 67 (7.8) | 31 (7.8) | 6 (1.5) |
| Combined toxicity [b] | 139 (35.1) | 90 (22.7) | 116 (29.3) | 51 (12.9) |

[a] Combined hematological toxicity refers to have any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[b] Combined GI toxicity refers to have any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.

The highest grade of hematological and GI toxicities was summarized in **Table 5**. Combined hematological and GI toxicities were prevalent during the first-line chemotherapy treatment, with 91.9% and 64.9% of breast cancer patients experiencing at least one specific toxicity. The incidence rates of grade 3 and above (grade ≥ 3) combined hematological and GI toxicity were 38.6% and 12.9%, respectively. Neutropenia was the most common chemotherapy-induced toxicity among hematological toxicities, with 22.0% and 29.5% of patients experiencing grade 2 and grade ≥ 3, respectively. In addition, the incidence rates of grade 2 and grade ≥ 3 nausea/vomiting were 22.0% and 10.1%, respectively, making this toxicity the most frequent GI toxicities among breast cancer patients.

Over 30% of patients recorded grade ≥ 3 combined hematological toxicity, and over 10% of patients recorded grade ≥ 3 combined GI toxicity across almost chemotherapy regimens at the first-line treatment (**Table 6**). Our study population received chemotherapy with sequential anthracycline and taxane more frequently than other regimens (70.3% vs. 29.3%). Significantly higher incidences of grade ≥ 3 combined hematological toxicity (41.4%) and grade ≥ 3 combined GI toxicity (12.9%) were

experienced by patients receiving sequential anthracycline and taxane (**Table 7**). Grade ≥ 3 neutropenia was also more frequently recorded among patients receiving sequential anthracycline and taxane than other regimens (33.2% vs. 20.7%; *P*=0.02).

No significant differences were observed among participants receiving sequential anthracycline and taxane for all hematological toxicities induced by anthracycline or by taxane post anthracycline. However, the incidence and severity of GI toxicities decreased after breast cancer patients began taxane treatment (**Table 7**). In comparison with grade ≥ 3 taxane-induced GI toxicities, patients experienced more grade ≥ 3 anthracycline-induced nausea/vomiting (7.1% vs. 2.5%). The incidences of anthracycline-induced and taxane-induced toxicities by using 5-FU (AC/EC vs. FAC/FEC), anthracycline-based types (AC/FAC vs. EC/FEC), and taxane types (paclitaxel vs. docetaxel) are shown in **Table 8**. Subgroups showed no significant differences for all hematological and GI toxicities.

**Table 6**: Incidence of grade ≥ 3 chemotherapy-induced toxicities by chemotherapy regimens

| | Grade | AC-P/FAC-P/ EC-P/FEC-P N=216 n (%) | AC-T/FAC-T/ EC-T/FEC-T N=64 n (%) | AC/EC/ FAC/FEC N=40 n (%) | AP/EP/ PC/PCBP N=34 n (%) | AT/ET/TC N=33 n (%) | Monotherapy N=9 n (%) |
|---|---|---|---|---|---|---|---|
| **Hematological toxicity** | | | | | | | |
| Anemia | ≥ 3 | 1 (0.5) | 2 (3.1) | 1 (2.5) | 1 (2.9) | 0 (0.0) | 1 (11.1) |
| Neutropenia | ≥ 3 | 71 (32.9) | 22 (34.4) | 10 (25.0) | 5 (14.7) | 7 (21.2) | 2 (22.2) |
| Lymphopenia | ≥ 3 | 11 (5.1) | 2 (3.1) | 1 (2.5) | 0 (0.0) | 1 (3.0) | 1 (11.1) |
| Thrombocytopenia | ≥ 3 | 24 (11.1) | 4 (6.3) | 2 (5.0) | 2 (5.9) | 3 (9.1) | 2 (22.2) |
| Combined toxicity [a] | ≥ 3 | 92 (42.6) | 24 (37.5) | 14 (35.0) | 8 (25.3) | 10 (30.3) | 5 (55.6) |
| **GI toxicity** | | | | | | | |
| Nausea/vomiting | ≥ 3 | 22 (10.2) | 7 (10.9) | 3 (7.5) | 2 (5.9) | 6 (18.2) | 0 (0.0) |
| Diarrhea | ≥ 3 | 5 (2.3) | 0 (0.0) | 0 (0.0) | 1 (2.9) | 0 (0.0) | 0 (0.0) |
| Stomatitis | ≥ 3 | 2 (0.9) | 1 (1.6) | 2 (5.0) | 1 (2.9) | 0 (0.0) | 0 (0.0) |
| Constipation | ≥ 3 | 4 (1.9) | 0 (0.0) | 0 (0.0) | 1 (2.9) | 1 (3.0) | 0(0.0) |
| Combined toxicity [b] | ≥ 3 | 28 (13.0) | 8 (12.5) | 4 (10.0) | 4 (11.8) | 7 (21.2) | 0 (0.0) |

[a] Combined hematological toxicity refers to have any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[b] Combined GI toxicity refers to have any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.

**Table 7**: Incidence of grade ≥ 3 chemotherapy-induced toxicities by sequential anthracycline and taxane

| | | Sequential anthracycline and taxane | | | | | |
|---|---|---|---|---|---|---|---|
| | | **No** | **Yes** | | **Anthracycline-** induced | **Taxane-**induced post Anthracycline | |
| | | N=116 | N=280 | *P*[1] | N=280 | N=280 | *P*[2] |
| | *Grade* | *n* (%) | *n* (%) | | *n* (%) | *n* (%) | |
| **Hematological toxicity** | | | | | | | |
| Anemia | ≥ 3 | 3 (2.6) | 3 (1.1) | 0.36* | 1 (0.4) | 2 (0.7) | 0.68 |
| Neutropenia | ≥ 3 | 24 (20.7) | 93 (33.2) | 0.02 | 61 (21.8) | 49 (17.5) | 0.12 |
| Lymphopenia | ≥ 3 | 3 (2.6) | 13 (4.6) | 0.42* | 10 (3.6) | 8 (2.9) | 0.45 |
| Thrombocytopenia | ≥ 3 | 9 (7.8) | 28 (10.0) | 0.57 | 14 (5.0) | 17 (6.1) | 0.45 |
| Combined toxicity [a] | ≥ 3 | 37 (31.9) | 116 (41.4) | 0.009 | 79 (28.2) | 65 (23.2) | 0.12 |
| **GI toxicity** | | | | | | | |
| Nausea/vomiting | ≥ 3 | 11 (9.5) | 29 (10.4) | 0.79 | 20 (7.1) | 7 (2.5) | 0.04 |
| Diarrhea | ≥ 3 | 1 (0.9) | 5 (1.8) | 0.67* | 2 (0.7) | 3 (1.1) | 0.40 |
| Stomatitis | ≥ 3 | 3 (2.6) | 3 (1.1) | 0.36* | 3 (1.1) | 0 (0.0) | 0.03 |
| Constipation | ≥ 3 | 2 (1.7) | 4 (1.4) | 1.00* | 4 (1.5) | 0 (0.0) | 0.42 |
| Combined toxicity [b] | ≥ 3 | 15 (12.9) | 36 (12.9) | 0.98 | 25 (7.1) | 10 (3.6) | 0.04 |

[1] p-value for chi-square tests; [a] p-value for fisher's exact test
[2] p-value for the equality of proportion
[a] Combined hematological toxicity refers to have any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[b] Combined GI toxicity refers to have any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.

**Table 8:** Incidence of ≥ 3 grade anthracycline-induced and taxane-induced toxicities among breast cancer patients who received sequential anthracycline and taxane

| | | **Anthracycline** | | | | | | **Taxane** post Anthracycline | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Using 5-FU | | | Anthracycline-based types | | | Taxane types | | |
| | | AC/EC | FAC/FEC | | AC/FAC | EC/FEC | | Paclitaxel | Docetaxel | |
| | | N=206 | N=74 | *P*[1] | N=203 | N=77 | *P*[1] | N=216 | N=64 | *P*[1] |
| | *Grade* | *n* (%) | *n* (%) | | *n* (%) | *n* (%) | | *n* (%) | *n* (%) | |
| **Hematological toxicity** | | | | | | | | | | |
| Anemia | ≥ 3 | 1 (0.5) | 0 (0.0) | 1.00* | 1 (0.5) | 0 (0.0) | 1.00* | 0 (0.0) | 2 (3.1) | 0.05* |
| Neutropenia | ≥ 3 | 48 (23.3) | 13 (17.6) | 0.33 | 48 (23.7) | 13 (16.9) | 0.22 | 38 (17.6) | 11 (17.2) | 0.94 |
| Lymphopenia | ≥ 3 | 10 (4.9) | 0 (0.0) | 0.07* | 10 (4.9) | 0 (0.0) | 0.07* | 6 (2.8) | 2 (3.1) | 1.00* |
| Thrombocytopenia | ≥ 3 | 11 (5.3) | 3 (4.1) | 1.00* | 9 (4.4) | 5 (6.5) | 0.54 | 13 (6.0) | 4 (6.3) | 1.00* |
| Combined toxicity [a] | ≥ 3 | 63 (30.6) | 16 (21.6) | 0.17 | 61 (30.1) | 18 (23.4) | 0.27 | 50 (23.2) | 15 (23.4) | 0.96 |
| **GI toxicity** | | | | | | | | | | |
| Nausea/vomiting | ≥ 3 | 14 (6.8) | 6 (8.1) | 0.71 | 14 (6.9) | 6 (7.8) | 0.80 | 6 (2.8) | 1 (1.6) | 1.00* |
| Diarrhea | ≥ 3 | 1 (0.5) | 1 (1.4) | 0.46* | 1 (0.5) | 1 (1.3) | 0.48* | 3 (1.4) | 0 (0.0) | 1.00* |
| Stomatitis | ≥ 3 | 2 (1.0) | 1 (1.4) | 1.00* | 2 (1.0) | 1 (1.3) | 1.00* | 0 (0.0) | 0 (0.0) | - |
| Constipation | ≥ 3 | 3 (1.5) | 1 (1.4) | 1.00* | 3 (1.5) | 1 (1.3) | 1.00* | 0 (0.0) | 0 (0.0) | - |
| Combined toxicity [b] | ≥ 3 | 17 (8.3) | 8 (10.8) | 0.51 | 17 (8.4) | 8 (10.4) | 0.60 | 9 (4.2) | 1 (1.6) | 0.32* |

[1] p-value for chi-square tests; [*] p-value for fisher's exact test
[a] Combined hematological toxicity refers to have any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[b] Combined GI toxicity refers to have any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.

**Tables 9-11** show the incidence of combined hematological toxicity (grade ≥ 3 vs. grade < 3) and combined GI toxicity (grade ≥ 3 vs. grade < 3) by selected demographic characteristics and clinical features. Besides administering sequential anthracycline and taxane, a significantly higher incidence of grade ≥ 3 combined hematological toxicity was experienced by patients who used G-CSF, drug dose reduction (RDI < 85%), and treatment discontinuance. In addition, grade ≥ 3 combined hematological toxicity was significantly more frequent among patients with triple-negative/basal-like and luminal/HER2-positive subtypes than other remaining subtypes. Patients living in rural areas and with comorbidity had lower grade ≥ 3 combined hematological toxicity than patients living in urban areas and had no comorbidity. There were no significant differences for combined GI toxicity by demographic characteristics and clinical factors, except for TNM cancer stage and chemotherapy discontinuance. A significantly higher incidence of grade ≥ 3 combined GI toxicity was experienced in breast cancer patients with early-stage (stage I) and patients who have chemotherapy discontinuance.

**Table 9:** Highest grade of chemotherapy-induced toxicity by selected demographic characteristics among study participants (total N=396)

| | N | Combined GI toxicity | | | Combined hematological toxicity | | |
|---|---|---|---|---|---|---|---|
| | | Grade <3 | Grade ≥3 | *P* [1] | Grade <3 | Grade ≥3 | *P* [1] |
| **Age group** | | | | | | | |
| < 40 | 61 | 51 (83.6) | 10 (16.4) | 0.41 | 40 (65.6) | 21 (34.4) | 0.13 |
| 40-49 | 153 | 131 (85.6) | 22 (14.4) | | 84 (54.9) | 69 (45.1) | |
| 50-59 | 135 | 119 (88.1) | 16 (11.9) | | 85 (63.0) | 50 (37.0) | |
| ≥ 60 | 47 | 44 (93.6) | 3 (6.4) | | 34 (72.3) | 13 (27.7) | |
| **Education** | | | | | | | |
| Primary school | 60 | 53 (88.3) | 7 (11.7) | 0.62 | 45 (75.0) | 15 (25.0) | 0.13 |
| Middle school | 168 | 150 (89.3) | 18 (10.7) | | 98 (58.3) | 70 (41.7) | |
| High school | 98 | 83 (84.7) | 15 (15.3) | | 59 (60.2) | 39 (39.8) | |
| College or higher | 70 | 59 (84.3) | 11 (15.7) | | 41 (58.6) | 29 (41.4) | |
| **Income** | | | | | | | |
| Low (T1) | 141 | 122 (86.5) | 19 (13.5) | 0.53 | 91 (64.5) | 50 (35.5) | 0.62 |
| Middle (T2) | 128 | 109 (85.2) | 19 (14.8) | | 77 (60.2) | 51 (39.8) | |
| High (T3) | 127 | 114 (89.8) | 13 (10.2) | | 75 (59.1) | 52 (40.9) | |
| **Residence** | | | | | | | |
| Urban area | 150 | 131 (87.3) | 19 (12.7) | 0.92 | 80 (53.3) | 70 (46.7) | 0.01 |
| Rural area | 246 | 214 (87.0) | 32 (13.0) | | 163 (66.3) | 83 (33.7) | |
| **Menopausal status** | | | | | | | |
| Pre-menopausal | 228 | 195 (85.5) | 33 (14.5) | 0.27 | 137 (60.1) | 91 (39.9) | 0.54 |
| Post-menopausal | 168 | 150 (89.3) | 18 (10.7) | | 106 (63.1) | 62 (36.9) | |
| **Family history of breast cancer** | | | | | | | |
| No | 380 | 331 (87.1) | 49 (12.9) | 0.96 | 234 (61.6) | 146 (38.4) | 0.67 |
| Yes | 16 | 14 (87.5) | 2 (12.5) | | 9 (56.3) | 7 (43.8) | |
| **BMI levels (kg/m$^2$)** | | | | | | | |
| Underweight (<18.5) | 42 | 38 (90.5) | 4 (9.5) | 0.43 | 26 (61.9) | 16 (38.1) | 0.07 |
| Normal weight (18.5-22.9) | 245 | 208 (84.9) | 37 (15.1) | | 153 (62.4) | 92 (37.6) | |
| Overweight (23-24.9) | 75 | 67 (89.3) | 8 (10.7) | | 38 (50.7) | 37 (49.3) | |
| Obese (≥25) | 34 | 32 (94.1) | 2 (5.9) | | 26 (76.5) | 8 (23.5) | |

[1] p-value for chi-square tests

**Table 10**: Highest grade of chemotherapy-induced toxicity by selected disease characteristics among study participants (total N=396)

| | N | Combined GI toxicity | | | Combined hematological toxicity | | |
|---|---|---|---|---|---|---|---|
| | | Grade <3 | Grade ≥3 | *P*[1] | Grade <3 | Grade ≥3 | *P*[1] |
| **Comorbidity** [a] | | | | | | | |
| No | 330 | 288 (87.3) | 42 (12.7) | 0.84 | 195 (59.1) | 135 (40.9) | 0.04 |
| Yes | 66 | 57 (86.4) | 9 (13.6) | | 48 (72.7) | 18 (27.3) | |
| **Pre-existing hematological condition** [b] | | | | | | | |
| No | 280 | 240 (85.7) | 40 (14.3) | 0.19 | 173 (61.8) | 107 (38.2) | 0.79 |
| Yes | 116 | 105 (90.5) | 11 (9.5) | | 70 (60.3) | 46 (39.7) | |
| **Pre-existing nephrological condition** [c] | | | | | | | |
| No | 319 | 274 (85.9) | 45 (14.1) | 0.14 | 205 (64.3) | 114 (35.7) | 0.06 |
| Yes | 77 | 71 (92.2) | 6 (7.8) | | 38 (49.4) | 39 (50.6) | |
| **Pre-existing hepatological condition** [d] | | | | | | | |
| No | 330 | 290 (87.9) | 40 (12.1) | 0.31 | 201 (60.9) | 129 (39.1) | 0.67 |
| Yes | 66 | 55 (83.3) | 11 (16.7) | | 42 (63.6) | 24 (36.4) | |
| **ER status** | | | | | | | |
| Negative | 152 | 128 (84.2) | 24 (15.8) | 0.17 | 88 (57.9) | 64 (42.1) | 0.26 |
| Positive | 244 | 217 (88.9) | 27 (11.1) | | 155 (63.5) | 89 (36.5) | |
| **PR status** | | | | | | | |
| Negative | 174 | 149 (85.6) | 25 (14.4) | 0.43 | 110 (63.2) | 64 (36.8) | 0.50 |
| Positive | 222 | 196 (88.3) | 26 (11.7) | | 133 (59.9) | 89 (40.1) | |
| **HER2 status** | | | | | | | |
| Negative | 213 | 184 (86.4) | 29 (13.6) | 0.64 | 134 (62.9) | 79 (37.1) | 0.50 |
| Positive | 183 | 161 (88.0) | 22 (12.0) | | 109 (59.6) | 74 (40.4) | |
| **Ki-67 (%)** | | | | | | | |
| <20% | 132 | 116 (87.9) | 16 (12.1) | 0.75 | 78 (59.1) | 54 (40.9) | 0.51 |
| ≥20% | 264 | 229 (86.7) | 35 (13.3) | | 165 (62.5) | 99 (37.5) | |
| **Tumor size stage** | | | | | | | |
| 1 | 101 | 83 (82.2) | 18 (17.8) | 0.21 | 57 (56.4) | 44 (43.6) | 0.36 |
| 2 | 230 | 204 (88.7) | 26 (11.3) | | 142 (61.7) | 88 (38.3) | |
| 3 | 34 | 32 (94.1) | 2 (5.9) | | 21 (61.8) | 13 (38.2) | |
| 4 | 31 | 26 (83.9) | 5 (16.1) | | 23 (74.2) | 8 (25.8) | |
| **Node stage** | | | | | | | |
| 0 | 211 | 178 (84.4) | 33 (15.6) | 0.05 | 128 (60.7) | 83 (39.3) | 0.19 |
| 1 | 104 | 98 (94.2) | 6 (5.8) | | 68 (65.4) | 36 (34.6) | |
| 2 | 61 | 51 (83.6) | 10 (16.4) | | 39 (63.9) | 22 (36.1) | |
| 3 | 20 | 18 (90.0) | 2 (10.0) | | 8 (40.0) | 12 (60.0) | |
| **TNM stage** | | | | | | | |
| Stage I | 76 | 59 (77.6) | 17 (22.4) | 0.01 | 46 (60.5) | 30 (39.5) | 0.98 |
| Stage II | 217 | 197 (90.8) | 20 (9.2) | | 134 (61.8) | 83 (38.2) | |
| Stage III-IV | 103 | 89 (86.4) | 14 (13.6) | | 63 (61.2) | 40 (38.8) | |
| **Breast cancer subtype** | | | | | | | |
| Luminal/HER2-negative | 163 | 146 (89.6) | 17 (10.4) | 0.09 | 111 (68.1) | 52 (31.9) | 0.01 |
| Luminal/HER2-positive | 97 | 85 (87.6) | 12 (12.4) | | 51 (52.6) | 46 (47.4) | |
| HER2 enriched | 86 | 76 (88.4) | 10 (11.6) | | 58 (67.4) | 28 (32.6) | |
| Triple-negative | 50 | 38 (76.0) | 12 (24.0) | | 23 (46.0) | 27 (54.0) | |
| **Histological subtype** | | | | | | | |
| IDC | 303 | 263 (86.8) | 40 (13.2) | 0.94 | 188 (62.0) | 115 (38.0) | 0.73 |
| Non-IDC | 43 | 38 (88.4) | 5 (11.6) | | 24 (55.8) | 19 (44.2) | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Unknown | 50 | 44 (88.0) | 6 (12.0) | | 31 (62.0) | 19 (38.0) | |

[1] p-value for chi-square tests
[a] Having diagnosis of specific comorbidities, including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Having at least one of hematological symptoms (grade≥1), including anemia, neutropenia, lymphopenia, and thrombocytopenia within 120 days prior to chemotherapy.
[c] Having at least one of nephrological symptoms (grade≥1), including high creatinine, proteinuria, and hematuria within 120 days prior to chemotherapy.
[d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT within 120 days prior to chemotherapy.

**Table 11**: Highest grade of chemotherapy-induced toxicity by characteristics related to treatment

| | N | Combined GI toxicity | | | Combined hematological toxicity | | |
|---|---|---|---|---|---|---|---|
| | | Grade <3 | Grade ≥3 | $P^1$ | Grade <3 | Grade ≥3 | $P^1$ |
| **Diagnosis delay [a]** | | | | | | | |
| No delay | 212 | 186 (87.7) | 26 (12.3) | 0.90 | 120 (56.6) | 92 (43.4) | 0.10 |
| Moderate delay | 114 | 99 (86.8) | 15 (13.2) | | 78 (68.4) | 36 (31.6) | |
| Serious delay | 70 | 60 (85.7) | 10 (14.3) | | 45 (64.3) | 25 (35.7) | |
| **Health system delay [b]** | | | | | | | |
| No delay | 378 | 328 (86.8) | 50 (13.2) | 0.49 | 233 (61.6) | 145 (38.4) | 0.60 |
| Delay | 18 | 17 (94.4) | 1 (5.6) | | 10 (55.6) | 8 (44.4) | |
| **Breast cancer surgery** | | | | | | | |
| No | 22 | 19 (86.4) | 3 (13.6) | 0.91 | 14 (63.6) | 8 (36.4) | 0.82 |
| Yes | 374 | 326 (87.2) | 48 (12.8) | | 229 (61.2) | 145 (38.8) | |
| **Chemotherapy timing** | | | | | | | |
| Neoadjuvant | 63 | 54 (85.7) | 9 (14.3) | 0.72 | 44 (69.8) | 19 (30.2) | 0.13 |
| Adjuvant | 333 | 291 (87.4) | 42 (12.6) | | 199 (59.8) | 134 (40.2) | |
| **Sequential anthracycline and taxane** | | | | | | | |
| No | 116 | 101 (87.1) | 15 (12.9) | 0.98 | 79 (68.1) | 37 (31.9) | 0.01 |
| Yes | 280 | 244 (87.1) | 36 (12.9) | | 164 (58.6) | 116 (41.4) | |
| **Taxane types** | | | | | | | |
| Paclitaxel | 252 | 220 (87.3) | 32 (12.7) | 0.58 | 153 (60.7) | 99 (39.3) | 0.69 |
| Docetaxel | 99 | 84 (84.8) | 15 (15.2) | | 64 (64.6) | 35 (35.4) | |
| No taxane | 45 | 41 (91.1) | 4 (8.9) | | 26 (57.8) | 19 (42.2) | |
| **Using 5-Fluorouracil** | | | | | | | |
| No | 313 | 274 (87.5) | 39 (12.5) | 0.63 | 191 (61.0) | 122 (39.0) | 0.79 |
| Yes | 83 | 71 (85.5) | 12 (14.5) | | 52 (62.7) | 31 (37.3) | |
| **Dose-dense chemotherapy** | | | | | | | |
| No | 349 | 301 (86.2) | 48 (13.8) | 0.16 | 220 (63.0) | 129 (37.0) | 0.06 |
| Yes | 47 | 44 (93.6) | 3 (6.4) | | 23 (48.9) | 24 (51.1) | |
| **Chemotherapy duration** | | | | | | | |
| T1 | 134 | 118 (88.1) | 16 (11.9) | 0.08 | 87 (64.9) | 47 (35.1) | 0.38 |
| T2 | 142 | 129 (90.8) | 13 (9.2) | | 81 (57.0) | 61 (43.0) | |
| T3 | 120 | 98 (81.7) | 22 (18.3) | | 75 (62.5) | 45 (37.5) | |
| **Using G-CSF** | | | | | | | |
| No | 290 | 253 (87.2) | 37 (12.8) | 0.91 | 188 (64.8) | 102 (35.2) | 0.02 |
| Yes | 106 | 92 (86.8) | 14 (13.2) | | 55 (51.9) | 51 (48.1) | |
| **Relative dose intensity (RDI) in chemotherapy [c]** | | | | | | | |
| RDI>=85% | 325 | 280 (86.2) | 45 (13.8) | 0.22 | 209 (64.3) | 116 (35.7) | 0.01 |
| RDI<85% | 71 | 65 (91.5) | 6 (8.5) | | 34 (47.9) | 37 (52.1) | |
| **Chemotherapy discontinuance** | | | | | | | |
| No | 363 | 320 (88.2) | 43 (11.8) | 0.04 | 242 (66.7) | 121 (33.3) | 0.001 |
| Yes | 33 | 25 (75.8) | 8 (24.2) | | 1 (3.0) | 32 (97.0) | |

[1] p-value for chi-square tests
[a] Diagnosis delay: a delay in diagnosis from the first signs/noticeable breast cancer symptoms to the diagnosis.
[b] Health system delay: a delay within the health care system from the first medical visit to the initiation of cancer treatment
[c] RDI: ratio of the dose intensity delivered to the reference standard dose intensity for a chemotherapy regimen

Multivariable analyses showed breast cancer patients with pre-existing nephrological condition had a significantly higher risk of grade ≥3 combined hematological toxicity (OR= 2.30; 95% CI: 1.32-4.01), while patients with comorbidity had a significantly lower risk of ≥3 combined hematological toxicity (OR = 0.49; 95% CI: 0.24-0.97). In breast cancer subtypes, patients with luminal/HER2-positive and triple-negative/basal-like subtypes were more likely to experience grade ≥ 3 combined hematological toxicity than patients with luminal/HER2-negative subtype. The adjusted ORs and 95% CIs for grade ≥3 combined hematological toxicity were 1.78 (1.02-3.10) for luminal/HER2-positive and 3.15 (1.56-6.34) for Triple-negative/basal-like. Despite patients who received chemotherapy with sequential anthracycline and taxane or received dose-dense chemotherapy being more likely to experience higher grade ≥ 3 combined hematological toxicity, we found no significant association in multivariate analyses (**Table 12**). Conversely, sequential anthracycline and taxane was not associated with an increased risk of grade ≥3 neutropenia (OR= 1.74; 95%CI: 0.96-3.17), but patients who received dose-dense chemotherapy were more likely to experience grade ≥3 neutropenia (OR= 2.64; 95% CI: 1.32-5.25). A significantly higher risk of grade ≥3 neutropenia was found for the pre-existing nephrological condition (OR = 1.86; 95%: 1.04-3.30), but not for comorbidity (OR= 0.56; 95% CI: 0.26-1.17). In addition, only breast cancer patients having triple-negative/basal-like subtype were more likely to experience grade ≥3 neutropenia (OR = 2.53; 95%CI: 1.22-5.24) compared with patients with luminal/HER2-negative subtype (**Table 13**). Last but not least, breast cancer patients who were living in rural areas were less likely to experience grade ≥3 combined hematological toxicity (OR = 0.48; 95% CI: 0.30-0.77) and grade ≥3 combined neutropenia (OR = 0.53; 95% CI: 0.32-0.87) compared with those living in urban areas (**Tables 12 and 13**).

Multivariable analyses in **Tables 14 and 15** showed that patients diagnosed at stage II and stage III-IV had a significantly lower risk of grade ≥3 combined GI toxicity than patients with stage I. Adjusted ORs and 95%CIs for grade ≥3 combined GI toxicity were 0.26 (0.12-0.59) and 0.47 (0.20-1.10) for cancer stage II and cancer stage III-IV. A similar grade ≥3 nausea/vomiting association pattern was observed for TNM cancer stages. Adjusted ORs and 95%CIs for grade ≥3 nausea/vomiting were 0.17 (0.07-0.41) and 0.29 (0.11-0.76) for cancer stage II and cancer stage III-IV. Patients with triple-

negative/basal-like subtype were more likely to experience grade ≥3 combined GI toxicity (OR = 3.60; 95% CI: 1.45-8.95) than patients with luminal/HER2-negative subtype. This association was not consistent for grade ≥3 nausea/vomiting.

**Table 12:** Association of demographic characteristics and clinical factors with combined hematological toxicity

| | **Combined hematological toxicity** (grade ≥3 vs. grade <3) | | |
|---|---|---|---|
| | *No*. of grade ≥3/ grade <3 | Model 1 Adjusted OR (95%CI)[1] | Model 2 Adjusted OR (95%CI)[2] |
| **Age group** | | | |
| < 40 | 21/ 40 | 1 | 1 |
| 40-49 | 69/ 84 | 1.64 (0.88-3.07) | 2.04 (1.04-4.00) |
| 50-59 | 50/ 85 | 1.25 (0.66-2.39) | 1.95 (0.95-4.02) |
| ≥ 60 | 13/ 34 | 0.76 (0.33-1.76) | 1.18 (0.45-3.08) |
| **Income levels** | | | |
| Low (T1) | 50/ 91 | 1 | 1 |
| Middle (T2) | 51/ 77 | 1.15 (0.69-1.90) | 1.15 (0.67-1.98) |
| High (T3) | 52/ 75 | 1.12 (0.69-1.92) | 1.21 (0.70-2.08) |
| **Residence** | | | |
| Urban area | 70/ 80 | 1 | 1 |
| Rural area | 83/ 163 | 0.59 (0.38-0.90) | 0.48 (0.30-0.77) |
| **BMI levels** | | | |
| Normal weight (18.5-22.9) | 92/ 153 | 1 | 1 |
| Underweight (<18.5) | 16/ 26 | 0.99 (0.50-1.97) | 0.97 (0.47-2.02) |
| Overweight (23-24.9) | 37/ 38 | 1.61 (0.94-2.75) | 1.69 (0.95-2.98) |
| Obese (≥25) | 8/ 26 | 0.50 (0.21-1.15) | 0.56 (0.23-1.36) |
| **Comorbidity** [a] | | | |
| No | 135/ 195 | 1 | 1 |
| Yes | 18/ 48 | 0.54 (0.28-1.01) | 0.49 (0.24-0.97) |
| **Pre-existing hematological condition** [b] | | | |
| No | 107/ 173 | 1 | 1 |
| Yes | 46/ 70 | 1.12 (0.71-1.77) | 0.90 (0.55-1.47) |
| **Pre-existing nephrological condition** [c] | | | |
| No | 114/ 205 | 1 | 1 |
| Yes | 39/ 38 | 1.90 (1.14-3.17) | 2.30 (1.32-4.01) |
| **Pre-existing hepatological condition** [d] | | | |
| No | 129/ 201 | 1 | 1 |
| Yes | 24/ 42 | 0.96 (0.54-1.68) | 1.11 (0.60-2.05) |
| **TNM stage** | | | |
| Stage I | 30/ 46 | 1 | 1 |
| Stage II | 83/ 134 | 0.99 (0.57-1.72) | 0.89 (0.48-1.62) |
| Stage III-IV | 40/ 63 | 1.11 (0.60-2.10) | 1.08 (0.55-2.12) |
| **Breast cancer subtype** | | | |
| Luminal/HER2-negative | 52/ 111 | 1 | 1 |
| Luminal/HER2-positive | 46/ 51 | 1.85 (1.09-3.14) | 1.78 (1.02-3.10) |
| HER2 enriched | 28/ 58 | 1.08 (0.61-1.91) | 0.89 (0.48-1.62) |
| Triple-negative | 27/ 23 | 2.98 (1.52-5.83) | 3.15 (1.56-6.34) |
| **Sequential anthracycline and taxane** | | | |
| No | 37/ 79 | 1 | 1 |
| Yes | 116/ 164 | 1.55 (0.96-2.51) | 1.47 (0.85-2.53) |
| **Dose-dense chemotherapy** | | | |
| No | 129/ 220 | 1 | 1 |
| Yes | 24/ 23 | 1.86 (1.00-3.49) | 1.80 (0.91-3.57) |

[1]Multivariable mode 1 was adjusted for age groups at diagnosis, income levels, and residence.

[2] Multivariable model 2 was the multivariable model 1 with additional adjustment for BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, TNM cancer stage, breast cancer subtype, sequential anthracycline and taxane and dose-dense chemotherapy.
[a] Having diagnosis of specific comorbidities, including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Having at least one of hematological symptoms (grade≥1), including anemia, neutropenia, lymphopenia, and thrombocytopenia within 120 days prior to chemotherapy.
[c] Having at least one of nephrological symptoms (grade≥1), including high creatinine, proteinuria, and hematuria within 120 days prior to chemotherapy.
[d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT within 120 days prior to chemotherapy.

**Table 13**: Association of demographic characteristics and clinical factors with neutropenia

| | **Neutropenia** (grade ≥3 vs. grade <3) | | |
|---|---|---|---|
| | *No*. of grade ≥3/ grade <3 | Model 1 Adjusted OR (95%CI)[1] | Model 2 Adjusted OR (95%CI)[2] |
| **Age group** | | | |
| < 40 | 19/ 42 | 1 | 1 |
| 40-49 | 48/ 105 | 1.05 (0.55-2.02) | 1.24 (0.62-2.50) |
| 50-59 | 41/ 94 | 1.07 (0.55-2.09) | 1.66 (0.79-3.50) |
| ≥ 60 | 9/ 38 | 0.55 (0.22-1.37) | 0.82 (0.29-2.32) |
| **Income levels** | | | |
| Low (T1) | 37/ 104 | 1 | 1 |
| Middle (T2) | 38/ 99 | 1.13 (0.66-1.94) | 1.13 (0.63-2.02) |
| High (T3) | 42/ 85 | 1.26 (0.73-2.16) | 1.29 (0.72-2.30) |
| **Residence** | | | |
| Urban area | 54/ 96 | 1 | 1 |
| Rural area | 63/ 183 | 0.62 (0.40-0.97) | 0.53 (0.32-0.87) |
| **BMI levels** | | | |
| Normal weight (18.5-22.9) | 72/ 173 | 1 | 1 |
| Underweight (<18.5) | 11/ 31 | 0.80 (0.38-1.69) | 0.84 (0.38-1.85) |
| Overweight (23-24.9) | 28/ 47 | 1.43 (0.82-2.49) | 1.54 (0.85-2.79) |
| Obese (≥25) | 6/ 28 | 0.54 (0.21-1.38) | 0.55 (0.20-1.50) |
| **Comorbidity** [a] | | | |
| No | 103/ 227 | 1 | 1 |
| Yes | 14/ 52 | 0.58 (0.29-1.14) | 0.56 (0.26-1.17) |
| **Pre-existing hematological condition** [b] | | | |
| No | 81/ 199 | 1 | 1 |
| Yes | 36/ 80 | 1.15 (0.71-1.86) | 1.01 (0.60-1.70) |
| **Pre-existing nephrological condition** [c] | | | |
| No | 88/ 231 | 1 | 1 |
| Yes | 29/ 48 | 1.63 (0.96-2.77) | 1.86 (1.04-3.30) |
| **Pre-existing hepatological condition** [d] | | | |
| No | 96/ 234 | 1 | 1 |
| Yes | 21/ 45 | 1.19 (0.66-2.14) | 1.39 (0.73-2.63) |
| **TNM stage** | | | |
| Stage I | 24/ 52 | 1 | 1 |
| Stage II | 68/ 149 | 1.01 (0.57-1.80) | 0.86 (0.46-1.61) |
| Stage III-IV | 25/ 78 | 0.76 (0.39-1.50) | 0.70 (0.34-1.44) |
| **Breast cancer subtype** | | | |
| Luminal/HER2-negative | 41/ 122 | 1 | 1 |
| Luminal/HER2-positive | 37/ 60 | 1.76 (1.01-3.05) | 1.69 (0.95-3.01) |
| HER2 enriched | 19/ 67 | 0.86 (0.46-1.62) | 0.71 (0.36-1.39) |
| Triple-negative/basal-like | 20/ 30 | 2.27 (1.15-4.52) | 2.53 (1.22-5.24) |
| **Sequential anthracycline and taxane** | | | |
| No | 24/ 92 | 1 | 1 |
| Yes | 93/ 187 | 2.00 (1.17-3.42) | 1.74 (0.96-3.17) |
| **Dose-dense chemotherapy** | | | |
| No | 94/ 255 | 1 | 1 |
| Yes | 23/ 24 | 2.73 (1.45-5.15) | 2.64 (1.32-5.25) |

[1]Multivariable mode 1 was adjusted for age groups at diagnosis, income levels, and residence.

[2] Multivariable model 2 was the multivariable model 1 with additional adjustment for BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, TNM cancer stage, breast cancer subtype, sequential anthracycline and taxane and dose-dense chemotherapy.
[a] Having diagnosis of specific comorbidities, including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Having at least one of hematological symptoms (grade≥1), including anemia, neutropenia, lymphopenia, and thrombocytopenia within 120 days prior to chemotherapy.
[c] Having at least one of nephrological symptoms (grade≥1), including high creatinine, proteinuria, and hematuria within 120 days prior to chemotherapy.

[d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT within 120 days prior to chemotherapy.

**Table 14**: Association of demographic characteristics and clinical factors with combined GI toxicities

| | **Combined GI toxicity** (grade ≥3 vs. grade <3) | | |
|---|---|---|---|
| | *No*. of grade ≥3/ grade <3 | Model 1 Adjusted OR (95%CI)[1] | Model 2 Adjusted OR (95%CI)[2] |
| **Age group** | | | |
| < 40 | 10/ 51 | 1 | 1 |
| 40-49 | 22/ 131 | 0.83 (0.36-1.88) | 0.86 (0.36-2.08) |
| 50-59 | 16/ 119 | 0.66 (0.28-1.57) | 0.58 (0.22-1.55) |
| 60+ | 3/ 44 | 0.33 (0.09-1.29) | 0.20 (0.04-2.94) |
| **Income levels** | | | |
| Low (T1) | 19/ 122 | 1 | 1 |
| Middle (T2) | 19/ 109 | 1.08 (0.54-2.16) | 1.37 (0.64-2.95) |
| High (T3) | 13/ 114 | 0.69 (0.32-1.49) | 0.89 (0.39-2.03) |
| **Residence** | | | |
| Urban area | 19/ 131 | 1 | 1 |
| Rural area | 32/ 214 | 1.02 (0.55-1.91) | 1.20 (0.61-2.36) |
| **BMI levels** | | | |
| Normal weight (18.5-22.9) | 37/ 208 | 1 | 1 |
| Underweight (<18.5) | 4/ 38 | 0.58 (0.19-1.74) | 0.46 (0.14-1.45) |
| Overweight (23-24.9) | 8/ 67 | 0.72 (0.32-1.64) | 0.74 (0.31-1.75) |
| Obese (≥25) | 2/ 32 | 0.36 (0.08-1.60) | 0.33 (0.07-1.58) |
| **Comorbidity** [a] | | | |
| No | 42/ 288 | 1 | 1 |
| Yes | 9/ 57 | 1.53 (0.66-3.57) | 1.92 (0.76-4.88) |
| **Pre-existing hematological condition** [b] | | | |
| No | 40/ 240 | 1 | 1 |
| Yes | 11/ 105 | 0.65 (0.32-1.32) | 0.70 (0.32-1.51) |
| **Pre-existing nephrological condition** [c] | | | |
| No | 45/ 274 | 1 | 1 |
| Yes | 6/ 71 | 0.49 (0.20-1.20) | 0.43 (0.16-1.13) |
| **Pre-existing hepatological condition** [d] | | | |
| No | 40/ 290 | 1 | 1 |
| Yes | 11/ 55 | 1.55 (0.73-3.27) | 1.80 (0.80-4.06) |
| **TNM stage** | | | |
| Stage I | 17/ 59 | 1 | 1 |
| Stage II | 20/ 197 | 0.32 (0.15-0.66) | 0.26 (0.12-0.59) |
| Stage III-IV | 14/ 89 | 0.51 (0.23-1.14) | 0.47 (0.20-1.10) |
| **Breast cancer subtype** | | | |
| Luminal/HER2-negative | 17/ 146 | 1 | 1 |
| Luminal/HER2-positive | 12/ 85 | 1.12 (0.51-2.49) | 1.24 (0.54-2.87) |
| HER2 enriched | 10/ 76 | 1.16 (0.50-2.71) | 1.43 (0.58-3.56) |
| Triple-negative/basal-like | 12/ 38 | 2.84 (1.23-6.56) | 3.60 (1.45-8.95) |
| **Sequential anthracycline and taxane** | | | |
| No | 15/ 101 | 1 | 1 |
| Yes | 36/ 244 | 0.83 (0.43-1.63) | 1.29 (0.60-2.77) |
| **Dose-dense chemotherapy** | | | |
| No | 48/ 301 | 1 | 1 |
| Yes | 3/ 44 | 0.43 (0.13-1.45) | 0.43 (0.12-1.52) |

[1]Multivariable mode 1 was adjusted for age groups at diagnosis, income levels, and residence.

[2] Multivariable model 2 was the multivariable model 1 with additional adjustment for BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, TNM cancer stage, breast cancer subtype, sequential anthracycline and taxane and dose-dense chemotherapy.
[a] Having diagnosis of specific comorbidities, including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Having at least one of hematological symptoms (grade≥1), including anemia, neutropenia, lymphopenia, and thrombocytopenia within 120 days prior to chemotherapy.
[c] Having at least one of nephrological symptoms (grade≥1), including high creatinine, proteinuria, and hematuria within 120 days prior to chemotherapy.
[d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT within 120 days prior to chemotherapy.

**Table 15**: Association of demographic characteristics and clinical factors with nausea/vomiting

| | **Nausea/vomiting** (grade ≥3 vs. grade <3) | | |
|---|---|---|---|
| | *No*. of grade ≥3/ grade <3 | Model 1 Adjusted OR (95%CI)[1] | Model 2 Adjusted OR (95%CI)[2] |
| **Age group** | | | |
| < 40 | 10/ 51 | 1 | 1 |
| 40-49 | 15/ 38 | 0.53 (0.22-1.26) | 0.50 (0.19-1.30) |
| 50-59 | 13/ 122 | 0.52 (0.21-1.29) | 0.35 (0.12-1.03) |
| 60+ | 2/ 45 | 0.21 (0.04-1.03) | 0.09 (0.01-0.61) |
| **Income levels** | | | |
| Low (T1) | 16/ 125 | 1 | 1 |
| Middle (T2) | 14/ 114 | 0.89 (0.41-1.93) | 1.17 (0.48-2.81) |
| High (T3) | 10/ 117 | 0.59 (0.25-1.38) | 0.75 (0.29-1.92) |
| **Residence** | | | |
| Urban area | 16/ 134 | 1 | 1 |
| Rural area | 24/ 222 | 0.88 (0.44-1.75) | 0.98 (0.45-2.10) |
| **BMI levels** | | | |
| Normal weight (18.5-22.9) | 33/ 212 | 1 | 1 |
| Underweight (<18.5) | 3/ 39 | 0.48 (0.14-1.67) | 0.40 (0.11-1.49) |
| Overweight (23-24.9) | 3/ 72 | 0.29 (0.09-0.98) | 0.27 (0.07-0.98) |
| Obese (≥25) | 1/ 33 | 0.22 (0.03-1.40) | 0.15 (0.02-1.33) |
| **Comorbidity** [a] | | | |
| No | 32/ 298 | 1 | 1 |
| Yes | 8/ 58 | 1.88 (0.76-4.68) | 2.91 (1.03-8.24) |
| **Pre-existing hematological condition** [b] | | | |
| No | 32/ 248 | 1 | 1 |
| Yes | 8/ 108 | 0.58 (0.26-1.32) | 0.74 (0.30-1.81) |
| **Pre-existing nephrological condition** [c] | | | |
| No | 35/ 284 | 1 | 1 |
| Yes | 5/ 72 | 0.54 (0.20-1.44) | 0.40 (0.13-1.21) |
| **Pre-existing hepatological condition** [d] | | | |
| No | 30/ 300 | 1 | 1 |
| Yes | 10/ 56 | 1.84 (0.83-4.09) | 2.27 (0.92-5.58) |
| **TNM stage** | | | |
| Stage I | 16/ 60 | 1 | 1 |
| Stage II | 15/ 202 | 0.25 (0.11-0.55) | 0.17 (0.07-0.41) |
| Stage III-IV | 9/ 94 | 0.32 (0.13-0.81) | 0.29 (0.11-0.76) |
| **Breast cancer subtype** | | | |
| Luminal/HER2-negative | 13/ 150 | 1 | 1 |
| Luminal/HER2-positive | 9/ 88 | 1.08 (0.44-2.66) | 1.23 (0.46-3.27) |
| HER2 enriched | 10/ 76 | 1.66 (0.68-4.06) | 2.56 (0.94-7.00) |
| Triple-negative/basal-like | 8/ 42 | 2.26 (0.86-5.93) | 2.90 (0.98-8.57) |
| **Sequential anthracycline and taxane** | | | |
| No | 11/ 105 | 1 | 1 |
| Yes | 29/ 251 | 0.94 (0.44-2.00) | 1.55 (0.63-3.81) |
| **Dose-dense chemotherapy** | | | |
| No | 37/ 312 | 1 | 1 |
| Yes | 3/ 44 | 0.59 (0.17-2.02) | 0.58 (0.16-2.19) |

[1]Multivariable mode 1 was adjusted for age groups at diagnosis, income levels, and residence.

[2] Multivariable model 2 was the multivariable model 1 with additional adjustment for BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, TNM cancer stage, breast cancer subtype, sequential anthracycline and taxane and dose-dense chemotherapy.
[a] Having diagnosis of specific comorbidities, including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Having at least one of hematological symptoms (grade≥1), including anemia, neutropenia, lymphopenia, and thrombocytopenia within 120 days prior to chemotherapy.
[c] Having at least one of nephrological symptoms (grade≥1), including high creatinine, proteinuria, and hematuria within 120 days prior to chemotherapy.
[d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT within 120 days prior to chemotherapy.

## 3. Discussion

In this study of 396 Vietnamese breast cancer patients, we found that a substantial proportion of them experienced severe (grade ≥ 3) hematological (38.6%) and GI (12.9%) toxicities associated with the admission of chemotherapeutic agents at the first-line treatment. Patients, particularly those who received chemotherapy with sequential anthracycline and taxane, were more likely to experience severe chemotherapy-induced hematological and GI toxicities, although no significant association was found. In multivariable analyses, we found that pre-existing nephrological condition was significantly associated with an increased risk of severe combined hematological toxicity and neutropenia. In addition, we found that the association with severe neutropenia was positively significant for dose-dense chemotherapy. Moreover, patients living in rural areas showed a lower risk of severe hematological toxicity than those living in urban areas. Furthermore, a significantly lower risk of severe combined GI toxicity in overall and nausea/vomiting in the toxicity-specific analysis were observed for patients diagnosed at stage II and stage III-IV. Finally, triple-negative/basal-like breast cancer was significantly associated with high risks of severe chemotherapy-induced hematological and GI toxicities compared with other breast cancer subtypes.

The incidence and severity of reported toxicities varied widely in previous studies due to various study designs; many were limited by data quality, result generalizability, or biased reporting between clinicians and cancer patients.[72,73] Almost all reports of chemotherapy-induced toxicities and their frequency came from clinical trials of new treatments and used clinician-reported toxicity ratings. It is well documented that clinical trial participants are biased towards healthier patients, and self-reported and clinical assessed toxicity differs.[72,73] Results from these settings may not reflect the frequency, severity, and burden of chemotherapy-induced toxicities in breast cancer patients receiving a real-world care. Nevertheless, our results are generally consistent with a recent systematic review and meta-analysis of seven clinical trials. Both hematological and non-hematological toxicities such as neutropenia, nausea, vomiting, and mucositis were common in breast cancer chemotherapy and mainly were more severe in anthracycline-based regimens.[211]

Few observational studies have investigated chemotherapy-induced adverse effects and factors associated with the severity of toxicities in breast cancer chemotherapy, including anthracycline-, taxane-, and non-anthracycline-based regimens. A 2014 population-based study evaluated the rates of first hospitalization

(i.e., interpreted as severe levels) caused by eight reasons, including neutropenia, infection, fever, thrombocytopenia, anemia, other adverse effects of chemotherapy, dehydration, and delirium, that occurred within six months of chemotherapy initiation among 3,567 breast cancer patients older than age 65 years from SEER/Texas Cancer Registry-Medicare database and 9,327 patients younger than age 65 years from the MarketScan database, diagnosed with stages I-IV breast cancer from 2003-2007. The study reported that among patients younger than age 65 years, the unplanned hospitalization rates ranged from 6.2% (dose-dense AC followed by paclitaxel) to 10.0% (TAC: docetaxel, doxorubicin, plus cyclophosphamide every three weeks); and around 6% of patients were hospitalized for neutropenia, fever, or infection. Among patients older than age 65 years, the rate of those who were admitted to the hospital ranged from 12.7% (TC: docetaxel and cyclophosphamide every three weeks) to 24.2% (TAC); and 12.4% of patients were hospitalized for neutropenia, fever, or infection.[212] In addition, that study suggested the regimens TAC and AC followed by docetaxel every three weeks (AC-T) were associated with the highest risk of chemotherapy-related hospitalization compared with the TC regimen for all age groups.[212] Moreover, the benefits of prophylactic G-CSF in reducing chemotherapy-related hospitalization rates and improving the ability of elderly patients to complete all cycles of chemotherapy were seen in SEER-Medicare patients.[213-215] In our study, few patients received prophylactic G-CSF treatment, likely due to limited resources. In Vietnam, a guideline for breast cancer treatment, released in July 2018,[52] recommended admitter with prophylactic G-CSF when receiving a dose-dense AC with sequential paclitaxel, a dose-dense AC followed by sequential weekly paclitaxel, and docetaxel plus cyclophosphamide (TC).[52] The National Health Insurance has approved reimbursement for myeloid growth factor supports only when the presence of chemotherapy-induced toxicities is documented or recorded. Therefore, G-CSF was often offered after the appearance of chemotherapy-induced toxicities among Vietnamese breast cancer patients, as observed in our study, i.e., patients who received G-CSF were more likely to be those with severe hematological toxicities. Our study found that patients who received dose-dense chemotherapy were more likely to have severe neutropenia. Over haft (52.1%) and almost of patients (97.0%) who had drug dose reduction (RDI < 85%) and chemotherapy discontinuance experienced severe combined hematological toxicity. These results highlight the importance and necessity of prophylactic G-CSF for breast cancer chemotherapy in Vietnam.

Another observed data from SEER registries among 1945 women aged 20-79 diagnosed with early-stage breast cancer from 2013-2014 measured seven toxicities, including nausea/vomiting, diarrhea, constipation, pain, arm edema, dyspnea, and breast skin irritation.[72] Approximately 45% of patients reported at least one severe/very severe toxicity, 9% reported unscheduled clinic visits for toxicity management, and 5% visited an emergency department or hospital approximately seven months after diagnosis. Nearly 25% of chemotherapy recipients endorsed severe/very severe nausea/vomiting during their cancer treatment. Women who received both chemotherapy and radiotherapy had 30% higher odds of more severe toxicity than those receiving only chemotherapy.[72] Similarly, high rates (~80%) of GI toxicities from a prospective study of 604 Italian women based on self-reports which documented toxicities while receiving adjuvant chemotherapy.[54] The GI toxicities were less common in our study participants. Our two followed-up surveys were carried out from 6 to 11 months for the first follow-up and from 12 to 18 months for the second follow-up after study enrollment. Not assessing side effects during active chemotherapy cycle(s) may have led to underestimating GI toxicities because most were based on self-reports.

Our study found that participants who received chemotherapy with sequential anthracycline and taxane were more likely to experience higher incidences of severe hematological and GI toxicities. Among participants receiving sequential anthracycline and taxane, the incidence and severity of GI toxicities were decreased after breast cancer patients began taxane, but no significant differences were observed for all hematological toxicities induced by anthracycline or by taxane post anthracycline. Our findings are in line with the previous studies and support the evidence that adding taxane (e.g., paclitaxel) sequentially to the anthracycline-based regimens does not increase the overall incidence and severity of toxicity.[216,217]

Severe chemotherapy-induced toxicities may lead to dose delay or dose reduction, chemotherapy discontinuance, and high costs for health care services to manage these side effects, which may result in premature death.[64-70] As a result, identifying patients at higher risk before chemotherapy, such as demographic characteristics and clinical factors, may have a significant clinical impact. It may enable caregivers to initiate supportive measures before the onset of complications.[218] Previous studies have suggested that race, age, comorbidity, and BMI may be associated with chemotherapy toxicities.[219,220] Racial differences in acute toxicities were notably documented in women with breast cancer who received FEC 100 chemotherapy (fluorouracil 500

mg/m$^2$, epirubicin 100 mg/m$^2$, and cyclophosphamide 500 mg/m$^2$) and TC (docetaxel plus cyclophosphamide) regimen.[221,222] There was a general trend toward a higher incidence and severity of hematological toxicity experienced by Asians than Caucasians (over 30% vs. < 5%) when G-CSF use was held consistent, whereas reporting of non-hematological toxicities (~20%) did not reveal significant differences across populations for both regimens.[221,222] Asian race and low BMI (underweight or normal weight, BMI <25 kg/m$^2$) were significantly associated with severe hematological toxicity during FEC 100 chemotherapy.[78] Evidence of a strong relationship between low/normal BMI and increased incidence of severe neutropenia was reported in a cohort study of 6,248 women with early-stage breast cancer.[57] However, our study did not find that severe hematological toxicity was significantly associated with age, comorbidity, BMI, TNM cancer stage, and other evaluated factors. Lower incidences of severe combined hematological toxicity and GI toxicity were documented in breast cancer patients with underweight (BMI<18.5 kg/m$^2$) and obese (≥ 25 kg/m$^2$) when compared to patients with normal weight (BMI 18.5-22.9 kg/m$^2$) and overweight (23.0-24.9 kg/m$^2$). We also found that overweight patients had a significantly lower risk of severe nausea/vomiting in the toxicity-specific analysis. The reasons behind the observed reduced risk of chemotherapy-induced toxicities among lean and obese cancer patients are unknown. One possible explanation is that obese patients are more likely to receive planned empirical dose reduction (aka. “dose capping”),[223] but this cannot explain why underweight patients had low toxicities. Further investigation on pharmacokinetic profiles of chemotherapy agents among obese patients and patients underweight is warranted.[224] The “dose capping” might also be partially explainable for the significantly inverse association with severe nausea/vomiting among patients diagnosed at stage III-IV. They had a higher prevalence of overweight and obese in our study.

We found that pre-existing nephrological condition (i.e., elevated creatinine, proteinuria, and hematuria before chemotherapy) was significantly associated with an increased risk of severe combined hematological toxicity, neutropenia in particular, whereas the pre-existing hematological and hepatological conditions were not significantly associated with hematological toxicity. A study involving 619 patients aged ≥ 65 with early-stage breast cancer, who received CMF, AC, or capecitabine reported that pretreatment renal function did not influence the occurrence of hematologic toxicity regardless of regimen, whereas an increased creatine clearance at baseline was highly related to the occurrence of non-hematologic toxicity for the AC regimen and

very mildly for the capecitabine regimen, but not related for the CMF regimen.[225] Another study revealed a positive association between severe chemotherapy-induced toxicities associated with decreased creatine clearance at baseline among older patients with cancer.[226] Inconsistent with our findings, lower pretreatment blood counts (e.g., WBC, ANC) and Hgb were previously suggested to be associated with chemotherapy-induced hematological toxicities.[96,216,218] Although not confirmed, we speculate that oncologists might have considered patients' pre-existing hematological condition in cancer treatment. Nevertheless, our finding on renal function and chemotherapy toxicity reinforces the importance of considering renal function before administering chemotherapy.

In our study, patients living in rural areas showed a lower risk of severe hematological toxicity than those living in urban areas. The physical and financial burden to our participants who resided in rural areas, including the cost of medical care services, travel time to the hospital, and distance from patients' homes to the hospital, might be considered by clinicians' decisions when selecting appropriate chemotherapy regimens and schedules. However, no differences in chemotherapy regimens and schedules were observed in residential areas. This suggests that other factors, such as lifestyle, physical function, physical activity, diet habits, gut microbiome, and family and social support, may contribute to the association between residential areas and the risk of hematological toxicity. For example, nutritional deficiencies (e.g., vitamin B12 and folate deficiencies) have been associated with neutropenia.[227] Further studies are needed to understand the impact of these factors on chemotherapy-induced toxicities in Vietnam.

Evidence of association between chemotherapy-induced toxicity with molecular subtypes of breast cancer is limited. Breast cancer is a known heterogeneous disease, and its molecular subtypes have different chemotherapy regimens and schedules as well as different therapeutic responses and clinical outcomes.[228] Thus, they may have different associations with chemotherapy-induced toxicities. Our study found that patients with triple-negative/basal-like breast cancer were significantly associated with high risks of severe chemotherapy-induced hematological and GI toxicities compared with other breast cancer subtypes, except for severe nausea/vomiting. In addition, a positive association with combined hematological toxicity was also observed for patients with luminal/HER2-positive when compared with patients with luminal/HER2-negative.

Our study is the first to investigate associations of severe chemotherapy-induced toxicities with demographic characteristics and clinical features among Vietnamese breast cancer patients. The participation rate for the patients approached by our research team was high (93.1%). In addition, the availability of blood and urine test results before and during each cycle of chemotherapy/hospital visit and detailed clinical information of breast cancer patients are strengths of this study. The chemotherapy-induced hematological toxicities were captured during the first-line chemotherapy treatment and 90 follow-up days after the treatment. In our research, non-hematological chemotherapy-induced toxicities (i.e., GI toxicities) were identified through a combination of patient self-report side effects at the two follow-ups and a review of the assessments recorded by treating physicians/nurses during each cycle of chemotherapy/hospital visit, which minimize the concern related to underestimated non-hematological chemotherapy-induced toxicities by clinicians. The patient self-report side effects information was collected through a structured questionnaire administered by trained interviewers following a standard protocol. However, as aforementioned, our study is limited by not collecting toxicity information during active chemotherapy. In addition, the response rates of follow-up surveys were moderate for the first follow-up (77.8%) and the second follow-up (62.6%). Patients diagnosed at early stages were more likely to complete the first follow-up than those diagnosed at late stages. The response rates of the first follow-up surveys respectively were 85.5%, 76.5%, and 74.8% for patients with stage I, stage II, and stage III-IV. The loss of follow-up would likely affect the statistical power for GI toxicity assessment but not hematological toxicity, as the latter was assessed solely via medical chart review. Furthermore, our study only accounted for ~35% of newly diagnosed breast cancer patients treated at Vietnam National Cancer Hospital and Hanoi Oncology Hospital from July 2017 to June 2018. Therefore, selection bias cannot be completely ruled out, and our findings may not be generalizable to all breast cancer cases, particularly those treated in other settings in Vietnam.

In conclusion, we found that a substantial proportion of breast cancer patients in Vietnam suffered severe hematological and GI toxicities during their first-line chemotherapy. Our study characterized the burden of chemotherapy-induced toxicity faced by patients that would be valuable to assist Vietnamese oncologists/clinicians in treatment planning, dose adjustments, and managing side effects. In addition, our study calls for further research on factors related to chemotherapy-induced toxicities and factors related to

interindividual variations in chemotherapy-induced toxicities that may facilitate the delivery of personalized treatment and improve treatment outcomes.

# CHAPTER 3

SPECIFIC AIM 2

## GI microbiome and its associations with clinical and non-clinical factors

To evaluate the associations between GI microbiome and sociodemographic and clinical factors among breast cancer patients.

## 1. Methods

### 1.1. Parent Study

This study was based on a prospective follow-up of 501 newly diagnosed Vietnamese breast cancer patients who were recruited into the VBCS. Details of breast cancer case recruitment in the VBCS have been described in the Aim 1. [26,207]

### 1.2. Population Selection

To be included in this analysis, newly diagnosed breast cancer patients in this study must have provided stool samples at baseline before systemic treatment (i.e., chemotherapy) regardless of their status of receiving breast cancer surgery. We excluded participants who were subsequently confirmed to have a benign tumor based on pathological reviews (n=9) and those diagnosed at stage 0 (n=2). We also excluded participants who did not donate stool samples (n=96) at the baseline survey. Stool samples from 4 participants were excluded due to low DNA yields (n=4). In addition, participants with incomplete medical chart reviews or missing treatment information were excluded (n=34). Finally, a total of 356 participants was included for Aim 2 (**Figure 5**).

### 1.3. Outcome Assessment

#### 1.3.1. Stool sample collection:

Stool samples were collected using fecal occult blood test (FOBT) cards in the enrollment for all breast cancer patients before neoadjuvant/adjuvant chemotherapy following a standard protocol. Trained study staff provided two FOBT cards for each patient with clear instructions to self-collect stool samples whether in-hospital or at home. The stool samples were transferred to the research laboratory within 24 hours after

collection, were stored with no additives at room temperature during transportation and frozen at -80°C until assays in order to minimize microbial growth. In our study, collecting stool samples after a breast cancer surgery and before chemotherapy were more common (54.5%) than collecting stool samples before a breast cancer surgery or before the initiation of chemotherapy (45.5%) among patients who received neoadjuvant chemotherapy and patients who received chemotherapy without a breast cancer surgery.

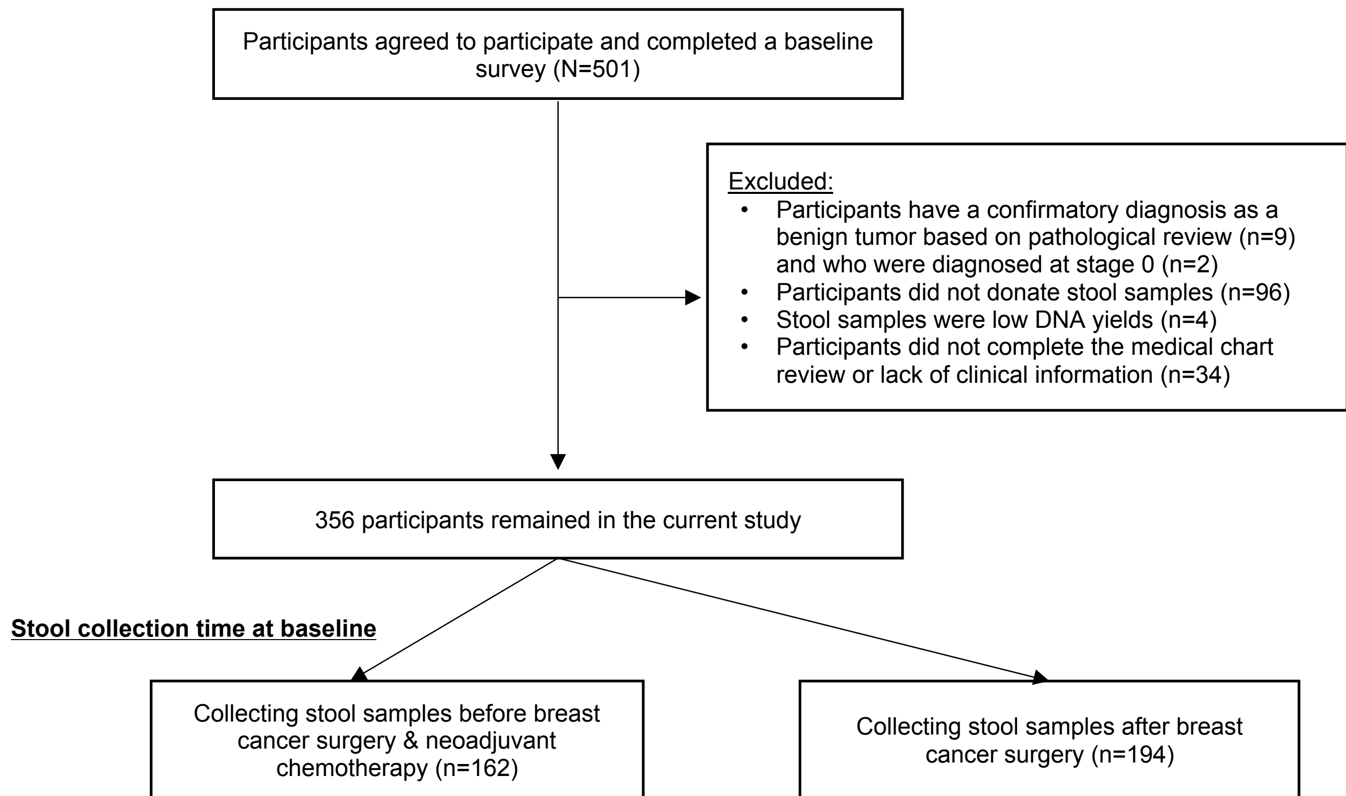



**Figure 5:** Flow diagram of study participant inclusion criteria for Aim 2

1.3.2. Microbiome profiling

***DNA extraction and shotgun metagenomic sequencing***

The DNeasy PowerSoil Kit (Qiagen) was used to isolate DNA from buccal samples, following protocols provided by the manufacturer. Then, the TruePrep DNA Library Prep Kit V2 or Nextera XT DNA Library Preparation Kit (Illumina) was utilized to build sequencing libraries from DNA samples for shotgun metagenomic sequencing, following manufacturers' instructions. Finally, sequencing was conducted at paired-end 150bp using the Illumina HiSeq System at BGI Americas. DNA extraction, library preparation, and

sequencing for all samples were performed in one batch. A total of 356 participants' stool samples, with sequencing data available, were included in downstream analyses.

***Sequencing data processing***

On average, 10.19 million (Min-Max: 4.05-10.90) raw sequencing reads were obtained for each sample from our study. Raw reads were processed by Trimmomatic v0.39 to trim low-quality bases,[229] after which reads with fewer than 105 nucleotides, i.e., 70% of original read lengths, were discarded. Then, Bowtie2 v2.3.0 was used to remove reads that could be mapped on the human genome (GRCh38).[230] After quality-trimming and human reads removal steps, respective averages of 10.14 million (Min-Max: 4.05-10.89) clean reads per sample were retained for downstream analyses. Clean reads were subjected to taxonomic profiling and estimating the absolute abundance of microbial taxa by using Kraken v2.1.1 and Bracken v2.6, with human bacterial genomes from the Unified Human Gastrointestinal Genome (UHGG) collection as the reference.[231-233] Within each sample, only taxa with a relative abundance of >0.001% were considered detected.[234,235]

1.4. Covariate Assessment

We collected information on covariates, including sociodemographic characteristics and clinical factors which are summarized in Table 16.

The mean age of 356 study participants was 48.8 years at diagnosis and treatment. Approximately 61.7% of patients lived in rural areas, and 39.6% of cases had attained a high school, college, or higher education. The percentages of underweight (BMI <18.5 kg/m$^2$), overweight (BMI: 23-24.9 kg/m$^2$), and obese (BMI ≥ 25 kg/m$^2$) in our breast cancer patients were 9.0%, 18.5%, and 9.0%, respectively. Comorbidity was reported by 18.5% of patients. Approximately 31.7 % and 18.0 % of breast cancer patients, respectively, experienced moderate (4-8 months) and serious (9 months) delays in diagnosis. Over half (54.8%) of participants were diagnosed at stage II, while 19.2% were diagnosed at stage I, and 26.0% were diagnosed at stage III or later.

**Table 16**: Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 2)

| | All eligible participants (N = 356) | |
|---|---|---|
| | n | % |
| **Age group** | | |
| < 40 | 48 | 13.5 |
| 40-49 | 135 | 37.9 |
| 50-59 | 120 | 33.7 |
| 60+ | 53 | 14.9 |
| **Education levels** | | |
| Primary school | 55 | 15.4 |
| Middle school | 160 | 44.9 |
| High school | 78 | 21.9 |
| College or higher | 63 | 17.7 |
| **Income levels** | | |
| Low (T1) | 129 | 36.2 |
| Middle (T2) | 114 | 32.0 |
| High (T3) | 113 | 31.7 |
| **Residence** | | |
| Urban area | 136 | 38.2 |
| Rural area | 220 | 61.7 |
| **Family history of breast cancer** | | |
| No | 342 | 96.1 |
| Yes | 14 | 3.9 |
| **Menopausal status** | | |
| Pre-menopausal | 290 | 55.1 |
| Post-menopausal | 66 | 44.9 |
| **BMI levels** ($kg/m^2$) | | |
| Underweight (<18.5) | 32 | 9.0 |
| Normal weight (18.5-22.9) | 226 | 63.5 |
| Overweight (23-24.9) | 66 | 18.5 |
| Obese (≥25) | 32 | 9.0 |
| **Comorbidity** [a] | | |
| No | 290 | 81.5 |
| Yes | 66 | 18.5 |
| **Diagnosis delay** [b] | | |
| No delay | 179 | 50.3 |
| Moderate delay | 113 | 31.7 |
| Serious delay | 64 | 18.0 |

| | All eligible participants (N = 356) | |
|---|---|---|
| | n | % |
| **ER status** | | |
| Negative | 138 | 38.8 |
| Positive | 218 | 61.2 |
| **PR status** | | |
| Negative | 163 | 45.8 |
| Positive | 193 | 54.2 |
| **HER2 status** | | |
| Negative | 196 | 55.1 |
| Positive | 160 | 44.9 |
| **Ki-67 levels** | | |
| <20% | 133 | 37.4 |
| ≥20% | 223 | 62.6 |
| **Breast cancer subtypes** | | |
| Luminal/HER2-negative | 144 | 40.4 |
| Luminal/HER2-positive | 87 | 24.4 |
| HER2 enriched | 86 | 20.5 |
| Triple-negative | 52 | 14.6 |
| **Tumor size stage** | | |
| 1 | 95 | 26.7 |
| 2 | 202 | 56.7 |
| 3 | 31 | 8.7 |
| 4 | 28 | 7.9 |
| **Node stage** | | |
| 0 | 192 | 53.9 |
| 1 | 98 | 27.5 |
| 2 | 47 | 13.2 |
| 3 | 20 | 5.3 |
| **TNM stage** | | |
| Stage I | 74 | 19.2 |
| Stage II | 217 | 54.8 |
| Stage III-IV | 85 | 26.0 |
| **Historical subtypes** | | |
| ICD | 264 | 74.1 |
| Non-ICD | 43 | 12.1 |
| Unknown | 49 | 13.8 |

[a] Having diagnosis of specific comorbidities including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at enrollment.
[b] Diagnosis delay: a delay in seeking medical care from the first self-discovery symptom onset to the first medical visit.

ER+, PR+, and HER2-positive accounted for 61.2%, 54.2%, and 44.9% of breast cancer patients. Most participants had luminal/HER2-negative subtypes (40.4%). The percentage of breast cancer patients with

HER2 enriched and triple-negative/basal-like subtypes was 20.5% and 14.6%, respectively. The majority of patients had invasive ductal carcinoma (74.1%), whereas 12.1% had unknown histological subtypes.

1.5. Statistical Analysis

To evaluate the associations between the GI microbiome and sociodemographic and clinic factors among breast cancer patients, we first evaluated overall microbial richness (alpha diversity) and composition (beta diversity). Because sequencing depth might affect both alpha and beta diversity estimates,[236] we first rarefied the species level absolute abundance, i.e., read counts, of every sample to the minimum number of clean reads (n= 3,578,947) among 356 samples, using the R function *vegan::rarefy* [237]. Then alpha diversity and beta diversity were calculated based on the rarefied species level absolute abundance data using the R functions *vegan::diversity* and *vegan::vegdist*, respectively.[237]

In our study, alpha diversity was measured by the Chao 1 richness index (Chao1 index), Shannon - Wiener diversity index (Shannon index), and inverse Simpson diversity index. Median and interquartile range (IQR) for alpha diversity indexes were calculated by patients' demographic and clinical characteristics and differences across these subgroups compared using non-parametric tests (i.e., Wilcoxon rank sum and Kruskal-Wallis rank sum tests). In addition, the Shannon index and inverse Simpson index were transformed to the square of the Shannon index and the square root of the inverse Simpson index to normalize their distributions before evaluating the association between alpha diversity indexes with selected sociodemographic and clinic factors. The Chao1 index, the square of Shannon index, and the square root of the inverse Simpson index within selected sociodemographic and clinic factor strata were estimated by the mean difference ($\beta$ Coefficients) and 95% CIs in linear regression models. Models were adjusted for covariates: age group, income levels, residence, menopausal status, body mass index (BMI, kg/m$^2$), comorbidity, diagnosis delay, breast cancer subtypes, and TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk). In our study, almost all participants (~94%) who received breast cancer surgery also received an antibiotics treatment within approximately one week after surgery (100%). Since breast cancer surgery and antibiotics post-surgery might substantially affect the gut microbiome profile, stool sample time, i.e., before breast cancer surgery and after breast cancer surgery may modify the association between the alpha diversity indexes with the clinical and non-clinical factors. We performed sub-

analyses with stratification by stool collection time, and the overall analysis was additionally adjusted for stool collection time relative to surgery. Beta diversity (the total variance of a gut microbial composition) was measured by Bray-Curtis dissimilarity matrix (Bray-Curtis), weighted UniFrac distance matrix (wUniFrac) and unweighted UniFrac distance (uwUniFrac). The Permutational Multivariate Analysis of Variance ( PERMANOVA) test was implemented to assess whether there was a difference regarding GI microbial composition according to sociodemographic and clinic factors.[238] R square and p values from PERMANOVA tests were produced in models adjusted for the aforementioned potential covariates and 999 permutations using the R functions *vegan.*[237] All statistical analyses were performed at two-sided tests, and associations with $P<0.05$ were considered statistically significant.

We evaluated the associations of GI microbial taxa with sociodemographic and clinic factors via logistic regression analyses. For individual taxa, we defined their presence as relative abundance ≥ 1/3,578,947 = 0.0000279% in a sample (i.e., ≥ 1 read when there were 3,578,947 reads, the minimum number of clean reads among 356 samples. In our study, common taxa were defined if present in >50% of samples; rare taxa were defined if shown in <50% of samples; We limited our analysis to those rare taxa in 10-50% of samples. For common taxa, centered log-ratio (clr) transformation [239] was utilized to normalize the absolute abundance of taxa at each taxonomic level from phylum to species, with zeros replaced by the minimum read count value of the whole dataset,[239] and general linear regression was used to evaluate associations between clr-transformed taxa abundance and the sociodemographic and clinic factors. For rare taxa, the negative binomial hurdle model that handles zero-inflated data were performed to evaluate the association of clinical and non-clinical factors with the abundance of taxa without transformation using R package “pscl”. $\beta$-Coefficients, standard error (SE) and p values for individual rare taxa were produced based on the zero-hurdle part of the model. Analyses were conducted for all 356 breast cancer participants and stratified by stool collection time with models adjusted for the aforementioned potential covariates. False discovery rate (FDR) was calculated at each taxonomic level separately for overall or stratified analyses to account for multiple testing. Association with FDR-corrected p-value ($P_{FDR}$) of <0.1 was considered statistically significant. All statistical analyses were performed using R version 3.6.3.

## 1.6. Statistical Power Estimation

*GI microbial richness*

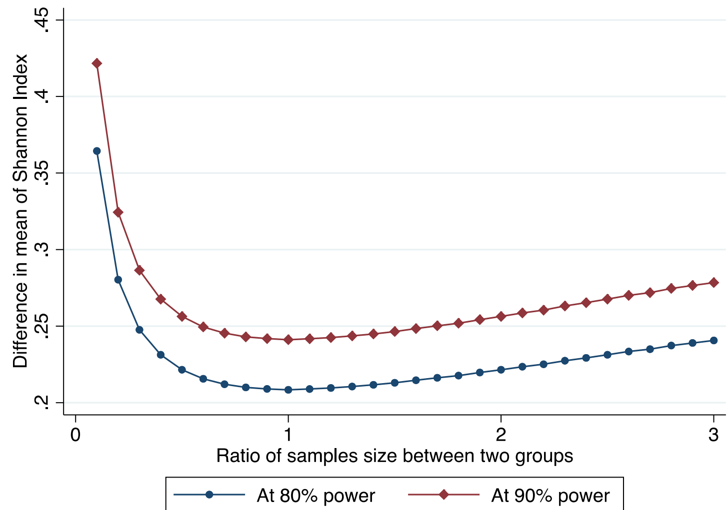



**Figure 6:** Power curve for a test of mean Shannon index according to sample size ratios between two groups

Among 356 fecal samples, we assumed the Shannon index would normally be distributed with a mean of 3.9 and a standard deviation of 0.7. We have 80% power to detect a true difference in the mean of Shannon index with a range from 0.21 to 0.36 according to the sample size ratios between two comparison groups with sample size ratios ranging from 0.1 to 3.0. (**Figure 6**).

*GI microbial taxa*

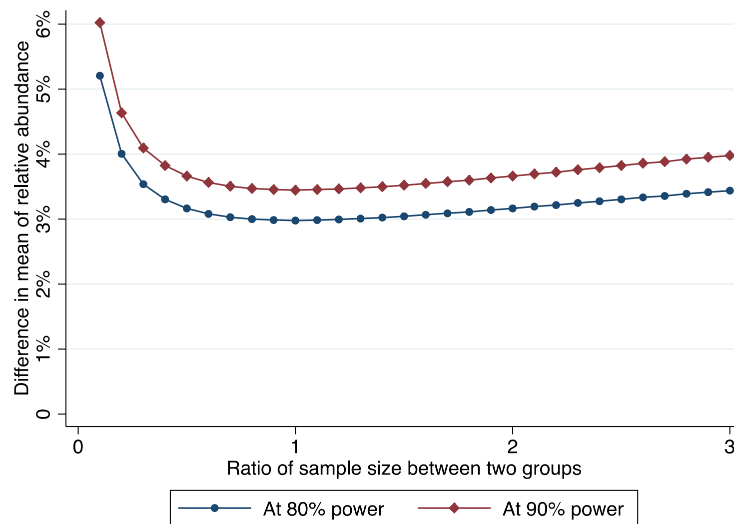



**Figure 7:** Power curve for a test of mean relative abundance according to sample size ratios between two groups

We assumed the relative abundance of the most common taxa at the phylum level (e.g., *Firmicutes* and *Bacteroides*) would normally be distributed with a standard deviation of 10% referred data from a case-only study among 31 breast cancer patients of Luu et al.[150] We have 80% power to detect a true difference in mean of relative abundance with a range from 2.9% to 5.2% according to the ratio of sample size between two groups ranged from 0.1 to 3.0. Likewise, at 90% power, we can detect a true difference in the mean of relative abundance with a range from 3.4% to 6.0% according to the same ratios of sample size between two groups (**Figure 7**).

## 2. Results

**Table 17**: Demographic characteristics and clinical factors of participants by stool collection time

| | Overall | Stool collection time | | |
|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy | Collection after breast cancer surgery | |
| | N=356<br>n | N=162<br>n (%) | N=194<br>n (%) | *P*[1] |
| **Age at diagnosis** | | 49.5±9.6 | 49.6±9.5 | |
| < 40 | 48 | 22 (13.6) | 26 (13.4) | 0.86 |
| 40-49 | 135 | 58 (35.8) | 77 (39.7) | |
| 50-59 | 120 | 58 (35.8) | 62 (32.0) | |
| ≥ 60 | 53 | 24 (14.8) | 29 (14.9) | |
| **Education** | | | | |
| Primary school | 55 | 28 (17.3) | 27 (13.9) | 0.18 |
| Middle school | 160 | 76 (46.9) | 84 (43.3) | |
| High school | 78 | 37 (22.8) | 41 (21.1) | |
| College or higher | 63 | 21 (13.0) | 42 (21.6) | |
| **Income** | | | | |
| Low (T1) | 129 | 73 (45.1) | 56 (28.9) | 0.003 |
| Middle (T2) | 114 | 49 (30.2) | 65 (33.5) | |
| High (T3) | 113 | 40 (24.7) | 73 (37.6) | |
| **Location** | | | | |
| Urban area | 136 | 46 (28.4) | 90 (46.4) | <0.001 |
| Rural area | 220 | 116 (71.6) | 104 (53.6) | |
| **Menopausal status** | | | | |
| Premenopausal | 196 | 82 (50.6) | 114 (58.8) | 0.12 |
| Postmenopausal | 160 | 80 (49.4) | 80 (41.2) | |
| **Family history of breast cancer** | | | | |
| No | 342 | 156 (96.3) | 186 (95.9) | 0.84 |
| Yes | 14 | 6 (3.7) | 8 (4.1) | |
| **BMI levels** | | | | |
| Under weight (<18.5) | 32 | 16 (9.9) | 16 (8.2) | 0.69 |
| Normal weight (18.5-22.9) | 226 | 98 (60.5) | 128 (66.0) | |
| Asian overweight (23.0-24.9) | 66 | 31 (19.1) | 35 (18.0) | |
| Asian obese (≥ 25) | 32 | 17 (10.5) | 15 (7.7) | |
| **Comorbidity** | | | | |
| No | 290 | 134 (82.7) | 156 (80.4) | 0.60 |
| Yes | 66 | 28 (17.3) | 38 (19.6) | |
| **Diagnosis delay** | | | | |
| No delay | 179 | 65 (40.1) | 114 (58.8) | 0.001 |
| Moderate delay | 113 | 58 (35.8) | 55 (28.3) | |
| Serious delay | 64 | 39 (24.1) | 25 (12.9) | |
| **ER status** | | | | |
| Negative | 138 | 72 (44.4) | 66 (34.0) | 0.04 |
| Positive | 218 | 90 (55.6) | 128 (66.0) | |
| **PR status** | | | | |
| Negative | 163 | 86 (53.1) | 77 (39.7) | 0.01 |
| Positive | 193 | 76 (46.9) | 117 (60.3) | |
| **HER2 status** | | | | |
| Negative | 196 | 88 (54.3) | 108 (55.7) | 0.80 |
| Positive | 160 | 74 (45.7) | 86 (44.3) | |
| **Ki-67 (%)** | | | | |
| < 20% | 133 | 53 (32.7) | 80 (41.2) | 0.10 |
| ≥ 20% | 223 | 109 (67.3) | 114 (58.8) | |
| **Breast cancer subtypes** | | | | |
| Luminal/HER2-negative | 144 | 58 (35.8) | 86 (44.3) | 0.03 |

| | Overall | Stool collection time | | |
|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy | Collection after breast cancer surgery | |
| | N=356<br>n | N=162<br>n (%) | N=194<br>n (%) | *P*[1] |
| Luminal/HER2-positive | 87 | 34 (21.0) | 53 (27.3) | |
| HER2 enriched | 73 | 40 (24.7) | 33 (17.0) | |
| Triple-negative/basal-like | 52 | 30 (18.5) | 22 (11.3) | |
| **TNM cancer stage** | | | | |
| I | 74 | 14 (8.6) | 60 (30.9) | <0.001 |
| II | 197 | 76 (46.9) | 121 (61.4) | |
| III-IV | 85 | 72 (44.5) | 13 (6.7) | |
| **Historical subtypes** | | | | |
| ICD | 264 | 111 (68.5) | 153 (78.9) | 0.06 |
| Non-ICD | 43 | 22 (13.6) | 21 (10.8) | |
| Unknown | 49 | 29 (17.9) | 20 (10.3) | |

[1] p-value for chi-square tests

**Table 17** shows demographic characteristics and clinical factors of 356 breast cancer patients by stool collection time. The mean age was 49.6 years for breast cancer cases who donated stool samples after breast cancer surgery and 49.5 years for patients who donated stool samples before a breast cancer surgery or before the initiation of chemotherapy among patients who received neoadjuvant chemotherapy and patients who received adjuvant chemotherapy without a breast cancer surgery. Compared with patients whose stool samples were collected after breast cancer surgery, patients with stool collection before breast cancer surgery and neoadjuvant chemotherapy were similar regarding age at diagnosis and educational attainment but more likely to have low income and live in rural areas. No differences were observed between cases and controls regarding menopausal status, family history of cancer, BMI levels, and comorbidity. On the other hand, compared with patients who collected stool samples before breast cancer surgery and neoadjuvant chemotherapy, patients with stool collection after breast cancer surgery tended to be diagnosed at early stages (30.9% vs. 8.6% for stage I and 61.4% vs. 46.9% for stage II), less likely experienced moderate (4-8 months) and serious delays (≥ 9 months) in diagnosis and had a lower percentage of breast cancer patients with HER2 enriched (17.0% vs. 24.7%) and triple-negative/basal-like subtypes (11.3% vs. 18.5%).

**Association of GI microbial richness and composition with stool sample collection time, sociodemographic and clinic factors**

After quality control and rarefaction, a total of 4206 OTUs were identified among 356 breast cancer patients, based on which alpha diversity indexes (Chao1, Shannon, and inverse Simpson indexes) and beta

dissimilarity matrices (Bray-Curtis dissimilarity, weighted UniFrac distance, and unweighted UniFrac distance) were calculated.

**Table 18**: Comparison of alpha diversity indexes by stool collection time

| | Stool collection time | | |
|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy | Collection after breast cancer surgery | |
| | (N=162) | (N= 194) | |
| **Alpha diversity indexes** | Median (IQR) | Median (IQR) | *P*[1] |
| Chao1 richness index | 1006 (301) | 866 (351) | $3.04x10^{-5}$ |
| Shannon-Wiener diversity index | 4.12 (0.82) | 3.74 (0.87) | $6.22x10^{-6}$ |
| Inverse Simpson diversity index | 18.6 (18.6) | 14.5 (12.3) | 0.002 |

[1] p-value for Wilcoxon rank sum test

In our study, patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy showed higher alpha diversity (i.e., higher microbial richness and evenness) than the patients with stool collection after breast cancer surgery with p-values of $3.04x10^{-5}$, $6.22x10^{-6}$, and 0.002 for the Chao1 richness index, Shannon-Wiener diversity index and Inverse Simpson diversity index, respectively (**Table 18**). In addition, the significantly lower Shannon index was consistently observed among patients whose stool samples were collected within seven days after breast cancer surgery (Week 1; n=72), patients who collected stool samples within 8-14 days after breast cancer surgery (Week 2; n=23), and patients those collected stool samples within 15-21 days after breast cancer surgery (Week 3; n=62) in comparison with patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy. However, no difference in the Shannon index was found among patients whose stool samples were collected before breast cancer surgery, and neoadjuvant chemotherapy and patients whose stool samples were collected after breast cancer surgery > 21 days (Week 4; n=37) (**Figure 8**).

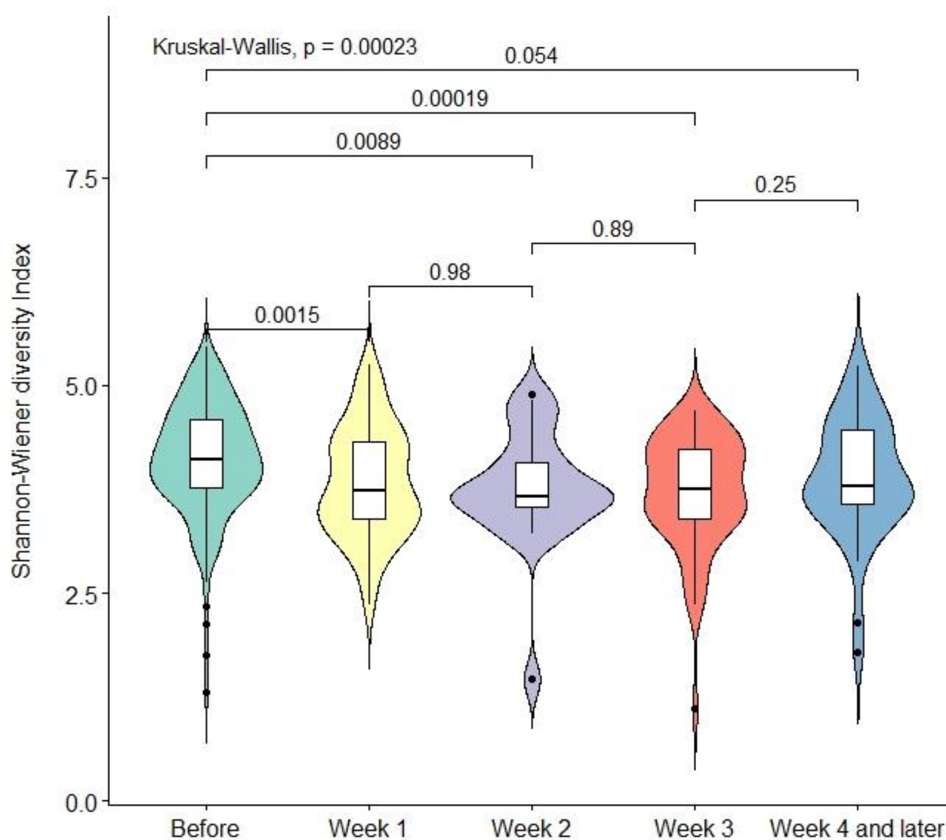



**Figure 8:** Alpha diversity at the species level (measured in term of the Shannon index) by stool collection time

(‘*Before*’: collected stool samples before breast cancer surgery and neoadjuvant chemotherapy (n=162); ‘*Week 1’*: collected stool samples within 7 days after breast cancer surgery (n=72); ‘*Week 2’*: collected stool samples within 8-14 days after breast cancer surgery (n=23); ‘*Week 3’*: collected stool samples within 15-21 days after breast cancer surgery (n=62); ‘*Week 4 and later’*: collected stool samples after breast cancer surgery > 21 days (n=37))

**Table 19-21** show the median and IQR of alpha diversity indexes (measured by Chao1, Shannon, and inverse Simpson indexes) by selected sociodemographic and clinic factors in analyses for overall and sub-analyses with stratification by stool collection time. The Shannon index was similar across subgroups defined by age at diagnosis, educational attainment, income, residence, menopausal status, family history of breast cancer, BMI levels, or other breast cancer-related characteristics except for diagnosis delay. Moderate and serious diagnosis delay groups were more likely to have lower alpha diversity than the no-delay group among patients who collected stool before breast cancer surgery (**Table 20**). A similar difference was observed for Chao1 and inverse Simpson indexes (**Table 19 and Table 20**). Surprisingly, we found a higher Chao1 and Shannon indexes among breast cancer patients with stage III-IV than patients with stage I in the overall analysis.

**Table 19**: Comparison of Chao1 index by demographic characteristics and clinical factors

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ |
| **Age at diagnosis** | | | | | | |
| < 40 | 1032 (372) | 0.08 | 1091 (390) | 0.35 | 1013 (330) | 0.13 |
| 40-49 | 921 (333) | | 1026 (288) | | 850 (361) | |
| 50-59 | 933 (313) | | 983 (288) | | 869 (305) | |
| 60+ | 879 (287) | | 945 (235) | | 849 (273) | |
| **Education** | | | | | | |
| Primary school | 955 (299) | 0.14 | 1047 (293) | 0.83 | 861 (259) | 0.16 |
| Middle school | 918 (334) | | 1001 (300) | | 843 (317) | |
| High school | 1011 (325) | | 980 (318) | | 1021 (377) | |
| College or higher | 875 (324) | | 1004 (304) | | 854 (345) | |
| **Income** | | | | | | |
| Low (T1) | 975 (316) | 0.09 | 1008 (312) | 0.94 | 911 (305) | 0.24 |
| Middle (T2) | 926 (344) | | 995 (339) | | 866 (361) | |
| High (T3) | 888 (371) | | 1011 (266) | | 841 (318) | |
| **Residence** | | | | | | |
| Urban area | 929 (358) | 0.75 | 1039 (324) | 0.66 | 869 (376) | 0.80 |
| Rural area | 935 (285) | | 990 (296) | | 866 (312) | |
| **Menopausal status** | | | | | | |
| Premenopausal | 940 (338) | 0.94 | 1026 (300) | 0.89 | 864 (372) | 0.65 |
| Postmenopausal | 923 (292) | | 990 (297) | | 872 (290) | |
| **Family history of breast cancer** | | | | | | |
| No | 934 (318) | 0.88 | 1005 (299) | 0.20 | 869 (355) | 0.18 |
| Yes | 944 (417) | | 1106 (180) | | 722 (277) | |
| **BMI levels** | | | | | | |
| Under weight (<18.5) | 906 (421) | 0.68 | 870 (633) | 0.42 | 986 (253) | 0.89 |
| Normal weight (18.5-22.9) | 956 (332) | | 1039 (274) | | 863 (365) | |
| Overweight (23.0-24.9) | 908 (308) | | 925 (277) | | 871 (328) | |
| Obese (≥ 25) | 879 (294) | | 934 (223) | | 861 (366) | |
| **Comorbidity** | | | | | | |
| No | 950 (324) | 0.10 | 1010 (296) | 0.33 | 874 (357) | 0.21 |
| Yes | 872 (303) | | 965 (247) | | 825 (284) | |
| **Diagnosis delay** | | | | | | |
| No delay | 940 (377) | 0.57 | 1084 (406) | 0.003 | 861 (392) | 0.59 |
| Moderate delay | 946 (297) | | 990 (242) | | 873 (321) | |
| Serious delay | 917 (284) | | 921 (293) | | 871 (271) | |
| **ER status** | | | | | | |
| Negative | 933 (360) | 0.92 | 996 (343) | 0.93 | 863 (365) | 0.69 |
| Positive | 935 (318) | | 1019 (278) | | 869 (342) | |
| **PR status** | | | | | | |
| Negative | 946 (332) | 0.61 | 990 (308) | 0.88 | 872 (402) | 0.96 |

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ |
| Positive | 916 (327) | | 1030 (297) | | 861 (338) | |
| **HER2 status** | | | | | | |
| Negative | 936 (349) | 0.97 | 1026 (342) | 0.39 | 850 (337) | 0.50 |
| Positive | 927 (291) | | 976 (224) | | 874 (356) | |
| **Ki-67 (%)** | | | | | | |
| < 20% | 934 (328) | 0.87 | 1006 (379) | 0.67 | 873 (300) | 0.78 |
| ≥ 20% | 936 (314) | | 1005 (290) | | 863 (396) | |
| **Breast cancer subtypes** | | | | | | |
| Luminal/HER2-negative | 913 (336) | 0.87 | 1026 (317) | 0.76 | 850 (333) | 0.89 |
| Luminal/HER2-positive | 934 (282) | | 995 (199) | | 865 (388) | |
| HER2 enriched | 920 (339) | | 976 (314) | | 888 (325) | |
| Triple-negative/basal-like | 953 (364) | | 1023 (283) | | 860 (427) | |
| **TNM cancer stage** | | | | | | |
| I | 880 (354) | 0.04 | 1033 (310) | 0.70 | 838 (326) | 0.09 |
| II | 914 (309) | | 1010 (303) | | 866 (332) | |
| III-IV | 980 (331) | | 978 (303) | | 1187 (404) | |

[1] p-value for Wilcoxon and Kruskal-Wallis rank sum tests.

**Table 20:** Comparison of Shannon index by demographic characteristics and clinical factors

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ |
| **Age at diagnosis** | | | | | | |
| < 40 | 4.11 (0.88) | 0.27 | 4.37 (0.82) | 0.81 | 4.06 (0.72) | 0.19 |
| 40-49 | 3.92 (0.96) | | 4.16 (0.75) | | 3.67 (0.96) | |
| 50-59 | 3.92 (0.89) | | 4.09 (0.84) | | 3.65 (0.72) | |
| 60+ | 3.91 (0.79) | | 4.12 (0.63) | | 3.74 (0.82) | |
| **Education** | | | | | | |
| Primary school | 4.07 (0.92) | 0.13 | 4.14 (0.66) | 0.73 | 3.67 (0.91) | 0.23 |
| Middle school | 3.88 (0.87) | | 4.10 (0.88) | | 3.63 (0.83) | |
| High school | 4.11 (0.85) | | 4.22 (0.82) | | 4.02 (0.92) | |
| College or higher | 3.81 (0.79) | | 3.94 (0.56) | | 3.74 (0.81) | |
| **Income** | | | | | | |
| Low (T1) | 4.06 (0.86) | 0.18 | 4.1 (0.77) | 0.99 | 3.94 (0.81) | 0.31 |
| Middle (T2) | 3.91 (0.97) | | 4.09 (0.97) | | 3.74 (1.05) | |
| High (T3) | 3.84 (0.91) | | 4.25 (0.70) | | 3.67 (0.74) | |
| **Location** | | | | | | |
| Urban area | 3.92 (0.94) | 0.68 | 4.15 (0.86) | 0.82 | 3.74 (0.93) | 0.74 |
| Rural area | 3.97 (0.88) | | 4.09 (0.82) | | 3.74 (0.77) | |
| **Menopausal status** | | | | | | |
| Premenopausal | 3.93 (0.93) | 0.72 | 4.14 (0.85) | 0.73 | 3.75 (0.89) | 0.76 |
| Postmenopausal | 3.94 (0.83) | | 4.1 (0.81) | | 3.74 (0.76) | |
| **Family history of breast cancer** | | | | | | |
| No | 3.94 (0.90) | 0.84 | 4.12 (0.83) | 0.51 | 3.74 (0.85) | 0.28 |
| Yes | 3.98 (1.04) | | 4.27 (0.49) | | 3.45 (0.95) | |
| **BMI levels** | | | | | | |
| Under weight (<18.5) | 3.93 (1.03) | 0.84 | 3.92 (1.64) | 0.42 | 4.03 (0.76) | 0.96 |
| Normal weight (18.5-22.9) | 3.99 (0.87) | | 4.21 (0.77) | | 3.74 (0.77) | |
| Overweight (23.0-24.9) | 3.83 (0.96) | | 3.92 (0.67) | | 3.66 (1.09) | |
| Obese (≥ 25) | 3.88 (0.91) | | 4.08 (0.64) | | 3.74 (1.01) | |
| **Comorbidity** | | | | | | |
| No | 3.97 (0.91) | 0.25 | 4.11 (0.83) | 0.90 | 3.76 (0.87) | 0.25 |
| Yes | 3.89 (0.73) | | 4.12 (0.56) | | 3.64 (0.60) | |
| **Diagnosis delay** | | | | | | |
| No delay | 3.98 (0.99) | 0.23 | 4.34 (0.93) | 0.0005 | 3.73 (0.97) | 0.85 |
| Moderate delay | 3.97 (0.76) | | 4.09 (0.76) | | 3.75 (0.78) | |
| Serious delay | 3.80 (0.67) | | 3.82 (0.76) | | 3.67 (0.67) | |
| **ER status** | | | | | | |
| Negative | 3.91 (0.90) | 0.84 | 4.09 (0.98) | 0.56 | 3.75 (1.02) | 0.76 |
| Positive | 3.97 (0.87) | | 4.14 (0.73) | | 3.74 (0.84) | |
| **PR status** | | | | | | |
| Negative | 3.93 (0.90) | 0.84 | 4.09 (0.89) | 0.39 | 3.77 (0.99) | 0.93 |
| Positive | 3.97 (0.88) | | 4.14 (0.73) | | 3.73 (0.83) | |
| **HER2 status** | | | | | | |

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ |
| Negative | 3.96 (0.86) | 0.99 | 4.10 (0.86) | 0.41 | 3.67 (0.86) | 0.44 |
| Positive | 3.94 (0.93) | | 4.13 (0.76) | | 3.77 (0.93) | |
| **Ki-67 (%)** | | | | | | |
| < 20% | 3.91 (0.84) | 0.62 | 4.13 (1.03) | 0.81 | 3.74 (0.77) | 0.97 |
| ≥ 20% | 3.97 (0.93) | | 4.11 (0.76) | | 3.74 (1.02) | |
| **Breast cancer subtypes** | | | | | | |
| Luminal/HER2-negative | 3.97 (0.87) | 0.89 | 4.13 (0.80) | 0.81 | 3.67 (0.85) | 0.65 |
| Luminal/HER2-positive | 3.92 (0.89) | | 4.16 (0.66) | | 3.73 (0.81) | |
| HER2 enriched | 3.94 (0.96) | | 4.11 (0.95) | | 3.93 (0.97) | |
| Triple-negative/basal-like | 3.91 (0.85) | | 4.08 (0.92) | | 3.70 (1.06) | |
| **TNM cancer stage** | | | | | | |
| I | 3.77 (0.86) | 0.05 | 4.13 (0.91) | 0.65 | 3.66 (0.84) | 0.10 |
| II | 3.92 (0.88) | | 4.14 (0.78) | | 3.74 (0.86) | |
| III-IV | 4.09 (0.90) | | 4.07 (0.88) | | 4.48 (0.98) | |

[1] p-value for Wilcoxon and Kruskal-Wallis rank sum tests.

**Table 21**: Comparison of inverse Simpson index by demographic characteristics and clinical factors

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ | Median (IQR) | $P^1$ |
| **Age at diagnosis** | | | | | | |
| < 40 | 17.4 (16.6) | 0.51 | 22.8 (15.5) | 0.88 | 15.5 (13.0) | 0.51 |
| 40-49 | 15.1 (15.7) | | 18.9 (21.2) | | 13.9 (11.9) | |
| 50-59 | 15.4 (14.5) | | 17.3 (16.9) | | 13.9 (12.0) | |
| 60+ | 16.6 (12.8) | | 20.1 (13.8) | | 15.2 (10.1) | |
| **Education** | | | | | | |
| Primary school | 17.4 (17.1) | 0.18 | 19.4 (23.2) | 0.61 | 12.2 (14.9) | 0.40 |
| Middle school | 15.3 (15.8) | | 17.0 (19.1) | | 14.1 (11.5) | |
| High school | 17.5 (16.7) | | 23.3 (16.7) | | 15.5 (16.1) | |
| College or higher | 14.4 (11.7) | | 17.0 (11.6) | | 13.5 (11.1) | |
| **Income** | | | | | | |
| Low (T1) | 16.5 (16.6) | 0.85 | 16.8 (19) | 0.47 | 15.8 (14.0) | 0.51 |
| Middle (T2) | 15.5 (16.2) | | 18.4 (24.5) | | 13.6 (13.1) | |
| High (T3) | 15.4 (13.8) | | 22.2 (14.9) | | 14.4 (10.1) | |
| **Location** | | | | | | |
| Urban area | 15.5 (15.1) | 0.78 | 20.2 (16.2) | 0.47 | 14.7 (10.9) | 0.63 |
| Rural area | 16.3 (15.7) | | 17.9 (19.1) | | 14.5 (13.3) | |
| **Menopausal status** | | | | | | |
| Premenopausal | 15.3 (15.8) | 0.53 | 17.9 (18.4) | 0.49 | 14.0 (14.1) | 0.96 |
| Postmenopausal | 16.5 (14.1) | | 19.2 (20.7) | | 15.1 (10.8) | |
| **Family history of breast cancer** | | | | | | |
| No | 15.8 (15.2) | 0.66 | 18.6 (18.9) | 0.79 | 14.6 (12.0) | 0.34 |
| Yes | 14.4 (17.2) | | 20.7 (13.6) | | 11.0 (11.4) | |
| **BMI levels** | | | | | | |
| Under weight (<18.5) | 16.5 (16.5) | 0.97 | 13.9 (28.2) | 0.76 | 18.4 (11.0) | 0.88 |
| Normal weight (18.5-22.9) | 15.8 (15.1) | | 20.1 (18.3) | | 14.3 (11.2) | |
| Overweight (23.0-24.9) | 15.3 (17.9) | | 17.2 (11.2) | | 13.2 (20.9) | |
| Obese (≥ 25) | 15.7 (10.2) | | 15.6 (16.5) | | 15.7 (9.4) | |
| **Comorbidity** | | | | | | |
| No | 16.0 (16.7) | 0.57 | 18.6 (18.9) | 0.97 | 14.5 (13.2) | 0.56 |
| Yes | 15.5 (10.2) | | 18.0 (11.7) | | 14.5 (7.8) | |
| **Diagnosis delay** | | | | | | |
| No delay | 17.0 (18.7) | 0.05 | 24.3 (22.1) | 0.0002 | 14.9 (13.4) | 0.83 |
| Moderate delay | 15.6 (14.5) | | 17.9 (16.6) | | 14.6 (11.9) | |
| Serious delay | 14.5 (10.4) | | 15.3 (11.0) | | 14.4 (9.0) | |
| **ER status** | | | | | | |
| Negative | 15.4 (18.7) | 0.60 | 16.7 (22.9) | 0.52 | 14.8 (12.7) | 0.60 |
| Positive | 16.1 (13.7) | | 19.5 (16.2) | | 14.3 (12.0) | |
| **PR status** | | | | | | |
| Negative | 15.5 (17.7) | 0.83 | 16.7 (19.4) | 0.26 | 15.0 (14.6) | 0.80 |
| Positive | 15.8 (14.2) | | 19.5 (16.1) | | 14.1 (11.1) | |
| **HER2 status** | | | | | | |

| | Overall | | Stool collection time | | | |
|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | |
| | (N=356) | | (N=162) | | (N= 194) | |
| | Median (IQR) | *P*[1] | Median (IQR) | *P*[1] | Median (IQR) | *P*[1] |
| Negative | 16.1 (15.5) | 0.96 | 18.9 (17.8) | 0.39 | 14.3 (13.2) | 0.40 |
| Positive | 15.5 (15.1) | | 17.3 (17.2) | | 14.6 (10.9) | |
| **Ki-67 (%)** | | | | | | |
| < 20% | 15.5 (14.6) | 0.92 | 18.9 (25.3) | 0.46 | 14.8 (12.2) | 0.83 |
| ≥ 20% | 16.1 (16.6) | | 18.4 (16.3) | | 14.2 (11.8) | |
| **Breast cancer subtypes** | | | | | | |
| Luminal/HER2-negative | 16.6 (14.8) | 0.51 | 20.2 (16.3) | 0.56 | 14.9 (12.2) | 0.40 |
| Luminal/HER2-positive | 15.1 (13.5) | | 19.0 (13.6) | | 13.7 (10.7) | |
| HER2 enriched | 16.5 (18.3) | | 16.5 (22.6) | | 15.7 (18.8) | |
| Triple-negative/basal-like | 15.1 (17.3) | | 17.1 (20.0) | | 12.1 (11.8) | |
| **TNM cancer stage** | | | | | | |
| I | 14.1 (11.7) | 0.20 | 18.8 (21.9) | 0.65 | 13.7 (10.0) | 0.07 |
| II | 15.6 (15.2) | | 19.6 (16.8) | | 14.2 (12.1) | |
| III-IV | 18.9 (19.1) | | 17.7 (20.4) | | 24.0 (19.8) | |

[1] p-value for Wilcoxon and Kruskal-Wallis rank sum tests.

Multivariable linear regression analyses show that mean alpha diversity indexes were lower for patients whose stool samples were collected after breast cancer surgery than those collected before breast cancer surgery and neoadjuvant chemotherapy. In the overall analysis and analysis for patients who collected stool before breast cancer surgery and neoadjuvant chemotherapy, all alpha diversity indexes were significantly decreased among those who experienced moderate and serious delay compared with no-delay groups (adjusted p-value <0.05 for all). In addition, patients aged ≥ 60 had significantly lower Chao1 and Shannon indexes than young patients aged < 40 years old in the analysis overall. In the analysis for patients whose stool samples were collected after breast cancer surgery, breast cancer patients with late-stage (stage III-IV) had significantly higher alpha diversity indexes than patients with stage I (adjusted p-value <0.05 for all) (**Table 22-25**).

**Table 22**: Association of selected demographic characteristics and clinical factors with alpha diversity indexes

| | Chao1 richness index | | Shannon-Wiener diversity index† | | Inverse Simpson index‡ | |
|---|---|---|---|---|---|---|
| | N=356 | | N=356 | | N=356 | |
| | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* |
| **Stool collection time** | | | | | | |
| Before BC surgery & neoadjuvant chemo | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| After BC surgery | -108.5 (-161.5, -55.4) | -107.3 (-170.3, -44.3) | -2.35 (-3.44, -1.27) | -2.46 (-3.75, -1.17) | -0.49 (-0.81, -0.17) | -0.52 (-0.90, -0.14) |
| **Age at diagnosis** | | | | | | |
| < 40 | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| 40-49 | -88.3 (-172.8, -3.9) | -96.9 (-183.3, -10.5) | -1.47 (-3.21, 0.27) | -1.66 (-3.44, 0.11) | -0.25 (-0.76, 0.26) | -0.28 (-0.81, 0.24) |
| 50-59 | -63.9 (-150.3, 22.6) | -106.8 (-217.0, 3.4) | -1.11 (-2.89, 0.67) | -2.13 (-4.39, 0.13) | -0.24 (-0.76, 0.29) | -0.52 (-1.19, 0.15) |
| 60+ | -117.0 (-217.7, -16.4) | -170.6 (-307, -34.3) | -1.60 (-3.68, 0.47) | -2.81 (-5.61, -0.02) | -0.20 (-0.81, 0.41) | -0.52 (-1.35, 0.31) |
| **Income** | | | | | | |
| Low (T1) | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Middle (T2) | -46.0 (-110.5, 18.6) | -43.4 (-108.7, 21.8) | -0.80 (-2.12, 0.53) | -0.78 (-2.12, 0.56) | 0.26 (-0.34, 0.86) | -0.07 (-0.47, 0.32) |
| High (T3) | -72.4 (-137.5, -7.3) | -62.2 (-129.8, 5.3) | -1.31 (-2.65, 0.03) | -1.08 (-2.47, 0.30) | 0.24 (-0.41, 0.88) | -0.18 (-0.59, 0.23) |
| **Location** | | | | | | |
| Urban area | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Rural area | -0.9 (-57.2, 55.4) | -29.0 (-88.2, 30.1) | 0.14 (-1.02, 1.29) | -0.35 (-1.57, 0.86) | -0.09 (-0.67, 0.49) | -0.11 (-0.47, 0.25) |
| **Menopausal status** | | | | | | |
| Premenopausal | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Postmenopausal | 4.5 (-49.3, 58.3) | 52.0 (-34.8, 138.8) | 0.27 (-0.83, 1.38) | 1.14 (-0.64, 2.92) | 0.24 (-0.27, 0.75) | 0.34 (-0.18, 0.87) |
| **BMI levels** | | | | | | |
| Normal weight | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Under weight | 57.1 (-38.8, 152.9) | 77.8 (-19.7, 175.3) | 1.06 (-0.90, 3.03) | 1.29 (-0.71, 3.29) | 0.13 (-0.44, 0.71) | 0.14 (-0.45, 0.73) |
| Overweight | 30.3 (-79.9, 140.4) | 49.7 (-63.1, 162.5) | 0.64 (-1.62, 2.90) | 0.72 (-1.59, 3.03) | 0.09 (-0.57, 0.76) | 0.01 (-0.68, 0.7) |
| Obese | 52.4 (-75.6, 180.5) | 91.8 (-41.5, 225.0) | 1.26 (-1.37, 3.89) | 1.56 (-1.17, 4.29) | 0.27 (-0.50, 1.04) | 0.25 (-0.56, 1.06) |
| **Comorbidity** | | | | | | |
| No | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Yes | -44.6 (-114.6, 25.5) | -38.1 (-117.7, 41.5) | -0.61 (-2.06, 0.83) | -0.73 (-2.36, 0.90) | -0.13 (-0.55, 0.29) | -0.23 (-0.71, 0.26) |
| **Diagnosis delay** | | | | | | |
| No delay | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Moderate delay | -6.2 (-67.2, 54.8) | -37.6 (-99.9, 24.8) | -0.18 (-1.43, 1.06) | -0.81 (-2.09, 0.47) | -0.16 (-0.52, 0.20) | -0.28 (-0.66, 0.10) |
| Serious delay | -40.5 (-114.3, 33.2) | -80.0 (-155.5, -4.4) | -1.30 (-2.81, 0.21) | -2.11 (-3.66, -0.56) | -0.53 (-0.97, -0.09) | -0.70 (-1.16, -0.24) |
| **Breast cancer subtypes** | | | | | | |
| Luminal/HER2-negative | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| Luminal/HER2-positive | 0.3 (-69.2, 69.8) | -6.8 (-77.0, 63.3) | -0.22 (-1.65, 1.20) | -0.36 (-1.79, 1.08) | -0.13 (-0.54, 0.29) | -0.17 (-0.6, 0.26) |
| HER2 enriched | 29.8 (-53.7, 113.3) | 4.2 (-70.1, 78.5) | 0.38 (-1.12, 1.87) | -0.01 (-1.53, 1.52) | 0.11 (-0.32, 0.55) | 0.02 (-0.43, 0.47) |
| Triple-negative | 0.3 (-69.2, 69.8) | 3.5 (-83.2, 90.1) | -0.21 (-1.92, 1.51) | -0.84 (-2.61, 0.94) | -0.18 (-0.68, 0.32) | -0.35 (-0.87, 0.18) |
| **TNM cancer stage** | | | | | | |
| I | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 |
| II | 10.1 (-58.5, 78.7) | -12.4 (-84.0, 59.1) | 0.27 (-1.14, 1.68) | -0.10 (-1.56, 1.37) | -0.25 (-1.21, 0.70) | 0.08 (-0.36, 0.51) |
| III-IV | 76.9 (-3.0, 156.8) | 7.5 (-83.9, 98.9) | 1.54 (-0.10, 3.18) | 0.20 (-1.67, 2.08) | -0.49 (-1.44, 0.46) | 0.19 (-0.37, 0.74) |

† Shannon index was transformed to the square of Shannon index; ‡ Inverse Simpson index was transformed to the square root of inverse Simpson index.

[a] Multivariable Model 1: multivariable linear regression model with adjustment for fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[b] Multivariable model 2 was the multivariable model 1 with additional adjustment for stool collection time (before breast cancer surgery & neoadjuvant chemotherapy vs. vs after breast cancer surgery), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
Abbreviation: BC: breast cancer.

**Table 23**: Association of selected demographic characteristics and clinical factors with Chao1 index by stool collection time

| | Stool collection time | | | |
|---|---|---|---|---|
| | Collecting before breast cancer surgery & neoadjuvant chemotherapy (N=162) | | Collecting after breast cancer surgery (N=194) | |
| | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* | Model 1[a] *β (95%CI)* | Model 2[c] *β (95%CI)* |
| **Age at diagnosis** | | | | |
| < 40 | 0.0 | 0.0 | 0.0 | 0.0 |
| 40-49 | -31.0 (-155.2, 93.2) | -62.5 (-194.7, 69.7) | -131.3 (-243.5, -19.1) | -146.1 (-263.7, -28.6) |
| 50-59 | -13.2 (-138.5, 112.2) | -74.3 (-241.7, 93.1) | -112.5 (-228.3, 3.4) | -145.2 (-295.2, 4.8) |
| 60+ | -90.5 (-236.8, 55.8) | -177 (-387.1, 33.2) | -143.6 (-279.0, -8.2) | -149.4 (-338.3, 39.5) |
| **Income** | | | | |
| Low (T1) | 0.0 | 0.0 | 0.0 | 0.0 |
| Middle (T2) | -21.3 (-112.4, 69.8) | -34.5 (-127.2, 58.2) | -51.7 (-143.0, 39.5) | -32.2 (-128.4, 64.0) |
| High (T3) | -21 (-119.0, 77.0) | -23.0 (-125.8, 79.8) | -83.1 (-173.1, 6.9) | -92.1 (-185.7, 1.5) |
| **Location** | | | | |
| Urban area | 0.0 | 0.0 | 0.0 | 0.0 |
| Rural area | -23.0 (-110.9, 65.0) | -25.8 (-116.8, 65.3) | -23.1 (-97.4, 51.2) | -41.8 (-121.9, 38.2) |
| **Menopausal status** | | | | |
| Premenopausal | 0.0 | 0.0 | 0.0 | 0.0 |
| Postmenopausal | 6.5 (-70.9, 83.9) | 66.4 (-59.5, 192.4) | -16.7 (-90.3, 56.8) | 34.5 (-88.4, 157.4) |
| **BMI levels** | | | | |
| Normal weight | 0.0 | 0.0 | 0.0 | 0.0 |
| Under weight | 108.0 (-26.2, 242.1) | 146.4 (5.4, 287.4) | 31.7 (-102.7, 166.2) | -17.8 (-158.0, 122.4) |
| Overweight | 66.7 (-87.7, 221.1) | 95.3 (-69.8, 260.3) | 9.5 (-143.6, 162.5) | -46.1 (-209.6, 117.4) |
| Obese | 112.6 (-61.1, 286.4) | 191.1 (-3.0, 385.1) | -17.7 (-200.9, 165.5) | -68.7 (-264.5, 127.2) |
| **Comorbidity** | | | | |
| No | 0.0 | 0.0 | 0.0 | 0.0 |
| Yes | -22.0 (-128.6, 84.7) | -51.9 (-176.6, 72.8) | -53.1 (-143.9, 37.7) | -39.6 (-149.7, 70.5) |
| **Diagnosis delay** | | | | |
| No delay | 0.0 | 0.0 | 0.0 | 0.0 |
| Moderate delay | -91.0 (-179.0, -3.0) | -120.8 (-219.7, -22) | 14.9 (-67.4, 97.2) | 12.5 (-73.5, 98.4) |
| Serious delay | -185.0 (-282.7, -87.4) | -199.8 (-304.1, -95.5) | 57.1 (-53.2, 167.4) | 81.6 (-33.9, 197.0) |
| **Breast cancer subtypes** | | | | |
| Luminal/HER2-negative | 0.0 | 0.0 | 0.0 | 0.0 |
| Luminal/HER2-positive | -19.5 (-130.3, 91.3) | -17.7 (-134.3, 99.0) | 6.8 (-81.1, 94.7) | 8.4 (-83.8, 100.6) |
| HER2 enriched | 38.7 (-77.4, 154.7) | 58.2 (-67.2, 183.5) | -39.4 (-160.9, 82.0) | -75.1 (-209.9, 59.7) |
| Triple-negative/basal-like | -19.5 (-130.3, 91.3) | -17.7 (-134.3, 99.0) | 6.8 (-81.1, 94.7) | 8.4 (-83.8, 100.6) |
| **TNM cancer stage** | | | | |
| I | 0.0 | 0.0 | 0.0 | 0.0 |
| II | -62.4 (-207.9, 83.0) | -51.1 (-199.2, 97.1) | -4.3 (-82.2, 73.5) | 1.5 (-84.3, 87.3) |
| III-IV | -83.7 (-228.4, 61.1) | -42.3 (-189.7, 105.1) | 170.1 (18.2, 322.0) | 186.4 (23.9, 348.8) |

[a] Multivariable Model 1: multivariable linear regression model with adjustment for fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[b] Multivariable model 2 was the multivariable model 1 with additional adjustment for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[c] Multivariable model 2 was the multivariable model 1 with additional adjustment for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Table 24**: Association of selected demographic characteristics and clinical factors with Shannon index by stool collection time

| | Stool collection time† | | | |
|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162) | | Collection after breast cancer surgery (N=194) | |
| | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* | Model 1[a] *β (95%CI)* | Model 2[c] *β (95%CI)* |
| **Age at diagnosis** | | | | |
| < 40 | 0.00 | 0.00 | 0.00 | 0.00 |
| 40-49 | -0.16 (-2.81, 2.48) | -0.98 (-3.79, 1.83) | -2.45 (-4.68, -0.22) | -2.76 (-5.09, -0.44) |
| 50-59 | -0.15 (-2.81, 2.52) | -1.83 (-5.39, 1.73) | -2.02 (-4.32, 0.28) | -2.52 (-5.48, 0.45) |
| 60+ | -0.82 (-3.94, 2.29) | -3.03 (-7.50, 1.44) | -2.33 (-5.02, 0.36) | -2.06 (-5.80, 1.68) |
| **Income** | | | | |
| Low (T1) | 0.00 | 0.00 | 0.00 | 0.00 |
| Middle (T2) | -0.03 (-1.97, 1.90) | -0.43 (-2.40, 1.54) | -1.18 (-2.99, 0.63) | -0.82 (-2.72, 1.09) |
| High (T3) | 0.05 (-2.03, 2.13) | 0.10 (-2.08, 2.29) | -1.80 (-3.58, -0.02) | -1.96 (-3.81, -0.11) |
| **Location** | | | | |
| Urban area | 0.00 | 0.00 | 0.00 | 0.00 |
| Rural area | -0.15 (-2.02, 1.71) | 0.11 (-1.82, 2.05) | -0.43 (-1.91, 1.04) | -0.83 (-2.41, 0.75) |
| **Menopausal status** | | | | |
| Premenopausal | 0.00 | 0.00 | 0.00 | 0.00 |
| Postmenopausal | 0.44 (-1.20, 2.08) | 1.75 (-0.93, 4.43) | -0.28 (-1.74, 1.17) | 0.35 (-2.08, 2.78) |
| **BMI levels** | | | | |
| Normal weight | 0.00 | 0.00 | 0.00 | 0.00 |
| Under weight | 2.60 (-0.23, 5.44) | 2.97 (-0.03, 5.97) | 0.17 (-2.50, 2.84) | -0.91 (-3.68, 1.87) |
| Asian overweight | 1.65 (-1.62, 4.91) | 1.43 (-2.08, 4.94) | 0.02 (-3.01, 3.06) | -1.21 (-4.45, 2.02) |
| Asian obese | 2.48 (-1.19, 6.16) | 3.08 (-1.04, 7.21) | -0.10 (-3.74, 3.54) | -1.26 (-5.14, 2.61) |
| **Comorbidity** | | | | |
| No | 0.00 | 0.00 | 0.00 | 0.00 |
| Yes | 0.34 (-1.92, 2.60) | -0.70 (-3.35, 1.95) | -1.13 (-2.92, 0.67) | -1.13 (-3.30, 1.05) |
| **Diagnosis delay** | | | | |
| No delay | 0.00 | 0.00 | 0.00 | 0.00 |
| Moderate delay | -1.89 (-3.73, -0.04) | -2.19 (-4.29, -0.09) | 0.15 (-1.48, 1.79) | 0.03 (-1.67, 1.73) |
| Serious delay | -4.27 (-6.32, -2.22) | -4.38 (-6.60, -2.16) | 0.62 (-1.57, 2.81) | 1.10 (-1.19, 3.38) |
| **Breast cancer subtypes** | | | | |
| Luminal/HER2-negative | 0.00 | 0.00 | 0.00 | 0.00 |
| Luminal/HER2-positive | -0.95 (-3.30, 1.40) | -0.90 (-3.38, 1.58) | 0.07 (-1.67, 1.80) | 0.07 (-1.75, 1.89) |
| HER2 enriched | -0.31 (-2.77, 2.15) | -0.09 (-2.75, 2.58) | 0.80 (-1.22, 2.83) | -2.24 (-4.91, 0.43) |
| Triple-negative/basal-like | -0.95 (-3.30, 1.40) | -0.90 (-3.38, 1.58) | -1.49 (-3.88, 0.91) | 0.07 (-1.75, 1.89) |
| **TNM cancer stage** | | | | |
| I | 0.00 | 0.00 | 0.00 | 0.00 |
| II | -0.88 (-3.96, 2.20) | -0.60 (-3.75, 2.55) | -0.17 (-1.71, 1.37) | 0.11 (-1.59, 1.81) |
| III-IV | -1.59 (-4.65, 1.48) | -0.70 (-3.84, 2.43) | 3.29 (0.28, 6.30) | 3.73 (0.52, 6.95) |

† Shannon index was transformed to the square of Shannon index

[a] Multivariable Model 1: multivariable linear regression model with adjustment for fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[b] Multivariable model 2 was the multivariable model 1 with additional adjustment for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[c] Multivariable model 2 was the multivariable model 1 with additional adjustment for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Table 25**: Association of selected demographic characteristics and clinical factors with inverse Simpson index by stool collection time

| | Stool collection time‡ | | | |
|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162) | | Collection after breast cancer surgery (N=194) | |
| | Model 1[a] *β (95%CI)* | Model 2[b] *β (95%CI)* | Model 1[a] *β (95%CI)* | Model 2[c] *β (95%CI)* |
| **Age at diagnosis** | | | | |
| < 40 | 0.00 | 0.00 | 0.00 | 0.00 |
| 40-49 | 0.16 (-0.66, 0.98) | -0.09 (-0.96, 0.79) | -0.55 (-1.18, 0.08) | -0.57 (-1.22, 0.08) |
| 50-59 | 0.04 (-0.79, 0.86) | -0.56 (-1.67, 0.55) | -0.47 (-1.12, 0.18) | -0.48 (-1.32, 0.35) |
| 60+ | 0.12 (-0.85, 1.09) | -0.62 (-2.01, 0.77) | -0.46 (-1.22, 0.30) | -0.22 (-1.28, 0.83) |
| **Income** | | | | |
| Low (T1) | 0.00 | 0.00 | 0.00 | 0.00 |
| Middle (T2) | -0.03 (-1.97, 1.90) | 0.11 (-0.51, 0.72) | -0.29 (-0.8, 0.21) | -0.20 (-0.74, 0.33) |
| High (T3) | 0.05 (-2.03, 2.13) | 0.20 (-0.48, 0.88) | -0.47 (-0.97, 0.03) | -0.50 (-1.02, 0.02) |
| **Location** | | | | |
| Urban area | 0.00 | 0.00 | 0.00 | 0.00 |
| Rural area | -0.15 (-2.02, 1.71) | 0.03 (-0.57, 0.63) | -0.11 (-0.53, 0.30) | -0.21 (-0.65, 0.24) |
| **Menopausal status** | | | | |
| Premenopausal | 0.00 | 0.00 | 0.00 | 0.00 |
| Postmenopausal | 0.44 (-1.20, 2.08) | 0.63 (-0.21, 1.46) | -0.08 (-0.49, 0.33) | 0.004 (-0.68, 0.69) |
| **BMI levels** | | | | |
| Normal weight | 0.00 | 0.00 | 0.00 | 0.00 |
| Under weight | 0.43 (-0.45, 1.32) | 0.47 (-0.46, 1.41) | 0.03 (-0.72, 0.78) | -0.29 (-1.07, 0.49) |
| Asian overweight | 0.26 (-0.76, 1.28) | 0.05 (-1.04, 1.15) | 0.05 (-0.80, 0.91) | -0.34 (-1.25, 0.57) |
| Asian obese | 0.56 (-0.59, 1.71) | 0.65 (-0.63, 1.93) | -0.04 (-1.06, 0.98) | -0.44 (-1.53, 0.65) |
| **Comorbidity** | | | | |
| No | 0.00 | 0.00 | 0.00 | 0.00 |
| Yes | 0.13 (-0.57, 0.83) | -0.31 (-1.13, 0.52) | -0.28 (-0.79, 0.22) | -0.31 (-0.92, 0.30) |
| **Diagnosis delay** | | | | |
| No delay | 0.00 | 0.00 | 0.00 | 0.00 |
| Moderate delay | -0.77 (-1.34, -0.20) | -0.81 (-1.46, -0.15) | 0.09 (-0.37, 0.54) | 0.05 (-0.43, 0.52) |
| Serious delay | -1.33 (-1.96, -0.70) | -1.34 (-2.03, -0.65) | -0.002 (-0.62, 0.61) | 0.14 (-0.50, 0.79) |
| **Breast cancer subtypes** | | | | |
| Luminal/HER2-negative | 0.00 | 0.00 | 0.00 | 0.00 |
| Luminal/HER2-positive | -0.48 (-1.21, 0.25) | -0.41 (-1.18, 0.36) | 0.03 (-0.46, 0.52) | 0.02 (-0.49, 0.54) |
| HER2 enriched | -0.27 (-1.03, 0.49) | -0.21 (-1.04, 0.62) | -0.41 (-1.08, 0.26) | -0.64 (-1.39, 0.11) |
| Triple negative/basal-like | -0.48 (-1.21, 0.25) | -0.41 (-1.18, 0.36) | 0.03 (-0.46, 0.52) | 0.02 (-0.49, 0.54) |
| **TNM cancer stage** | | | | |
| I | 0.00 | 0.00 | 0.00 | 0.00 |
| II | -0.36 (-1.43, 0.70) | -0.12 (-1.10, 0.86) | 0.01 (-0.42, 0.44) | 0.11 (-0.37, 0.58) |
| III-IV | -0.59 (-1.64, 0.45) | -0.15 (-1.12, 0.83) | 1.09 (0.25, 1.93) | 1.22 (0.32, 2.13) |

‡ Inverse Simpson index was transformed to the square root of inverse Simpson index.

[a] Multivariable Model 1: multivariable linear regression model with adjustment for fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[b] Multivariable model 2 was the multivariable model 1 with additional adjustment for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[c] Multivariable model 2 was the multivariable model 1 with additional adjustment for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Figure 9** shows principal coordinate analysis (PCoA) of beta diversity (measured by Bray-Curtis dissimilarity matrix) by stool collection time. A significant difference in the Bray-Curtis dissimilarity matrix was also found between two groups defined by stool collection time, with p-value of 0.001 for PERMANOVA test with 999 permutations. Likewise, significant differences were consistently observed for the weighted UniFracs distance and unweighted UniFracs distance matrixes (p-vale of 0.001 for both). Stool collection time explained 2.6% to 3.2% variations in beta-diversity in the analysis overall (**Table 26**)

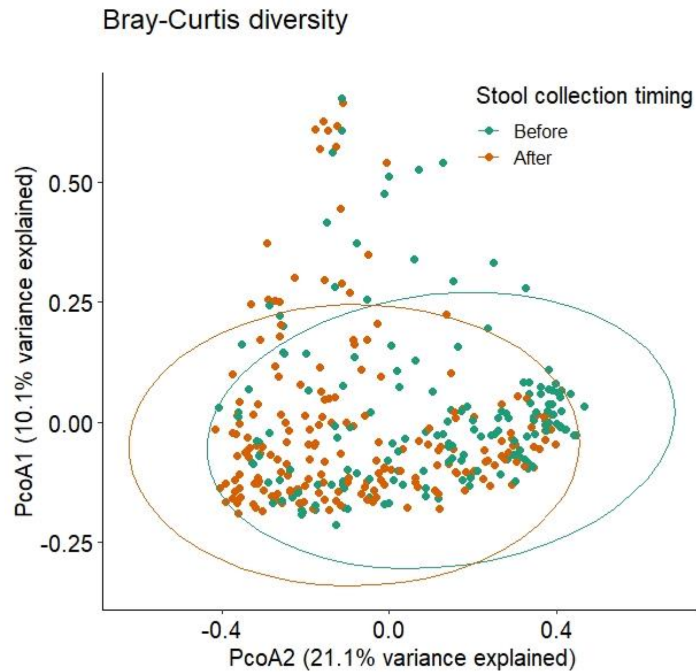



**Figure 9:** Principal coordinate analysis (PCoA) of Bray-Curtis diversity by stool collection time

(‘*Before*’: collected stool samples before breast cancer surgery and chemotherapy (n=162); ‘*After*’: collected stool samples after breast cancer surgery (n=194)

Besides stool collection time, PERMANOVA analysis showed that income and diagnosis delay were significant factors for explaining the variation of beta diversity in 356 participants’ stool samples. The combined analysis and analyses stratified by stool collection time showed that income levels (tertile distribution) explained 1.0% to 1.8% of variations in beta-diversity, 2.0% to 3.3% of variations among patients who collected stool before breast cancer surgery, and neoadjuvant chemotherapy, and 1.2% to 1.6% among those collected stool after breast cancer surgery in the model 2 (all p for *PERMANOVA* <0.05). In addition, diagnosis delay significantly explained 1.8% to 2.5% of variations in beta-diversity among patients who collected stool before breast cancer surgery and neoadjuvant chemotherapy (**Table 26**).

**Table 26**: PERMANOVA test difference of the beta diversity between demographic characteristic and clinical factors

| | Bray-Curtis dissimilarity matrix | | Unweighted UniFracs distance matrix | | Weighted UniFracs distance matrix | |
|---|---|---|---|---|---|---|
| | $R^2$(*p for PERMANOVA test*) | | $R^2$(*p for PERMANOVA test*) | | $R^2$(*p for PERMANOVA test*) | |
| | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| **Overall (N=356)** [a] | | | | | | |
| Stool collection time | 3.2% (0.001) | 3.2% (0.001) | 2.6% (0.001) | 2.6% (0.001) | 2.8% (0.001) | 2.8% (0.001) |
| Age groups at diagnosis | 0.9% (0.19) | 1.0% (0.18) | 1.0% (0.13) | 1.0% (0.13) | 1.0% (0.16) | 1.0% (0.14) |
| Income levels | 2.4% (0.001) | 1.8% (0.001) | 1.3% (0.001) | 1.0% (0.004) | 2.4% (0.001) | 1.7% (0.001) |
| Residence | 0.7% (0.007) | 0.3% (0.15) | 0.5% (0.038) | 0.4% (0.07) | 0.6% (0.02) | 0.3% (0.40) |
| Menopausal status | 0.3% (0.32) | 0.2% (0.69) | 0.2% (0.63) | 0.2% (0.78) | 0.3% (0.30) | 0.2% (0.67) |
| BMI levels | 1.0% (0.23) | 1.0% (0.15) | 0.9% (0.31) | 0.9% (0.28) | 1.0% (0.20) | 0.9% (0.21) |
| Comorbidity | 0.5% (0.05) | 0.3% (0.18) | 0.5% (0.02) | 0.4% (0.10) | 0.4% (0.12) | 0.3% (0.26) |
| Diagnosis delay | 0.5% (0.51) | 0.6% (0.32) | 0.5% (0.62) | 0.6% (0.24) | 0.5% (0.57) | 0.6% (0.29) |
| Breast cancer subtypes | 1.0% (0.13) | 0.8% (0.40) | 0.9% (0.28) | 0.8% (0.43) | 1.0% (0.19) | 0.8% (0.38) |
| TNM cancer stags | 1.1% (0.004) | 0.4% (0.75) | 1.0% (0.02) | 0.4% (0.81) | 1.0% (0.014) | 0.4% (0.75) |
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)** [b] | | | | | | |
| Age groups at diagnosis | 2.1% (0.20) | 2.2% (0.18) | 2.1% (0.23) | 2.1% (0.21) | 2.3% (0.16) | 2.2% (0.13) |
| Income levels | 3.5% (0.001) | 3.3% (0.001) | 2.1% (0.025) | 2.0% (0.03) | 3.4% (0.001) | 3.3% (0.001) |
| Residence | 1.4% (0.011) | 1.0% (0.07) | 0.8% (0.14) | 0.7% (0.25) | 1.5% (0.13) | 1.0% (0.06) |
| Menopausal status | 0.5% (0.68) | 0.5% (0.64) | 0.5% (0.84) | 0.4% (0.87) | 0.5% (0.62) | 0.5% (0.69) |
| BMI levels | 2.2% (0.16) | 2.2% (0.11) | 2.3% (0.13) | 2.3% (0.09) | 2.5% (0.09) | 2.5% (0.07) |
| Comorbidity | 0.7% (0.23) | 0.7% (0.25) | 0.6% (0.37) | 0.6% (0.44) | 0.6% (0.53) | 0.7% (0.27) |
| Diagnosis delay | 1.5% (0.17) | 1.8% (0.05) | 2.1% (0.012) | 2.5% (0.003) | 1.8% (0.10) | 2.0% (0.03) |
| Breast cancer subtypes | 1.5% (0.79) | 1.4% (0.91) | 1.6% (0.74) | 1.6% (0.73) | 1.7% (0.57) | 1.5% (0.76) |
| TNM cancer stags | 1.0% (0.70) | 1.0% (0.77) | 1.0% (0.77) | 0.9% (0.90) | 1.1% (0.63) | 0.9% (0.80) |
| **Stool collection after breast cancer surgery (N=194)** [c] | | | | | | |
| Age groups at diagnosis | 1.4% (0.70) | 1.4% (0.66) | 1.6% (0.32) | 1.6% (0.32) | 1.4% (0.58) | 1.5% (0.52) |
| Income levels | 1.6% (0.02) | 1.6% (0.03) | 1.1% (0.25) | 1.2% (0.18) | 1.6% (0.03) | 1.6% (0.04) |
| Residence | 0.5% (0.50) | 0.5% (0.34) | 0.6% (0.17) | 0.7% (0.09) | 0.4% (0.59) | 0.6% (0.33) |
| Menopausal status | 0.6% (0.32) | 0.4% (0.75) | 0.4% (0.69) | 0.4% (0.81) | 0.6% (0.22) | 0.4% (0.64) |
| BMI levels | 1.4% (0.73) | 1.3% (0.73) | 1.2% (0.84) | 1.3% (0.88) | 1.2% (0.84) | 1.1% (0.92) |
| Comorbidity | 0.7% (0.13) | 0.6% (0.29) | 0.7% (0.14) | 0.5% (0.37) | 0.6% (0.20) | 0.4% (0.62) |
| Diagnosis delay | 1.2% (0.20) | 1.1% (0.40) | 1.0% (0.58) | 0.9% (0.74) | 1.3% (0.16) | 1.2% (0.29) |
| Breast cancer subtypes | 1.6% (0.38) | 1.6% (0.41) | 1.5% (0.55) | 1.4% (0.61) | 1.6% (0.45) | 0.8% (0.59) |
| TNM cancer stags | 1.1% (0.31) | 0.8% (0.76) | 1.1% (0.35) | 1.0% (0.44) | 1.0% (0.44) | 0.7% (0.83) |

Model 1: was adjusted for fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[a] Model 2 was the model 1 with additional adjustment for stool collection time (before breast cancer surgery & neoadjuvant chemotherapy vs. vs after breast cancer surgery), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, and TNM cancer stage.

[b] Model 2 was the model 1 with additional adjustment for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, and TNM cancer stage.

[c] Model 2 was the model 1 with additional adjustment for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, and TNM cancer stage.

**Association of GI microbial taxa with sociodemographic and clinic factors**

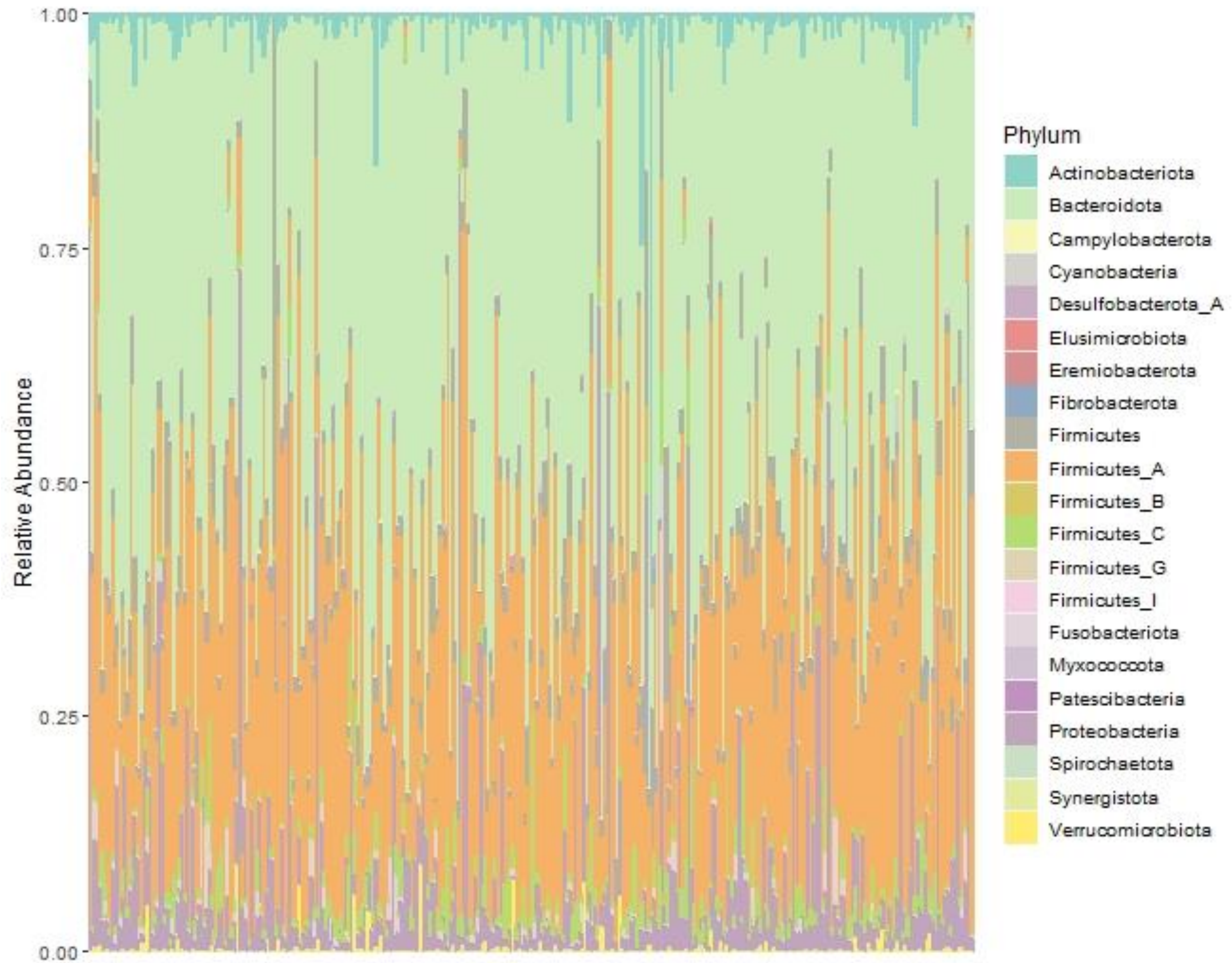



**Figure 10:** Relative abundance of the gut microbiome at phylum levels

Among 356 stool samples, a total of 21 phyla, 29 classes, 77 orders, 244 families, 1278 genera, and 4206 species was identified and estimated absolute abundances. At the phylum level, the GI microbiota had high proportions (i.e., relative abundance) of *Bacteroidota, Proteobacteria* and *Firmicutes* with a dominance of *Firmicutes A* and smaller proportion of other phyla including *Actinobacteriota, Campylobacterota, Cyanobacteria, Desulfobacterota A, Elusimicrobiota, Eremiobacterota, Fibrobacterota, Fusobacteriota, Myxococcota, Patescibacteria, Spirochaetota, Synergistota* and *Verrucomicrobiota* **(Figure 10)**. Compared with patients whose stool samples were collected after breast cancer surgery, patients whose stool samples were collected before breast cancer surgery and chemotherapy showed a higher relative abundance of phyla *Firmicutes A, Firmicutes B, Cyanobacteria, Spirochaetota* but a lower relative abundance of phyla *Bacteroidota,* and *Firmicutes C* at the phylum level (all p-values of Wilcoxon rank sum test < 0.05) (**Figure 11**).

**Figure 11:** Boxplot of relative abundance at phylum level by stool collection time

('*Before*': collected stool samples before breast cancer surgery and neoadjuvant chemotherapy (n=162); '*After*': collected stool samples after breast cancer surgery (n=194))

After excluding rare taxa with a prevalence < 10% of 356 samples, a total of 17 phyla, 23 classes, 52 orders, 137 families, 646 genera, and 1989 species were included to evaluate the associations of GI microbial taxa with sociodemographic and clinic factors.

**Table 27**: Association of stool collection time with GI microbial taxa at phylum to order levels

| ***Taxonomy*** | Relative abundance, median (%) | Prevalence (%) | *β (SE)* for after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy[1] | P-value | $P_{FDR}$[2] |
|---|---|---|---|---|---|
| Phylum *Actinobacteriota* | | | | | |
| Class *Actinobacteria* | 0.075 | 97.2 | 0.935 (0.280) | $9.32 \times 10^{-4}$ | 0.006 |
| Order *Actinomycetales* | 0.075 | 97.2 | 0.921 (0.295) | $1.94 \times 10^{-3}$ | 0.010 |
| Order *Mycobacteriales* | 0.004 | 19.9 | 1.075 (0.366) | $3.34 \times 10^{-3}$ | 0.057 |
| Class Coriobacteriia | 0.361 | 99.7 | -0.447 (0.111) | $6.92 \times 10^{-5}$ | 0.001 |
| Order *Coriobacteriales* | 0.361 | 99.7 | -0.456 (0.115) | $9.18 \times 10^{-5}$ | 0.003 |
| Phylum Cyanobacteria | 0.040 | 98.9 | -0.303 (0.108) | $5.30 \times 10^{-3}$ | 0.019 |
| Class Vampirovibrionia | 0.040 | 98.9 | -0.283 (0.115) | 0.015 | 0.041 |
| Order *Gastranaerophilales* | 0.040 | 98.9 | -0.293 (0.113) | 0.010 | 0.036 |
| Phylum *Firmicutes* | 1.722 | 100.0 | 0.252 (0.123) | 0.041 | 0.098 |
| Class *Bacilli* | 1.722 | 100.0 | 0.272 (0.119) | 0.023 | 0.053 |
| Order *Bacillales* | 0.018 | 75.6 | -0.818 (0.269) | $2.50 \times 10^{-3}$ | 0.011 |
| Order *Staphylococcales* | 0.001 | 50.6 | 0.503 (0.179) | $5.21 \times 10^{-3}$ | 0.020 |
| Phylum *Firmicutes A* | | | | | |
| Class *Clostridia* | | | | | |
| Order *Clostridiales* | 0.034 | 97.2 | -0.549 (0.26) | 0.036 | 0.089 |
| Order *Eubacteriales* | 0.003 | 65.4 | 0.825 (0.217) | $1.67 \times 10^{-4}$ | 0.003 |
| Order *Lachnospirales* | 14.503 | 100.0 | -0.301 (0.083) | $3.07 \times 10^{-4}$ | 0.003 |
| Order *Monoglobales* | 0.021 | 93.5 | -0.348 (0.161) | 0.032 | 0.086 |
| Order *Peptostreptococcales* | 0.288 | 100.0 | 0.350 (0.111) | $1.72 \times 10^{-3}$ | 0.010 |
| Order *Tissierellales* | 0.017 | 93.3 | 0.689 (0.191) | $3.57 \times 10^{-4}$ | 0.003 |
| Phylum *Firmicutes B* | 0.006 | 86.8 | -0.410 (0.149) | $6.19 \times 10^{-3}$ | 0.019 |
| Class *Peptococcia* | 0.006 | 84.0 | -0.489 (0.153) | $1.57 \times 10^{-3}$ | 0.006 |
| Order *Peptococcales* | 0.006 | 84.0 | -0.499 (0.153) | $1.27 \times 10^{-3}$ | 0.009 |
| Phylum Firmicutes C | 1.165 | 100.0 | 0.402 (0.135) | $3.02 \times 10^{-3}$ | 0.018 |
| Class *Negativicutes* | 1.165 | 100.0 | 0.423 (0.133) | $1.59 \times 10^{-3}$ | 0.006 |
| Phylum *Spirochaetota* | 0.002 | 59.3 | -0.464 (0.147) | $1.72 \times 10^{-3}$ | 0.018 |
| Class *Spirochaetia* | 0.001 | 54.5 | -0.298 (0.141) | 0.036 | 0.063 |
| Order *Treponematales* | 0.001 | 51.4 | -0.315 (0.145) | 0.031 | 0.086 |
| Phylum *Verrucomicrobiota* | | | | | |
| Class *Lentisphaeria* | 0.006 | 80.6 | -0.481 (0.228) | 0.036 | 0.063 |
| Order *Victivallales* | 0.006 | 80.6 | -0.491 (0.225) | 0.029 | 0.086 |
| Order *Opitutales*† | 0.004[¥] | 17.4 | -0.952 (0.352) | $6.90 \times 10^{-3}$ | 0.059 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
† Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for taxa abundance.
[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
[¥] Median relative abundance for rare taxa was calculated among carriers

**Table 27** presents significant associations ($P_{FDR}$ < 0.1) of stool collection time with the abundances of 28 taxa at phylum to order levels. Among these 28 taxa, most of them were defined as common taxa, except for the order *Opitutales*. The stool collection after breast cancer surgery was inversely associated with three phyla, five classes, and ten orders and was positively associated with two phyla, two classes, and six orders. In addition, there were 42 families, 239 genera, and 827 species that had significant differences between groups of stool collection time in linear regression analysis after FDR correction ($P_{FDR}$ < 0.1). Compared with the stools collected before breast cancer surgery and neoadjuvant chemotherapy, the stool collected after breast cancer surgery was significantly associated with lower abundances of 23 families, 136 genera, and 497 species, and with higher abundances of 19 families, 103 genera, and 375 species. Generally, the relative abundance of approximately 40% of investigated taxa differed significantly by stool collection time.

No association between TNM cancer stages with GI microbial taxa was found in the analysis for all participants. Among patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, stage II breast cancer patients have a significant lower proportion of carriers of phylum *Synergistota* ($\beta$= -1.678; $P_{FDR}$=*0.095),* comparing to stage I breast cancer patients. The proportion of carriers of family *MGYG-HGUT-04147* in the phylum *Firmicutes A* was significantly higher in stage III-IV breast cancer patients than in stage I patients ($\beta$= 3.351; $P_{FDR}$=*0.062)* among patients whose stool samples were collected after breast cancer surgery (**Table 28**).

In terms of breast cancer subtype, ten common taxa showed a significantly difference in relative abundance between luminal/HER2-positive and luminal/HER2-negative breast cancer in the overall analysis. In the phylum *Firmicutes A*, eight taxa, including two families *MGYG-HGUT-03214, Ruminococcaceae,* and 5 *MGYG-HGUT* species (i.e., *03409, 02224, 03323, 03675*, and *04336)* were more abundant, while only species *MGYG-HGUT-04359* showed a lower abundance among patients with luminal/HER2-positive breast cancer (FDR corrected p-values <0.1). In the phylum *Bacteroidota,* the abundance of species *MGYG-HGUT-01038* was significantly higher in luminal/HER2-positive than luminal/HER2-negative breast cancer patients ($\beta$= 0.479; $P_{FDR}$=*0.080)*. In addition, the proportion of carriers of family *Atopobiaceae,* belonging to the phylum *Actinobacteriota,* was lower in luminal/HER2-positive breast cancer among patients whose stool samples were collected after breast cancer surgery. Furthermore, *Campylobacterota* was significantly less prevalent among

patients with HER2 enriched subtype ($\beta$= -1.238; $P_{FDR}$=0.072) than luminal/HER2-negative breast cancer*.* The genus *Acutalibacter*, belonging to the phylum *Firmicutes A,* showed a higher abundance among patients with HER2 enriched breast cancer in the analysis overall ($\beta$= 0.60; $P_{FDR}$=*0.085)* and in those with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy only analysis ($\beta$= 0.877; $P_{FDR}$=*0.076).* Finally, no association was found between triple-negative breast cancer and GI microbial taxa (**Table 29**).

**Table 28**: Association of TNM cancer stage with GI microbial taxa by stool collection time

| Taxonomy | RA, median (%) | Pre (%) | β (SE) for stage II vs stage I | P | $P_{FDR}$[3] | β (SE) for stage III-IV vs. stage I | P | $P_{FDR}$[3] |
|---|---|---|---|---|---|---|---|---|
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)** [1] | | | | | | | | |
| Phylum *Synergistota* ┼ | 0.003¥ | 45.2 | -1.678 (0.715) | 0.019 | 0.095 | -1.307 (0.713) | 0.067 | 0.333 |
| **Stool collection after breast cancer surgery (N=194)** [2] | | | | | | | | |
| Phylum *Firmicutes A* | | | | | | | | |
| Family *MGYG-HGUT-04147* ┼ | 0.002¥ | 11.2 | 0.372 (0.752) | 0.621 | 0.848 | 3.351 (1.011) | 0.001 | 0.062 |

┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for taxa abundance. ¥ Median relative abundance for rare taxa was calculated among carriers.
[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, patient delay, and TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 29**: Association of breast cancer subtypes with GI microbial taxa by stool collection time

| Taxonomy | RA, median (%) | Pre (%) | *β (SE) for* Luminal/HER2-positive *vs.* Luminal/HER2-negative | *P* | $P_{FDR}$[4] | *β (SE) for* HER2 enriched *vs.* Luminal/HER2-negative | *P* | $P_{FDR}$[4] | *β (SE) for* Triple-negative *vs.* Luminal/HER2-negative | *P* | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Overall (N=356)** [1] | | | | | | | | | | | |
| Phylum *Bacteroidota* | | | | | | | | | | | |
| Species *MGYG-HGUT-01038* | 0.004 | 84.0 | 0.479 (0.138) | $6.14 \times 10^{-4}$ | 0.080 | 0.283 (0.147) | 0.054 | 0.992 | 0.450 (0.171) | 0.009 | 0.363 |
| Phylum *Campylobacterota* ┼ | 0.004¥ | 17.1 | -0.357 (0.383) | 0.351 | 0.585 | -1.238 (0.505) | 0.014 | 0.072 | -0.095 (0.467) | 0.839 | 0.870 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Species *MGYG-HGUT-04359* | 0.002 | 67.4 | -0.756 (0.188) | $7.32 \times 10^{-5}$ | 0.055 | -0.552 (0.200) | 0.006 | 0.832 | -0.622 (0.233) | 0.008 | 0.363 |
| Family *MGYG-HGUT-03214* | 0.035 | 93.3 | 0.509 (0.162) | 0.002 | 0.092 | 0.007 (0.171) | 0.966 | 0.987 | 0.223 (0.200) | 0.266 | 0.699 |
| Species *MGYG-HGUT-03214* | 0.035 | 93.3 | 0.498 (0.150) | 0.001 | 0.098 | -0.009 (0.159) | 0.954 | 0.998 | 0.213 (0.186) | 0.254 | 0.727 |
| Genus *Acutalibacter* | 0.004 | 81.5 | 0.349 (0.155) | 0.025 | 0.461 | 0.600 (0.164) | $3.07 \times 10^{-4}$ | 0.085 | 0.258 (0.192) | 0.179 | 0.632 |
| Species *MGYG-HGUT-03409* | 0.006 | 79.5 | 0.560 (0.171) | 0.001 | 0.100 | 0.320 (0.181) | 0.079 | 0.992 | 0.505 (0.212) | 0.017 | 0.363 |
| Family *Ruminococcaceae* | 5.933 | 100 | 0.465 (0.154) | 0.003 | 0.092 | -0.064 (0.163) | 0.694 | 0.987 | -0.001 (0.190) | 0.995 | 0.995 |
| Species *MGYG-HGUT-02224* | 0.008 | 86.0 | 0.549 (0.160) | $6.46 \times 10^{-4}$ | 0.080 | -0.046 (0.169) | 0.784 | 0.992 | 0.189 (0.197) | 0.339 | 0.751 |
| Species *MGYG-HGUT-03323* | 0.008 | 87.1 | 0.501 (0.140) | $4.11 \times 10^{-4}$ | 0.080 | -0.047 (0.149) | 0.755 | 0.992 | 0.214 (0.174) | 0.218 | 0.719 |
| Species *MGYG-HGUT-03675* | 0.013 | 86.8 | 0.555 (0.157) | $4.56 \times 10^{-4}$ | 0.080 | -0.018 (0.166) | 0.914 | 0.996 | 0.214 (0.194) | 0.270 | 0.735 |
| Species *MGYG-HGUT-04336* | 0.063 | 89.0 | 0.839 (0.246) | $7.44 \times 10^{-4}$ | 0.080 | -0.087 (0.261) | 0.740 | 0.992 | 0.243 (0.305) | 0.426 | 0.796 |
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)** [2] | | | | | | | | | | | |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Genus *Acutalibacter* | 0.004 | 81.5 | 0.569 (0.255) | 0.027 | 0.465 | 0.877 (0.235) | $2.75 \times 10^{-4}$ | 0.076 | 0.886 (0.274) | 0.002 | 0.281 |
| **Stool collection after breast cancer surgery (N=194)** [3] | | | | | | | | | | | |
| Phylum *Actinobacteriota* | | | | | | | | | | | |
| Family *Atopobiaceae* ┼ | 0.004¥ | 43.8 | -1.460 (0.434) | 0.001 | 0.051 | -1.446 (0.536) | 0.007 | 0.234 | -1.142 (0.613) | 0.063 | 0.876 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa.
¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, and TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$<0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 30:** Association of ER status with GI microbial taxa by stool collection time

| *Taxonomy* | RA, median (%) | Pre (%) | Before breast cancer surgery & neoadjuvant chemotherapy (N=162) [2] | | | After breast cancer surgery (N= 194) [3] | | | Overall (N=356)[1] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *β (SE) for ER positive vs ER negative* | *P* | $P_{FDR}$[4] | *β (SE) for ER positive vs ER negative* | *P* | $P_{FDR}$[4] | *β (SE) for ER positive vs ER negative* | *P* | $P_{FDR}$[4] |
| Phylum *Elusimicrobiota* ┼ | 0.002 | 16.9[¥] | 0.119 (0.843) | 0.888 | 0.888 | 1.58 (0.687) | 0.022 | 0.108 | 1.047 (0.448) | 0.02 | 0.098 |

┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa.

¥ Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, HER2 status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, HER2 status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, HER2 status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 31**: Association of HER2 status with GI microbial taxa by stool collection time

| Taxonomy | RA, median (%) | Pre (%) | Before breast cancer surgery & neoadjuvant chemotherapy (N=162) [2] β (SE) for HER2-positive vs negative | P | $P_{FDR}$[4] | After breast cancer surgery (N= 194) [3] β (SE) for HER2-positive vs negative | P | $P_{FDR}$[4] | Overall (N=356)[1] β (SE) for HER2-positive vs negative | P | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phylum *Campylobacterota* ┼ | 0.005 | 17.1¥ | -0.635 (0.540) | 0.239 | 0.512 | -0.656 (0.446) | 0.141 | 0.176 | -0.625 (0.320) | 0.050 | 0.085 |
| Phylum *Elusimicrobiota* ┼ | 0.005 | 12.6¥ | 0.464 (0.562) | 0.409 | 0.512 | -1.612 (0.745) | 0.030 | 0.076 | -0.42 (0.369) | 0.256 | 0.320 |
| Phylum *Firmicutes I* ┼ | 0.003 | 27.0¥ | 0.213 (0.442) | 0.629 | 0.629 | 0.866 (0.381) | 0.023 | 0.076 | 0.598 (0.269) | 0.026 | 0.066 |
| Phylum *Synergistota* ┼ | 0.003 | 45.2¥ | 0.537 (0.386) | 0.164 | 0.512 | 0.614 (0.344) | 0.074 | 0.123 | 0.544 (0.237) | 0.022 | 0.066 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa.
¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, ER status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, ER status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, ER status, PR status, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

Regrading to hormone receptor status, we only found a higher proportion of carriers of phylum *Elusimicrobiota* in breast cancer patients with ER+ than ER- ($\beta$= 1.047; $P_{FDR}$=*0.098)* (**Table 30**). **Table 31** presents four phyla with a significantly differential proportion of carriers between HER2-positive and HER2-negative breast cancer patients. The proportion of carriers of phylum *Firmicutes I* ($\beta$= 0.598; $P_{FDR}$=*0.098) and Synergistota* ($\beta$= 0.544; $P_{FDR}$=*0.066) were* higher, while the proportion of carriers of phylum *Campylobacterota* was lower among patients with HER2-positive breast cancer in the overall analysis. Among patients whose stool samples were collected after breast cancer surgery, a significantly higher proportion of carriers of phylum *Firmicutes I* was consistently observed in HER2-positive breast cancer ($\beta$= 0.866; $P_{FDR}$=*0.076)*. We also found that HER2-positive breast cancer was associated with a decreased proportion of carriers of phylum *Elusimicrobiota* ($\beta$= -1.612; $P_{FDR}$=*0.076),* comparing with HER2-negative breast cancer.

Compared with patients aged < 40 years old at diagnosis I, patients aged 60+ had a higher abundance of phylum *Proteobacteria* ($\beta$= 1.167; $P_{FDR}$=*0.044),* which was driven by the class *Gammaproteobacteria* ($\beta$= 1.198; $P_{FDR}$=*0.043).* In addition, the proportion of carriers of three taxa, including the genus *Senegalimassilia* (belonging to the phylum *Actinobacteriota*), the genus *Anaeromassilibacillus* (belonging to the phylum *Firmicutes A*), and the order *Acholeplasmatales* (belonging to the phylum *Firmicutes*), were significantly lower among patients aged 60+ with all FDR-corrected p-value <0.1. Among patients whose stool samples were collected after breast cancer surgery, we also found an increased proportion of carriers of family *Rhodocyclaceae,* belonging to the phylum *Proteobacteria,* in older patients aged 60 and above ($\beta$= 3.362; $P_{FDR}$=*0.034)* (**Table 32**) compared with patients aged < 40 years old.

**Table 33** presents significant associations of high income (tertile 3) with the abundances of 16 taxa at phylum to order levels. Among these 16 taxa, most of them were common taxa, except for the phyla *Elusimicrobiota* and *Synergistota.* The proportions of carriers of the latter two phyla were significantly lower among patients having high income compared with patients having low income (tertile 1) with $\beta$= -1.00 ($P_{FDR}$=*0.082) for Elusimicrobiota* and $\beta$= -0.628 ($P_{FDR}$=*0.082*) for *Synergistota.* In addition, high income was significantly associated with a higher abundance of two phyla, two classes, and six orders, and a lower abundance of order *Bacillales* (belonging to the phylum *Firmicutes)* and the phylum *Spirochaetota* (two taxa of this phylum, the class *Spirochaetia* and the order *Treponematales) (*all $P_{FDR}$ < 0.1). Moreover, 25 families, 86

genera, and 270 species differed significantly between patients with high income and those with low income in the combined analysis after FDR correction ($P_{FDR} < 0.1$). Significantly higher abundance of two orders, *Lachnospirales* and *Tissierellales,* lower abundance of *Treponematales,* and proportion of carriers of phylum *Elusimicrobiota* were consistently observed among patients with high income in patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy. Likewise, a significantly higher abundance of phylum *Firmicutes* and two taxa of this phylum, including the class *Bacilli* and the order *Erysipelotrichales, was observed* among patients with high income in patients whose d stool samples were collected after breast cancer surgery.

In the overall analysis, a total of 34 common taxa showed a significant difference by residence after FDR-correction (all $P_{FDR}<0.1$). The majority of them (26 of 34) belong to the phylum *Firmicutes A*, and the eight remaining taxa belong to four phyla *Actinobacteriota, Bacteroidota, Firmicutes*, and *Fusobacteriota.* In the phylum *Firmicutes A*, patients living in rural areas showed a significantly lower abundance of most taxa, particularly four genera and 15 species within the order *Oscillospirales*. Meanwhile, only four taxa, including the family *Peptostreptococcaceae,* the genus *Terrisporobacter,* the species *MGYG-HGUT-04232*, and the family *Helcococcaceae* (belonging to two orders, *Peptostreptococcales* and *Tissierellales*) were found to be significantly more abundant among patients living in rural areas than urban areas. In the phylum *Bacteroidota,* the species *Bacteroides intestinalis* was less abundant among patients living in rural areas ($\beta$= -0.645; $P_{FDR}=0.083$). In the phylum *Actinobacteriota,* the family *Coriobacteriaceae* showed a higher abundance among patients living in rural areas ($\beta$= 0.403; $P_{FDR}=0.095$), driven by the genus *Collinsella* ($\beta$= 0.463; $P_{FDR}$=0.099). Similarly, in the phylum *Firmicutes,* the family *Streptococcaceae* along with three species of this family, including *Streptococcus salivarius, Streptococcus sp001556435,* and *Streptococcus vestibularis,* were also more abundant among patients living in rural areas (all $P_{FDR}<0.1$). Furthermore, we found a significantly higher proportion of carriers of phylum *Firmicutes I* among patients living in rural areas than those living in urban areas ($\beta$= 0.700; $P_{FDR}$=0.089) (**Table 34**).

**Table 32**: Association of age groups with GI microbial taxa by stool collection time

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for* age 40-49 *vs.* age <40 | *P* | $P_{FDR}$[4] | *β (SE) for* age 50-59 *vs.* age <40 | *P* | $P_{FDR}$[4] | *β (SE) for* age 60+ *vs.* age <40 | *P* | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Overall (N=356)**[1] | | | | | | | | | | | |
| Phylum *Actinobacteriota* | | | | | | | | | | | |
| Order *Mycobacteriales* ┼ | 0.004¥ | 19.9 | -0.66 (0.410) | 0.108 | 0.534 | -1.551 (0.587) | 0.008 | 0.070 | -0.801 (0.723) | 0.268 | 0.700 |
| Order *Propionibacteriales* ┼ | 0.004¥ | 16.6 | -0.64 (0.432) | 0.138 | 0.534 | -1.687 (0.633) | 0.008 | 0.070 | -1.911 (0.849) | 0.024 | 0.207 |
| Genus *Senegalimassilia* ┼ | 0.025¥ | 40.4 | -0.956 (0.372) | 0.010 | 0.524 | -1.330 (0.481) | 0.006 | 0.474 | -2.153 (0.613) | $4.41\times10^{-4}$ | 0.061 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Genus *Anaeromassilibacillus* ┼ | 0.003¥ | 34.3 | -0.884 (0.378) | 0.019 | 0.524 | -1.522 (0.520) | 0.003 | 0.474 | -2.423 (0.651) | $2.00\times10^{-4}$ | 0.055 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Order *Acholeplasmatales* ┼ | 0.003¥ | 22.8 | -0.173 (0.402) | 0.668 | 0.833 | -0.695 (0.540) | 0.198 | 0.550 | -2.139 (0.771) | 0.006 | 0.094 |
| Phylum *Proteobacteria* | 3.367 | 100.0 | 0.268 (0.253) | 0.289 | 0.755 | 0.286 (0.322) | 0.375 | 0.686 | 1.167 (0.399) | 0.004 | 0.044 |
| Class *Gammaproteobacteria* | 3.213 | 100.0 | 0.276 (0.255) | 0.279 | 0.652 | 0.292 (0.325) | 0.370 | 0.760 | 1.198 (0.402) | 0.003 | 0.043 |
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)**[2] | | | | | | | | | | | |
| Phylum *Actinobacteriota* | | | | | | | | | | | |
| Family *Eggerthellaceae* | 0.101 | 99.2 | -0.209 (0.216) | 0.334 | 0.934 | -0.926 (0.273) | 0.001 | 0.064 | -0.416 (0.343) | 0.227 | 0.853 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Family *Anaerotignaceae* | 0.038 | 91.9 | 0.129 (0.370) | 0.729 | 0.955 | 1.444 (0.469) | 0.002 | 0.087 | 1.531 (0.589) | 0.010 | 0.492 |
| **Stool collection after breast cancer surgery (N=194)** [3] | | | | | | | | | | | |
| Phylum *Elusimicrobiota* ┼ | 0.005¥ | 12.6 | -2.355 (0.895) | 0.008 | 0.042 | -2.333 (1.290) | 0.071 | 0.353 | -1.95 (1.477) | 0.187 | 0.672 |
| Class *Elusimicrobia* ┼ | 0.005¥ | 12.6 | -2.355 (0.895) | 0.008 | 0.076 | -2.333 (1.290) | 0.071 | 0.635 | -1.95 (1.477) | 0.187 | 0.926 |
| Order *Elusimicrobiales* ┼ | 0.005¥ | 12.6 | -2.355 (0.895) | 0.008 | 0.072 | -2.333 (1.290) | 0.071 | 0.400 | -1.95 (1.477) | 0.187 | 0.966 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Order *Bacillales A* ┼ | 0.002¥ | 27.5 | -1.887 (0.608) | 0.002 | 0.033 | -0.095 (0.704) | 0.893 | 0.966 | -0.608 (0.986) | 0.537 | 0.992 |
| Family *Planococcaceae* ┼ | 0.002¥ | 27.5 | -1.887 (0.608) | 0.002 | 0.098 | -0.095 (0.704) | 0.893 | 0.952 | -0.608 (0.986) | 0.537 | 0.735 |
| Phylum *Fusobacteriota* | 0.058 | 99.7 | 1.028 (0.406) | 0.012 | 0.147 | 1.409 (0.518) | 0.007 | 0.087 | 1.352 (0.653) | 0.040 | 0.159 |
| Class *Fusobacteriia* | 0.058 | 99.7 | 1.108 (0.411) | 0.008 | 0.109 | 1.434 (0.525) | 0.007 | 0.097 | 1.341 (0.661) | 0.044 | 0.205 |
| Phylum *Proteobacteria* | | | | | | | | | | | |
| Family *Rhodocyclaceae* ┼ | 0.003¥ | 33.7 | 1.960 (0.658) | 0.003 | 0.098 | 2.044 (0.786) | 0.009 | 0.313 | 3.362 (0.967) | 0.001 | 0.034 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance. ┼Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa. ¥ Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 33**: Association of income levels with GI microbial taxa by stool collection time at phylum to order levels

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for tertile 2 vs. tertile 1* | *P* | $P_{FDR}$[3] | *β (SE) for tertile 3 vs. tertile 1* | *P* | $P_{FDR}$[3] |
|---|---|---|---|---|---|---|---|---|
| **Overall (N=356)**[1] | | | | | | | | |
| Phylum *Elusimicrobiota* ┼ | 0.005[¥] | 12.6 | -0.514 (0.393) | 0.191 | 0.478 | -1.000 (0.469) | 0.033 | 0.082 |
| Phylum *Firmicutes A* | | | | | | | | |
| Order *Eubacteriales* | 0.003 | 65.4 | 0.215 (0.225) | 0.339 | 0.687 | 0.579 (0.232) | 0.013 | 0.058 |
| Order *Lachnospirales* | 14.503 | 100.0 | 0.066 (0.086) | 0.440 | 0.787 | 0.270 (0.089) | 0.003 | 0.022 |
| Order *Peptostreptococcales* | 0.288 | 100.0 | 0.409 (0.115) | $4.13\text{x}10^{-4}$ | 0.011 | 0.238 (0.119) | 0.045 | 0.159 |
| Order *Tissierellales* | 0.017 | 93.3 | 0.625 (0.198) | 0.002 | 0.015 | 0.572 (0.205) | 0.006 | 0.039 |
| Phylum *Firmicutes* | 1.722 | 100.0 | 0.012 (0.127) | 0.925 | 0.987 | 0.387 (0.132) | 0.004 | 0.034 |
| Class *Bacilli* | 1.722 | 100.0 | 0.038 (0.123) | 0.758 | 0.829 | 0.389 (0.128) | 0.002 | 0.017 |
| Order *Bacillales* | 0.018 | 75.6 | -0.921 (0.278) | 0.001 | 0.012 | -1.097 (0.288) | $1.68\text{x}10^{-4}$ | 0.003 |
| Order *Erysipelotrichales* | 0.873 | 99.7 | 0.137 (0.124) | 0.268 | 0.625 | 0.537 (0.128) | $3.48\text{x}10^{-5}$ | 0.001 |
| Order *Staphylococcales* | 0.001 | 50.6 | 0.288 (0.185) | 0.121 | 0.424 | 0.505 (0.192) | 0.009 | 0.052 |
| Phylum *Fusobacteriota* | 0.058 | 99.7 | 0.87 (0.239) | $3.15\text{x}10^{-4}$ | 0.004 | 0.656 (0.247) | 0.008 | 0.034 |
| Class *Fusobacteriia* | 0.058 | 99.7 | 0.896 (0.240) | $2.24\text{x}10^{-4}$ | 0.003 | 0.659 (0.249) | 0.008 | 0.039 |
| Order *Fusobacteriales* | 0.058 | 99.7 | 0.859 (0.250) | 0.001 | 0.011 | 0.645 (0.258) | 0.013 | 0.058 |
| Phylum *Spirochaetota* | 0.002 | 59.3 | -0.171 (0.152) | 0.261 | 0.783 | -0.429 (0.157) | 0.007 | 0.034 |
| Class *Spirochaetia* | 0.001 | 54.5 | -0.244 (0.147) | 0.097 | 0.454 | -0.540 (0.152) | $4.26\text{x}10^{-4}$ | 0.006 |
| Order *Treponematales* | 0.001 | 51.4 | -0.261 (0.150) | 0.083 | 0.341 | -0.566 (0.156) | $3.23\text{x}10^{-4}$ | 0.004 |
| Phylum *Synergistota* ┼ | 0.003[¥] | 45.2 | -0.017 (0.275) | 0.950 | 0.950 | -0.628 (0.291) | 0.031 | 0.082 |
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)**[2] | | | | | | | | |
| Phylum *Campylobacterota* ┼ | 0.004[¥] | 17.1 | 1.334 (0.548) | 0.015 | 0.075 | -0.036 (0.719) | 0.961 | 0.961 |
| Phylum *Elusimicrobiota* ┼ | 0.005[¥] | 12.6 | -0.376 (0.563) | 0.505 | 0.545 | -1.684 (0.774) | 0.030 | 0.088 |
| Phylum *Firmicutes A* | | | | | | | | |
| Order *Lachnospirales* | 14.503 | 100.0 | 0.146 (0.111) | 0.192 | 0.420 | 0.383 (0.123) | 0.002 | 0.039 |
| Order *Tissierellales* | 0.017 | 93.3 | 0.992 (0.299) | 0.001 | 0.030 | 0.906 (0.332) | 0.007 | 0.083 |
| Phylum *Firmicutes I* ┼ | 0.003 | 27.0 | -0.638 (0.480) | 0.184 | 0.307 | -1.267 (0.602) | 0.035 | 0.088 |
| Phylum *Firmicutes* | | | | | | | | |
| Order *Bacillales* | 0.018 | 75.6 | -1.225 (0.383) | 0.002 | 0.030 | -0.866 (0.424) | 0.043 | 0.300 |
| Phylum *Fusobacteriota* | 0.058 | 99.7 | 1.183 (0.365) | 0.001 | 0.018 | 0.808 (0.404) | 0.048 | 0.321 |
| Class *Fusobacteriia* | 0.058 | 99.7 | 1.207 (0.364) | 0.001 | 0.016 | 0.795 (0.403) | 0.051 | 0.290 |
| Order *Fusobacteriales* | 0.058 | 99.7 | 1.157 (0.384) | 0.003 | 0.036 | 0.801 (0.426) | 0.062 | 0.300 |
| Phylum *Spirochaetota* | | | | | | | | |
| Order *Treponematales* | 0.001 | 51.4 | -0.441 (0.227) | 0.053 | 0.234 | -0.787 (0.251) | 0.002 | 0.039 |

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for tertile 2 vs. tertile 1* | *P* | $P_{FDR}$[3] | *β (SE) for tertile 3 vs. tertile 1* | *P* | $P_{FDR}$[3] |
|---|---|---|---|---|---|---|---|---|
| **Stool collection after breast cancer surgery (N=194)[3]** | | | | | | | | |
| Phylum *Firmicutes A* | | | | | | | | |
| Order *Peptostreptococcales* | 0.288 | 100.0 | 0.461 (0.184) | 0.013 | 0.464 | 0.495 (0.179) | 0.006 | 0.090 |
| Phylum *Firmicutes* | 1.722 | 100.0 | -0.110 (0.199) | 0.580 | 0.819 | 0.530 (0.193) | 0.007 | 0.081 |
| Class *Bacilli* | 1.722 | 100.0 | -0.067 (0.190) | 0.723 | 0.834 | 0.550 (0.185) | 0.003 | 0.047 |
| Order *Bacillales* | 0.018 | 75.6 | -0.537 (0.426) | 0.210 | 0.890 | -1.119 (0.415) | 0.008 | 0.090 |
| Order *Erysipelotrichales* | 0.873 | 99.7 | 0.083 (0.192) | 0.664 | 0.890 | 0.688 (0.186) | $3.02 \times 10^{-4}$ | 0.011 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance. ┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa. ¥ Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}<0.1$ is considered statistically significant.

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 34**: Association of residence (rural areas vs. urban areas) with GI microbial taxa by stool collection time

| | | | Before breast cancer surgery & neoadjuvant chemotherapy (N=162) [2] | | | After breast cancer surgery (N= 194) [3] | | | Overall (N= 356) [1] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] | *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] | *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] |
| Phylum *Actinobacteriota* | | | | | | | | | | | |
| Family *Coriobacteriaceae* | 0.199 | 99.7 | 0.340 (0.233) | 0.147 | 0.576 | 0.475 (0.186) | 0.012 | 0.386 | 0.403 (0.144) | 0.005 | 0.095 |
| Genus *Collinsella* | 0.169 | 99.7 | 0.322 (0.235) | 0.173 | 0.79 | 0.603 (0.200) | 0.003 | 0.114 | 0.463 (0.150) | 0.002 | 0.099 |
| Family *Eggerthellaceae* | | | | | | | | | | | |
| Genus *MGYG-HGUT-01153* ┼ | 0.003¥ | 39.9 | 0.487 (0.431) | 0.259 | 0.74 | -1.189 (0.386) | 0.002 | 0.082 | -0.375 (0.263) | 0.154 | 0.482 |
| Genus *MGYG-HGUT-04041* ┼ | 0.013¥ | 14.0 | 0.323 (0.586) | 0.582 | 0.843 | -3.756 (1.120) | 0.001 | 0.07 | -0.851 (0.394) | 0.031 | 0.237 |
| Phylum *Bacteroidota* | | | | | | | | | | | |
| Species *Bacteroides intestinalis* | 0.025 | 94.9 | -0.877 (0.332) | 0.009 | 0.523 | -0.436 (0.282) | 0.123 | 0.462 | -0.645 (0.207) | 0.002 | 0.083 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Family *CAG-727* | 0.011 | 81.5 | -0.674 (0.371) | 0.071 | 0.485 | -0.752 (0.328) | 0.023 | 0.401 | -0.650 (0.236) | 0.006 | 0.095 |
| Order *Christensenellales* | | | | | | | | | | | |
| Family *CAG-138* | 0.002 | 61.0 | -0.832 (0.452) | 0.068 | 0.485 | -0.679 (0.330) | 0.041 | 0.424 | -0.705 (0.263) | 0.008 | 0.095 |
| Genus *MGYG-HGUT-04245* ┼ | 0.003¥ | 35.4 | -0.608 (0.416) | 0.144 | 0.655 | -1.347 (0.408) | 0.001 | 0.07 | -0.846 (0.269) | 0.002 | 0.107 |
| Family *UBA1750* ┼ | 0.005¥ | 43.8 | -0.283 (0.426) | 0.506 | 0.869 | -1.199 (0.370) | 0.001 | 0.04 | -0.708 (0.257) | 0.006 | 0.196 |
| Genus *UBA7102* ┼ | 0.005¥ | 43.8 | -0.283 (0.426) | 0.506 | 0.821 | -1.199 (0.370) | 0.001 | 0.07 | -0.708 (0.257) | 0.006 | 0.153 |
| Order *Lachnospirales* | | | | | | | | | | | |
| Genus *KLE1615* | 0.036 | 90.4 | -0.107 (0.319) | 0.737 | 0.942 | -1.049 (0.312) | 0.001 | 0.088 | -0.640 (0.221) | 0.004 | 0.102 |
| Species *Ruminococcus_A_sp000432335* | 0.002 | 78.1 | 0.389 (0.188) | 0.040 | 0.523 | 0.402 (0.184) | 0.03 | 0.279 | 0.409 (0.128) | 0.002 | 0.083 |
| Order *Oscillospirales* | | | | | | | | | | | |
| Genus *CAG-177* | 0.002 | 59.0 | -0.716 (0.523) | 0.173 | 0.79 | -1.212 (0.397) | 0.003 | 0.114 | -0.947 (0.309) | 0.002 | 0.099 |
| Species *CAG-177 sp003538135* | 0.001 | 50.8 | -0.816 (0.447) | 0.070 | 0.523 | -0.881 (0.350) | 0.013 | 0.216 | -0.814 (0.268) | 0.003 | 0.092 |
| Genus *RUG806* ┼ | 0.003¥ | 41.6 | -0.240 (0.420) | 0.567 | 0.841 | -1.237 (0.383) | 0.001 | 0.07 | -0.642 (0.261) | 0.014 | 0.204 |
| Genus *UBA737* | 0.002 | 53.7 | -0.793 (0.392) | 0.045 | 0.698 | -0.734 (0.264) | 0.006 | 0.129 | -0.705 (0.216) | 0.001 | 0.099 |

| Taxonomy | RA, median (%) | Pre (%) | Before breast cancer surgery & neoadjuvant chemotherapy (N=162) [2] β (SE) for rural area vs. urban area | P | $P_{FDR}$[4] | After breast cancer surgery (N= 194) [3] β (SE) for rural area vs. urban area | P | $P_{FDR}$[4] | Overall (N= 356) [1] β (SE) for rural area vs. urban area | P | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Genus *MGYG-HGUT-00468* ┼ | 0.009¥ | 16.9 | 0.477 (0.623) | 0.444 | 0.821 | -2.04 (0.645) | 0.002 | 0.072 | -0.640 (0.340) | 0.06 | 0.335 |
| Genus *MGYG-HGUT-02723* ┼ | 0.006¥ | 20.5 | -0.393 (0.509) | 0.440 | 0.821 | -2.042 (0.601) | 0.001 | 0.07 | -0.861 (0.324) | 0.008 | 0.157 |
| Species *CAG-110_sp000434635* | 0.001 | 51.1 | -0.792 (0.270) | 0.004 | 0.496 | -0.591 (0.204) | 0.004 | 0.129 | -0.659 (0.159) | $4.29x10^{-5}$ | 0.032 |
| Species *MGYG-HGUT-04262* | 0.002 | 57.0 | -0.588 (0.317) | 0.066 | 0.523 | -0.577 (0.223) | 0.010 | 0.197 | -0.567 (0.181) | 0.002 | 0.083 |
| Species *MGYG-HGUT-00741* | 0.001 | 50.0 | -0.463 (0.293) | 0.117 | 0.546 | -0.647 (0.210) | 0.002 | 0.124 | -0.545 (0.169) | 0.001 | 0.083 |
| Species *MGYG-HGUT-02710* | 0.002 | 56.5 | -0.510 (0.320) | 0.113 | 0.546 | -0.84 (0.272) | 0.002 | 0.124 | -0.647 (0.202) | 0.002 | 0.083 |
| Species *CAG-83_sp003539495* | 0.001 | 52.2 | -0.538 (0.348) | 0.124 | 0.546 | -0.788 (0.260) | 0.003 | 0.124 | -0.632 (0.204) | 0.002 | 0.083 |
| Species *MGYG-HGUT-00703* | 0.001 | 50.6 | -0.414 (0.296) | 0.165 | 0.551 | -0.708 (0.212) | 0.001 | 0.124 | -0.533 (0.170) | 0.002 | 0.083 |
| Species *MGYG-HGUT-04463* | 0.002 | 60.4 | -0.484 (0.308) | 0.119 | 0.546 | -0.686 (0.224) | 0.003 | 0.124 | -0.587 (0.177) | 0.001 | 0.083 |
| Species *EO_sp003522105* | 0.000 | 50.0 | -0.486 (0.352) | 0.170 | 0.551 | -0.738 (0.228) | 0.001 | 0.124 | -0.612 (0.194) | 0.002 | 0.083 |
| Species *MGYG-HGUT-04276* | 0.002 | 55.3 | -0.616 (0.337) | 0.069 | 0.523 | -0.846 (0.236) | $4.48x10^{-4}$ | 0.124 | -0.716 (0.193) | $2.39x10^{-4}$ | 0.075 |
| Species *MGYG-HGUT-01166* | 0.003 | 63.5 | -0.333 (0.239) | 0.166 | 0.551 | -0.554 (0.221) | 0.013 | 0.217 | -0.488 (0.160) | 0.002 | 0.092 |
| Genus *MGYG-HGUT-04580* | 0.001 | 52.0 | -0.376 (0.199) | 0.061 | 0.699 | -0.391 (0.177) | 0.029 | 0.234 | -0.399 (0.129) | 0.002 | 0.099 |
| Species *MGYG-HGUT-04580* | 0.001 | 52.0 | -0.428 (0.206) | 0.040 | 0.523 | -0.382 (0.180) | 0.035 | 0.298 | -0.416 (0.132) | 0.002 | 0.083 |
| Species *MGYG-HGUT-00834* | 0.002 | 56.2 | -0.324 (0.300) | 0.283 | 0.666 | -0.762 (0.221) | 0.001 | 0.124 | -0.547 (0.175) | 0.002 | 0.083 |
| Species *Oscillibacter sp000436875* | 0.001 | 54.5 | 0.326 (0.184) | 0.078 | 0.523 | 0.333 (0.150) | 0.028 | 0.278 | 0.357 (0.115) | 0.002 | 0.083 |
| Genus *UBA866* | 0.000 | 50.0 | -0.636 (0.276) | 0.023 | 0.586 | -0.558 (0.214) | 0.01 | 0.148 | -0.595 (0.164) | $3.29x10^{-4}$ | 0.091 |
| Species *MGYG-HGUT-01597* | 0.000 | 50.0 | -0.690 (0.282) | 0.016 | 0.523 | -0.550 (0.218) | 0.013 | 0.216 | -0.613 (0.168) | $2.98x10^{-4}$ | 0.075 |
| Order *Peptostreptococcales* | | | | | | | | | | | |
| Genus *Mogibacterium* | 0.003 | 71.6 | 0.059 (0.316) | 0.853 | 0.988 | -1.026 (0.261) | $1.26x10^{-4}$ | 0.017 | -0.511 (0.197) | 0.010 | 0.129 |
| Family *Peptostreptococcaceae* | 0.121 | 97.8 | 0.211 (0.314) | 0.504 | 0.705 | 0.876 (0.301) | 0.004 | 0.287 | 0.595 (0.211) | 0.005 | 0.095 |
| Genus *Terrisporobacter* | 0.002 | 58.4 | -0.158 (0.334) | 0.637 | 0.92 | 1.388 (0.346) | $9.00x10^{-5}$ | 0.017 | 0.851 (0.248) | 0.001 | 0.093 |
| Species *MGYG-HGUT-04232* | 0.002 | 53.4 | -0.079 (0.308) | 0.799 | 0.941 | 1.119 (0.298) | 0 | 0.124 | 0.683 (0.219) | 0.002 | 0.083 |
| Order *Tissierellales* | | | | | | | | | | | |
| Family *Helcococcaceae* | 0.001 | 53.4 | 0.923 (0.246) | $2.86x10^{-4}$ | 0.018 | 0.371 (0.272) | 0.175 | 0.557 | 0.525 (0.184) | 0.005 | 0.095 |
| Phylum *Firmicutes I* ┼ | 0.003¥ | 27.0 | 0.722 (0.509) | 0.156 | 0.3 | 0.772 (0.403) | 0.056 | 0.262 | 0.700 (0.295) | 0.018 | 0.089 |

| *Taxonomy* | RA, median (%) | Pre (%) | Before breast cancer surgery & neoadjuvant chemotherapy (N=162)[2]: *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] | After breast cancer surgery (N= 194)[3]: *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] | Overall (N= 356)[1]: *β (SE) for rural area vs. urban area* | *P* | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phylum *Firmicutes* | | | | | | | | | | | |
| Family *Streptococcaceae* | 0.083 | 99.7 | 0.501 (0.281) | 0.076 | 0.485 | 0.502 (0.260) | 0.055 | 0.424 | 0.495 (0.186) | 0.008 | 0.095 |
| Species *Streptococcus salivarius* | 0.011 | 86.0 | 0.710 (0.350) | 0.045 | 0.523 | 0.803 (0.371) | 0.032 | 0.287 | 0.864 (0.254) | 0.001 | 0.083 |
| Species *Streptococcus sp001556435* | 0.004 | 80.3 | 0.529 (0.311) | 0.091 | 0.523 | 0.767 (0.323) | 0.019 | 0.238 | 0.732 (0.222) | 0.001 | 0.083 |
| Species *Streptococcus vestibularis* | 0.005 | 79.8 | 0.726 (0.327) | 0.028 | 0.523 | 0.661 (0.346) | 0.058 | 0.325 | 0.753 (0.236) | 0.002 | 0.083 |
| Phylum *Fusobacteriota* | | | | | | | | | | | |
| Genus *Leptotrichia* | 0.001 | 55.9 | 0.555 (0.246) | 0.025 | 0.586 | 0.380 (0.207) | 0.067 | 0.331 | 0.466 (0.153) | 0.003 | 0.099 |
| Phylum *Verrucomicrobiota* | | | | | | | | | | | |
| Family *UBA1829* ┼ | 0.007[¥] | 36.2 | 0.077 (0.408) | 0.851 | 0.99 | -1.406 (0.419) | 0.001 | 0.040 | -0.602 (0.263) | 0.022 | 0.298 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa.
¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 35** shows significant associations (*$P_{FDR}$<0.1)* of BMI levels with GI microbial taxa by stool sample collection time. In the overall analysis, 18 common taxa showed significantly different abundances between underweight and normal-weight patients. Four taxa belonging to the phylum *Firmicutes A,* including the order *Lachnospirales,* family *Lachnospiraceae*, order *TANB77*, and family *CAG-508*, showed a lower abundance among underweight patients. Other four taxa, in which two taxa belong to the phylum *Firmicutes C,* including the order *Acidaminococcales* and family *Acidaminococcaceae,* order *Peptococcales* (belonging to the phylum *Firmicutes B),* and order *RF39* (belonging to the phylum *Firmicutes),* were also found to be less abundant among underweight patients compared to patients with normal weight. Conversely, ten common taxa were significantly more abundant among underweight patients with all *$P_{FDR}$<0.1* than those with normal weight. Among them, we found seven taxa belong to three orders, including three orders *Haloplasmatales*, *Lactobacillales,* Staphylococcales, and their four families, *Turicibacteraceae, Aerococcaceae, Enterococcaceae,* and *Lactobacillaceae.* Among the remaining three taxa, we found a higher abundance of two belonging to the phylum *Firmicutes C*, including the order *Veillonellales* and family *Veillonellaceae,* and one taxon belongs to the phylum *Actinobacteriota,* the family *Actinomycetaceae* among underweight patients. The analysis included all participants found no significant association between overweight and obese with GI microbial taxa. However, the genus *Agathobaculum,* belonging to the phylum *Firmicutes A,* was less abundant among obese breast cancer patients in patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy ($\beta$= -1.143; $P_{FDR}$=0.047*).*

Relative abundances of six taxa, including three orders, *Acidaminococcales, Lactobacillales, Staphylococcales*, and three families *Acidaminococcaceae, Aerococcaceae, and Enterococcaceae,* among underweight patients, differed significantly from those with normal weight in patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy. In the analysis among patients whose stool samples were collected after breast cancer surgery, we found a decreased abundance of phylum *Cyanobacteria* and two taxa of this phylum, including the class *Vampirovibrionia* and the order *Gastranaerophilales* among underweight patients *(*all *$P_{FDR}$<0.1),* driven by the family *Gastranaerophilaceae* ($\beta$= -0.797; $P_{FDR}$=0.083*) (***Table 35***).*

Significant associations between diagnosis delay and GI microbial taxa are shown in **Table 36.** No association between moderate and serious delay with gut microbial taxa was found in the analysis that included all participants and the analysis for patients whose stool samples were collected after breast cancer surgery. We found that 61 taxa, most of which were rare taxa, showed significantly different proportions of carriers among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy (all $P_{FDR} < 0.1$). Significantly decreased proportions of carriers were observed for 57 taxa; 55 of them are members of the class *Clostridia* belonging to the phylum *Firmicutes A*, including six families, 48 genera, and the species *Roseburia hominis* (i.e., member of the family *Lachnospiraceae*). In addition, the proportions of carriers of two taxa, including the phylum *Elusimicrobiota* and family *CAG-433* (belonging to the phylum *Firmicutes*), were significantly lower among patients who experienced serious delay compared with no-delay groups with all FDR-corrected p-value <0.1. Among patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, we also found an increased proportion of carriers of two genera, including *Lancefieldella* (belong to the phylum *Actinobacteriota)* and *CAG-56* (belong to the phylum *Firmicutes A),* along with two species belonging to the phylum *Bacteroidota*, including *Bacteroides A mediterraneensis* and *MGYG-HGUT-04188,* in patients who experienced a serious delay.

**Table 35**: Association of BMI levels with GI microbial taxa by stool collection time

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for* underweight *vs.* normal weight | *P* | $P_{FDR}$[4] | *β (SE) for* overweight *vs.* normal weight | *P* | $P_{FDR}$[4] | *β (SE) for* obese *vs.* normal weight | *P* | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Overall (N=356)** [1] | | | | | | | | | | | |
| Phylum *Actinobacteriota* | | | | | | | | | | | |
| Family *Actinomycetaceae* | 0.010 | 83.1 | 1.106 (0.374) | 0.003 | 0.068 | -0.236 (0.274) | 0.390 | 0.970 | -0.058 (0.375) | 0.877 | 0.999 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Order *Lachnospirales* | 14.503 | 100.0 | -0.407 (0.128) | 0.002 | 0.014 | -0.09 (0.094) | 0.340 | 0.851 | 0.023 (0.128) | 0.858 | 0.948 |
| Family *Lachnospiraceae* | 14.278 | 100.0 | -0.351 (0.128) | 0.007 | 0.075 | -0.064 (0.094) | 0.494 | 0.972 | 0.065 (0.128) | 0.613 | 0.999 |
| Order *TANB77* | 0.061 | 99.4 | -0.724 (0.222) | 0.001 | 0.014 | 0.143 (0.162) | 0.377 | 0.851 | -0.048 (0.222) | 0.828 | 0.948 |
| Family *CAG-508* | 0.055 | 98.9 | -0.737 (0.239) | 0.002 | 0.068 | 0.167 (0.175) | 0.341 | 0.970 | 0.017 (0.239) | 0.944 | 0.999 |
| Phylum *Firmicutes B* | | | | | | | | | | | |
| Order *Peptococcales* | 0.006 | 84.0 | -0.626 (0.238) | 0.009 | 0.047 | -0.394 (0.174) | 0.024 | 0.420 | -0.051 (0.238) | 0.831 | 0.948 |
| Phylum *Firmicutes C* | | | | | | | | | | | |
| Order *Acidaminococcales* | 0.533 | 96.1 | -1.077 (0.381) | 0.005 | 0.035 | -0.297 (0.279) | 0.288 | 0.851 | -0.534 (0.382) | 0.163 | 0.890 |
| Family *Acidaminococcaceae* | 0.533 | 96.1 | -1.038 (0.386) | 0.007 | 0.075 | -0.276 (0.282) | 0.329 | 0.970 | -0.495 (0.386) | 0.201 | 0.789 |
| Order *Veillonellales* | 0.098 | 99.7 | 0.903 (0.363) | 0.014 | 0.059 | 0.252 (0.266) | 0.345 | 0.851 | 0.570 (0.364) | 0.118 | 0.890 |
| Family *Veillonellaceae* | 0.052 | 99.7 | 1.022 (0.389) | 0.009 | 0.078 | 0.050 (0.284) | 0.860 | 0.997 | 0.480 (0.389) | 0.218 | 0.789 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Order *Haloplasmatales* | 0.002 | 60.7 | 0.806 (0.337) | 0.017 | 0.068 | -0.172 (0.247) | 0.486 | 0.851 | -1.012 (0.338) | 0.003 | 0.104 |
| Family *Turicibacteraceae* | 0.002 | 60.7 | 0.873 (0.346) | 0.012 | 0.095 | -0.157 (0.253) | 0.536 | 0.972 | -0.993 (0.347) | 0.004 | 0.310 |
| Order *Lactobacillales* | 0.224 | 100.0 | 1.193 (0.330) | $3.40\times10^{-4}$ | 0.006 | 0.068 (0.241) | 0.780 | 0.987 | -0.279 (0.330) | 0.399 | 0.890 |
| Family *Aerococcaceae* | 0.002 | 57.3 | 0.945 (0.338) | 0.006 | 0.075 | 0.157 (0.248) | 0.528 | 0.972 | -0.13 (0.339) | 0.701 | 0.999 |
| Family *Enterococcaceae* | 0.024 | 95.8 | 1.872 (0.481) | $1.18\times10^{-4}$ | 0.008 | 0.16 (0.352) | 0.650 | 0.988 | 0.173 (0.481) | 0.720 | 0.999 |
| Family *Lactobacillaceae* | 0.008 | 83.4 | 1.303 (0.448) | 0.004 | 0.068 | -0.484 (0.328) | 0.141 | 0.941 | -0.119 (0.449) | 0.791 | 0.999 |
| Order *RF39* | 0.043 | 98.6 | -0.67 (0.256) | 0.009 | 0.047 | 0.060 (0.188) | 0.751 | 0.987 | 0.288 (0.257) | 0.262 | 0.890 |
| Order *Staphylococcales* | 0.001 | 50.6 | 1.178 (0.277) | $2.75\times10^{-5}$ | 0.001 | 0.234 (0.203) | 0.250 | 0.851 | 0.099 (0.278) | 0.723 | 0.948 |
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)** [2] | | | | | | | | | | | |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Order *Lachnospirales* | | | | | | | | | | | |
| Genus *Agathobacter* | 0.161 | 96.9 | -1.702 (0.500) | 0.001 | 0.082 | -0.903 (0.371) | 0.016 | 0.918 | -0.476 (0.498) | 0.341 | 0.984 |
| Genus *CAG-603* | 0.003 | 74.7 | -1.043 (0.309) | 0.001 | 0.082 | -0.574 (0.229) | 0.013 | 0.918 | -0.615 (0.307) | 0.047 | 0.949 |
| Genus *Dorea* | 0.342 | 99.4 | -0.782 (0.241) | 0.001 | 0.082 | -0.306 (0.179) | 0.090 | 0.918 | -0.209 (0.240) | 0.384 | 0.984 |
| Genus *Agathobaculum* | 0.075 | 97.8 | -0.705 (0.297) | 0.019 | 0.175 | -0.244 (0.221) | 0.271 | 0.918 | -1.143 (0.296) | $1.70\times10^{-4}$ | 0.047 |
| Order *Peptostreptococcales* | 0.288 | 100.0 | 0.664 (0.218) | 0.003 | 0.033 | 0.068 (0.162) | 0.674 | 0.994 | -0.166 (0.217) | 0.444 | 0.977 |
| Phylum *Firmicutes C* | | | | | | | | | | | |
| Order *Acidaminococcales* | 0.533 | 96.1 | -1.734 (0.562) | 0.002 | 0.033 | -0.431 (0.417) | 0.303 | 0.994 | -1.002 (0.559) | 0.075 | 0.977 |

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for* underweight *vs.* normal weight | *P* | $P_{FDR}$[4] | *β (SE) for* overweight *vs.* normal weight | *P* | $P_{FDR}$[4] | *β (SE) for* obese *vs.* normal weight | *P* | $P_{FDR}$[4] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Family *Acidaminococcaceae* | 0.533 | 96.1 | -1.758 (0.572) | 0.003 | 0.059 | -0.432 (0.424) | 0.311 | 0.996 | -1.006 (0.569) | 0.079 | 0.767 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Order *Lactobacillales* | 0.224 | 100.0 | 1.568 (0.442) | 0.001 | 0.018 | 0.182 (0.328) | 0.580 | 0.994 | -0.657 (0.439) | 0.137 | 0.977 |
| Family *Aerococcaceae* | 0.002 | 57.3 | 1.624 (0.481) | 0.001 | 0.033 | -0.097 (0.357) | 0.786 | 0.996 | -0.490 (0.478) | 0.307 | 0.768 |
| Genus *Granulicatella* | 0.002 | 57.3 | 1.848 (0.507) | $3.78 \times 10^{-4}$ | 0.082 | -0.105 (0.376) | 0.780 | 0.959 | -0.500 (0.504) | 0.323 | 0.984 |
| Family *Enterococcaceae* | 0.024 | 95.8 | 2.546 (0.685) | $2.91 \times 10^{-4}$ | 0.020 | 0.399 (0.508) | 0.434 | 0.996 | 0.303 (0.681) | 0.657 | 0.952 |
| Genus *Enterococcus A* | 0.005 | 86.2 | 1.817 (0.560) | 0.001 | 0.082 | 0.234 (0.416) | 0.574 | 0.927 | -0.212 (0.557) | 0.705 | 0.987 |
| Genus *Enterococcus B* | 0.005 | 69.4 | 2.637 (0.829) | 0.002 | 0.084 | 0.584 (0.615) | 0.344 | 0.918 | -0.127 (0.825) | 0.878 | 0.998 |
| Order *Staphylococcales* | 0.001 | 50.6 | 1.069 (0.375) | 0.005 | 0.044 | 0.478 (0.278) | 0.088 | 0.844 | -0.032 (0.373) | 0.931 | 0.977 |
| **Stool collection after breast cancer surgery (N=194)** [3] | | | | | | | | | | | |
| Phylum *Cyanobacteria* | 0.040 | 98.9 | -0.700 (0.228) | 0.002 | 0.030 | 0.257 (0.165) | 0.120 | 0.869 | 0.169 (0.235) | 0.472 | 0.957 |
| Class *Vampirovibrionia* | 0.040 | 98.9 | -0.846 (0.244) | 0.001 | 0.009 | 0.219 (0.176) | 0.216 | 0.826 | 0.112 (0.252) | 0.658 | 0.952 |
| Order *Gastranaerophilales* | 0.040 | 98.9 | -0.893 (0.236) | 0.000 | 0.007 | 0.197 (0.170) | 0.248 | 0.921 | 0.121 (0.243) | 0.619 | 0.950 |
| Family *Gastranaerophilaceae* | 0.039 | 98.9 | -0.797 (0.242) | 0.001 | 0.083 | 0.229 (0.175) | 0.191 | 0.903 | 0.195 (0.250) | 0.436 | 0.983 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
[1] Models were adjusted for stool collection time, age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 36:** Association of diagnosis delay with GI microbial taxa by stool collection time

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for moderate vs. no delay* | *P* | $P_{FDR}$[2] | *β (SE) for serious vs. no delay* | *P* | $P_{FDR}$[2] |
|---|---|---|---|---|---|---|---|---|
| **Stool collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)** [2] | | | | | | | | |
| Phylum *Actinobacteriota* | | | | | | | | |
| Genus *Lancefieldella* ┼ | 0.003¥ | 24.2 | 0.112 (0.573) | 0.845 | 0.923 | 1.322 (0.542) | 0.015 | 0.096 |
| Phylum *Bacteroidota* | | | | | | | | |
| Species *Bacteroides A mediterraneensis* | 0.016 | 93.5 | 0.367 (0.252) | 0.148 | 0.706 | 1.063 (0.266) | $1.06x10^{-4}$ | 0.080 |
| Species *MGYG-HGUT-04188* | 0.003 | 75.6 | 0.182 (0.321) | 0.571 | 0.918 | 1.233 (0.339) | $3.86x10^{-4}$ | 0.097 |
| Phylum *Elusimicrobiota* ┼ | 0.005¥ | 12.6 | -1.688 (0.711) | 0.018 | 0.088 | -0.545 (0.647) | 0.400 | 0.935 |
| Phylum *Firmicutes A* | | | | | | | | |
| Class *Clostridia* | | | | | | | | |
| Order *4C28d-15* | 0.027 | 93.8 | -1.113 (0.337) | 0.001 | 0.042 | -0.754 (0.355) | 0.036 | 0.567 |
| Family *CAG-314* ┼ | 0.003¥ | 28.1 | -1.407 (0.552) | 0.011 | 0.165 | -1.922 (0.669) | 0.004 | 0.058 |
| Family *CAG-552* ┼ | 0.007¥ | 13.2 | -1.722 (0.650) | 0.008 | 0.165 | -2.190 (0.774) | 0.005 | 0.058 |
| Genus *UBA10677* ┼ | 0.005¥ | 11.5 | -1.679 (0.669) | 0.012 | 0.165 | -2.141 (0.796) | 0.007 | 0.080 |
| Family *CAG-727* | | | | | | | | |
| Genus *QALS01* ┼ | 0.008¥ | 30.6 | -0.877 (0.481) | 0.068 | 0.262 | -2.122 (0.638) | 0.001 | 0.052 |
| Genus *UBA10281* ┼ | 0.014¥ | 20.5 | -1.120 (0.535) | 0.036 | 0.204 | -2.187 (0.673) | 0.001 | 0.052 |
| Genus *UBA1259* ┼ | 0.004¥ | 16.3 | -1.529 (0.582) | 0.009 | 0.165 | -1.612 (0.636) | 0.011 | 0.094 |
| Family *CAG-917* ┼ | 0.014¥ | 48.0 | -0.980 (0.491) | 0.046 | 0.260 | -1.611 (0.530) | 0.002 | 0.055 |
| Genus *CAG-349* ┼ | 0.012¥ | 43.3 | -0.667 (0.469) | 0.155 | 0.399 | -1.219 (0.502) | 0.015 | 0.096 |
| Genus *CAG-475* ┼ | 0.005¥ | 14.3 | -0.875 (0.588) | 0.137 | 0.365 | -2.108 (0.87) | 0.015 | 0.096 |
| Order *Christensenellales* | | | | | | | | |
| Family *CAG-138* | | | | | | | | |
| Genus *Phil1* ┼ | 0.006¥ | 29.2 | -0.631 (0.482) | 0.190 | 0.424 | -1.417 (0.540) | 0.009 | 0.089 |
| Genus *UBA1685* ┼ | 0.003¥ | 34.6 | -1.027 (0.501) | 0.041 | 0.204 | -2.329 (0.603) | $1.12x10^{-4}$ | 0.031 |
| Family *CAG-74* | | | | | | | | |
| Genus *MGYG-HGUT-01658* ┼ | 0.005¥ | 46.9 | 0.095 (0.471) | 0.840 | 0.923 | -1.193 (0.503) | 0.018 | 0.099 |
| Genus *MGYG-HGUT-01823* ┼ | 0.005¥ | 17.4 | -0.673 (0.526) | 0.201 | 0.438 | -2.681 (0.850) | 0.002 | 0.056 |
| Genus *MGYG-HGUT-02098* ┼ | 0.005¥ | 37.1 | -0.470 (0.459) | 0.305 | 0.542 | -1.374 (0.534) | 0.010 | 0.089 |
| Genus *MGYG-HGUT-03224* ┼ | 0.013¥ | 48.0 | -0.671 (0.482) | 0.164 | 0.410 | -1.198 (0.506) | 0.018 | 0.099 |
| Genus *MGYG-HGUT-04052* ┼ | 0.005¥ | 10.1 | -1.989 (0.735) | 0.007 | 0.165 | -3.082 (0.910) | 0.001 | 0.052 |
| Family *QALW01* | | | | | | | | |
| Genus *MGYG-HGUT-01665* ┼ | 0.003¥ | 25.3 | -1.114 (0.536) | 0.038 | 0.204 | -1.724 (0.671) | 0.010 | 0.089 |
| Genus *MGYG-HGUT-03967* ┼ | 0.003¥ | 33.7 | -0.714 (0.448) | 0.111 | 0.331 | -1.304 (0.500) | 0.009 | 0.089 |
| Order *Lachnospirales* | | | | | | | | |
| Genus *MGYG-HGUT-01118* ┼ | 0.005¥ | 27.2 | -0.048 (0.461) | 0.917 | 0.970 | -1.308 (0.540) | 0.015 | 0.096 |
| Family *Lachnospiraceae* | | | | | | | | |

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for moderate vs. no delay* | *P* | $P_{FDR}$[2] | *β (SE) for serious vs. no delay* | *P* | $P_{FDR}$[2] |
|---|---|---|---|---|---|---|---|---|
| Genus *CAG-56* | 0.022 | 96.3 | 1.159 (0.312) | 2.88x10$^{-4}$ | 0.080 | 0.257 (0.329) | 0.436 | 0.703 |
| Genus *CAG-791* ┼ | 0.002¥ | 24.2 | -1.266 (0.511) | 0.013 | 0.165 | -1.384 (0.561) | 0.014 | 0.096 |
| Genus *MGYG-HGUT-03215* ┼ | 0.005¥ | 42.1 | 0.104 (0.484) | 0.829 | 0.921 | -1.225 (0.501) | 0.014 | 0.096 |
| Genus *MGYG-HGUT-04548* ┼ | 0.003¥ | 48.0 | -0.316 (0.491) | 0.519 | 0.716 | -1.358 (0.498) | 0.006 | 0.080 |
| Species *Roseburia_hominis* | 0.007 | 80.9 | -0.330 (0.366) | 0.369 | 0.896 | -1.409 (0.386) | 3.75x10$^{-4}$ | 0.097 |
| Genus *UBA4285* ┼ | 0.003¥ | 48.9 | -0.305 (0.429) | 0.478 | 0.693 | -1.122 (0.464) | 0.016 | 0.096 |
| Family *UBA1390* ┼ | 0.003¥ | 36.2 | -1.121 (0.474) | 0.018 | 0.165 | -1.296 (0.515) | 0.012 | 0.100 |
| Genus *UBA1390* ┼ | 0.003¥ | 36.2 | -1.121 (0.474) | 0.018 | 0.165 | -1.296 (0.515) | 0.012 | 0.094 |
| Order *Monoglobales* | | | | | | | | |
| Family *MGYG-HGUT-02683* ┼ | 0.002¥ | 13.5 | -0.646 (0.562) | 0.250 | 0.566 | -3.176 (1.137) | 0.005 | 0.058 |
| Genus *MGYG-HGUT-02683* ┼ | 0.002¥ | 13.5 | -0.646 (0.562) | 0.250 | 0.499 | -3.176 (1.137) | 0.005 | 0.076 |
| Family *MGYG-HGUT-04133* ┼ | 0.002¥ | 10.1 | -2.220 (1.078) | 0.039 | 0.260 | -4.163 (1.355) | 0.002 | 0.055 |
| Genus *MGYG-HGUT-04133* ┼ | 0.002¥ | 10.1 | -2.220 (1.078) | 0.039 | 0.204 | -4.163 (1.355) | 0.002 | 0.065 |
| Family *Acutalibacteraceae* | | | | | | | | |
| Genus *Anaeromassilibacillus* ┼ | 0.003¥ | 34.3 | -1.353 (0.501) | 0.007 | 0.165 | -1.337 (0.530) | 0.012 | 0.094 |
| Genus *CAG-488* ┼ | 0.004¥ | 40.7 | -0.977 (0.468) | 0.037 | 0.204 | -1.837 (0.525) | 4.66x10$^{-4}$ | 0.052 |
| Genus *MGYG-HGUT-02705* ┼ | 0.002¥ | 16.0 | -1.618 (0.624) | 0.010 | 0.165 | -2.268 (0.754) | 0.003 | 0.065 |
| Genus *MGYG-HGUT-03278* ┼ | 0.002¥ | 11.8 | -1.409 (0.659) | 0.033 | 0.197 | -3.186 (1.126) | 0.005 | 0.075 |
| Genus *RUG806* ┼ | 0.003¥ | 41.6 | -1.044 (0.470) | 0.026 | 0.174 | -1.450 (0.504) | 0.004 | 0.071 |
| Family *Butyricicoccaceae* | | | | | | | | |
| Genus *MGYG-HGUT-02627* ┼ | 0.003¥ | 25.8 | -0.415 (0.479) | 0.386 | 0.615 | -1.733 (0.604) | 0.004 | 0.071 |
| Family *CAG-272* | | | | | | | | |
| Genus *CAG-724* ┼ | 0.003¥ | 41.9 | -2.049 (0.526) | 9.69x10$^{-5}$ | 0.027 | -1.595 (0.532) | 0.003 | 0.065 |
| Genus *CAG-841* ┼ | 0.004¥ | 24.4 | -2.001 (0.561) | 3.59x10$^{-4}$ | 0.050 | -0.784 (0.502) | 0.118 | 0.260 |
| Genus *MGYG-HGUT-00386* ┼ | 0.002¥ | 27.5 | -0.409 (0.457) | 0.371 | 0.602 | -1.354 (0.569) | 0.017 | 0.099 |
| Genus *MGYG-HGUT-00468* ┼ | 0.009¥ | 16.9 | -1.461 (0.623) | 0.019 | 0.165 | -2.431 (0.738) | 0.001 | 0.052 |
| Genus *MGYG-HGUT-02723* ┼ | 0.006¥ | 20.5 | -1.357 (0.525) | 0.010 | 0.165 | -1.733 (0.596) | 0.004 | 0.071 |
| Genus *MGYG-HGUT-03963* ┼ | 0.012¥ | 15.4 | -0.892 (0.56) | 0.111 | 0.331 | -1.663 (0.687) | 0.016 | 0.096 |
| Genus *MGYG-HGUT-03979* ┼ | 0.003¥ | 16.3 | -1.306 (0.564) | 0.021 | 0.165 | -1.606 (0.670) | 0.017 | 0.098 |
| Genus *MGYG-HGUT-04098* ┼ | 0.002¥ | 12.6 | -2.822 (0.972) | 0.004 | 0.165 | -2.416 (0.995) | 0.015 | 0.096 |
| Genus *MGYG-HGUT-04214* ┼ | 0.002¥ | 10.4 | -2.524 (1.001) | 0.012 | 0.165 | -3.801 (1.477) | 0.010 | 0.089 |
| Family *CAG-382* | | | | | | | | |
| Genus *UBA1206* ┼ | 0.005¥ | 21.9 | -0.994 (0.516) | 0.054 | 0.243 | -1.595 (0.613) | 0.009 | 0.089 |
| Family *Oscillospiraceae* | | | | | | | | |
| Genus *MGYG-HGUT-02840* ┼ | 0.002¥ | 25.3 | -1.360 (0.554) | 0.014 | 0.165 | -1.626 (0.605) | 0.007 | 0.080 |
| Genus *UBA738* ┼ | 0.008¥ | 34.6 | -0.425 (0.447) | 0.342 | 0.574 | -1.199 (0.491) | 0.015 | 0.096 |
| Family *Ruminococcaceae* | | | | | | | | |
| Genus *CAG-115* ┼ | 0.006¥ | 43.8 | -0.580 (0.463) | 0.210 | 0.451 | -1.284 (0.497) | 0.010 | 0.089 |

| *Taxonomy* | RA, median (%) | Pre (%) | *β (SE) for moderate vs. no delay* | *P* | $P_{FDR}$[2] | *β (SE) for serious vs. no delay* | *P* | $P_{FDR}$[2] |
|---|---|---|---|---|---|---|---|---|
| Genus *D5* ┼ | 0.002¥ | 19.4 | -0.990 (0.544) | 0.069 | 0.262 | -1.763 (0.647) | 0.006 | 0.080 |
| Genus *MGYG-HGUT-00425* ┼ | 0.005¥ | 15.4 | -1.130 (0.599) | 0.059 | 0.248 | -2.632 (0.881) | 0.003 | 0.065 |
| Genus *MGYG-HGUT-00450* ┼ | 0.002¥ | 12.9 | -1.571 (0.614) | 0.010 | 0.165 | -2.506 (0.864) | 0.004 | 0.071 |
| Genus *MGYG-HGUT-00537* ┼ | 0.003¥ | 23.0 | -0.634 (0.473) | 0.180 | 0.421 | -1.922 (0.598) | 0.001 | 0.052 |
| Genus *MGYG-HGUT-04104* ┼ | 0.002¥ | 19.9 | -1.216 (0.534) | 0.023 | 0.165 | -1.653 (0.613) | 0.007 | 0.080 |
| Family *UBA644* ┼ | 0.004¥ | 11.2 | -1.163 (0.725) | 0.109 | 0.447 | -2.418 (0.930) | 0.009 | 0.090 |
| Order *Peptostreptococcales* | | | | | | | | |
| Genus *MGYG-HGUT-01191* ┼ | 0.006¥ | 24.4 | -0.950 (0.513) | 0.064 | 0.258 | -1.626 (0.578) | 0.005 | 0.075 |
| Genus *GCA-900066495* ┼ | 0.006¥ | 19.7 | -0.225 (0.512) | 0.661 | 0.810 | -1.929 (0.709) | 0.007 | 0.080 |
| Order *TANB77* | | | | | | | | |
| Genus *UBA1234* ┼ | 0.003¥ | 14.0 | -1.138 (0.605) | 0.060 | 0.248 | -1.796 (0.750) | 0.017 | 0.098 |
| Phylum *Firmicutes* | | | | | | | | |
| Family *CAG-433* ┼ | 0.029¥ | 15.4 | -0.722 (0.533) | 0.175 | 0.514 | -2.259 (0.746) | 0.002 | 0.055 |

Common taxa (prevalence≥50% in the population): linear regression was conducted for clr- (centered log-ratio) transformed taxa abundance.
┼ Rare taxa: 10% ≤prevalence< 50% in the population; Rare taxa: negative binomial hurdle model was conducted for read counts of rare taxa.
¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

## 3. Discussion

The study of 356 Vietnamese breast cancer patients evaluated the associations of GI microbial richness and composition as well as individual microbial taxa with demographic and clinical factors. Stool samples from over half (54.5%) of study participants, collected after their breast cancer surgery, were associated with significantly lower alpha diversity indexes, beta diversity explained, and significant differences in taxa abundance of approximately 40% of investigated GI microbial taxa. We observed no significant associations between alpha or beta diversity and clinical characteristics, except for significantly higher alpha diversity indexes for stage III-IV breast cancer patients whose stool samples were collected after breast cancer surgery. However, we found that delay in diagnosis and treatment, particularly among patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, was significantly associated with lower alpha diversity indexes, beta diversity, and reduced proportions of carriers of species *Roseburia hominis* and *Firmicutes* populations, including 48 genera and six families within the class *Clostridia* (belonging to the phylum *Firmicutes A*). Furthermore, many GI microbial taxa were associated with age, income, and residence, including significantly lower abundances or proportions of carriers among old patients, patients having low income, and living in rural areas. Finally, we also found that some GI microbial taxa were significantly associated with cancer stages, breast cancer subtypes, ER and HER2 status, and BMI levels.

Few studies have investigated the GI microbiome among breast cancer patients according to demographic characteristics and clinical features. These studies suggest that the GI microbiome may vary by clinical stages, histo-prognostic grades, clinicopathological factors, and breast cancer risk factors such as BMI, age at menarche, and menopausal status.[145] The GI microbial composition among 31 patients with early-stage breast cancer differed according to clinical stages, histo-prognostic grades and BMI, which was reported by Luu in a case-only study in 2017.[150] Luu and colleagues found that the percentage and the absolute numbers of specific bacterial groups such as *Clostridium leptum*, *Clostridium coccoides*, *Facecalibacterium prausnitzii,* and *Blautia* species, were significantly higher in clinical stage II/ III breast cancer patients than in clinical stage 0 and I patients, and *Blautia* species was also were extraordinarily higher in patients with grade III than in those with grade I/II breast cancer.[150]

*Firmicutes*, *Faecalibacterium prausnitzii*, and *Blautia* species showed a significantly declined abundance in overweight and obese patients compared to those with normal BMI.[150] In the analysis of all 356 participants, we observed a higher Chao1 and Shannon indexes among stage II and stage III-IV breast cancer patients than patients with stage I. However, in multivariable analyses, we only found one significant association with higher alpha diversity indexes for stage III-IV breast cancer among patients whose stool samples were collected after breast cancer surgery. Moreover, in the analysis for this sub-population, we found a significantly higher proportion of carriers of family *MGYG-HGUT-04147* *(*belonging to phylum *Firmicutes A)* in stage III-IV breast cancer patients than in stage I patients. In addition, underweight breast cancer patients had a significantly lower abundance of orders *Lachnospirales* and *Acidaminococcales* along with its family *Acidaminococcaceae,* and a higher abundance of order Staphylococcales, families *Aerococcaceae* and *Enterococcaceae* in both analyses of all patients and subgroup of patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy. We found that the genus *Agathobaculum* was significantly less abundant among obese breast cancer patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. The species *Agathobaculum butyriciproducenes,* a butyrate-producing and lactate-producing bacterium linked to healthy plant-based foods*,* is a member of this genus *Agathobaculum*.

An early study reported that patients with different clinicopathological factors (i.e., ER, PR, HER2, and Ki-67 expression) and reproductive factors showed different gut microbiome. Wu and colleagues found that HER2 status and age at menarche were significantly associated with high alpha diversity indexes of GI microbiome and specific microbial composition.[152] Yang, Wang, and co-workers also observed a distinct enrichment of particular gut microbiota by different clinicopathological characteristics, including ER, PR, Ki-67 levels, HER2 status, and tumor type (malignant vs. benign).[151] Members of the family *Prevotellaceae* were more abundant in patients with PR+ or ER+, whereas some bacteria, including *Hydrogenophilus, Lactobacillus,* and *Acinetobacter,* were more abundant in breast cancer patients with PR- and ER- tumors. In addition, enrichment of *Megasphaera* was observed in patients with ER+ and HER2-positive tumors.[151] These findings suggested that there might be specific gut microbiome

communities among breast cancer patients according to different clinicopathological factors.[151] However, in our study, we found no significant association between alpha and beta diversity with menopausal status, the status of ER, PR, Ki-67 levels, HER2, or breast cancer subtypes. In GI microbial taxa analysis, we found a significantly higher proportion of carriers of phyla *Firmicutes I* and *Synergistota* and a lower proportion of carriers of phylum *Campylobacterota* among patients with HER2-positive breast cancer than HER2-negative breast cancer in the analysis of all patients. Among patients whose stool samples were collected after breast cancer surgery, a significantly higher proportion of carriers of phylum *Firmicutes I* was consistently observed in HER2-positive breast cancer compared with HER2-negative breast cancer. We also found that HER2-positive breast cancer was associated with a decreased proportion of carriers of phylum *Elusimicrobiota*. We observed a higher proportion of carriers of phylum *Elusimicrobiota* in ER+ breast cancer patients than those with ER- breast cancer. We also found no association between menopausal status and GI microbial taxa. In terms of breast cancer subtypes, the family *Ruminococcaceae* and four of its species were significantly more abundant among luminal/HER2-positive breast cancer patients than luminal/HER2-negative breast cancer patients. In addition, the genus *Acutalibacter* showed a higher abundance among patients with HER2 enriched breast cancer compared with luminal/HER2-negative breast cancer patients. Conversely, *Campylobacterota* was significantly less prevalent among patients with HER2 enriched breast cancer than luminal/HER2-negative breast cancer. However, the mechanism underlying these associations remains unknown. Finally, no significant association between triple-negative/basal-like breast cancer with GI microbial taxa was found in our study.

Our study suggested that age, income, and geographic residence were associated with the GI microbiome in breast cancer patients, with significantly declined abundances or proportions of carriers of many GI microbiota among elderly patients, patients with low income, and living in rural areas. Notably, our study found that those who experienced a delay in diagnosis and treatment (aka. diagnosis delay), particularly those whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, had significantly lower alpha diversity index and variation in beta diversity. Noteworthy, significant associations between diagnosis delay and 61 microbial taxa were found among patients with

stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. Most of them, which are members of the class *Clostridia*, including six families, 48 genera, and particularly the species *Roseburia hominis*, showed significantly decreased proportions of carriers among patients who experienced to a serious delay in diagnosis and treatment. *Roseburia hominis*, a gut anaerobic bacterium belonging to the family *Lachnospiraceae* is a well-known butyrate-producing bacterium.[240] Butyrate is an essential metabolite in the human colon. Many evidence suggested the benefits of butyrate to intestinal health as well as overall health.[241] Butyrate acts as the preferred energy source for colonic epithelia cells and contributes to maintain the gut barrier functions by exerting anti-inflammation effects, decreasing the luminal pH to reduce bile salt solubility, inhibiting ammonia absorption, and hampering the invasion of pathogens as well.[242] *Roseburia hominis*, together with *Faecalibacterium prausnitzii* could also serve as probiotics to restore beneficial flora.[241] Conversely, we found an increased proportion of carriers of genus *Lancefieldella* (belong to the phylum *Actinobacteriota)* and species *Bacteroides A mediterraneensis* (belong to the phylum *Bacteroidota)* among patients with serious diagnosis delay.

Among 356 Vietnamese breast cancer patients, delay in diagnosis and treatment was common (49.7%), accounting for 60% of patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy. In our study population, the median time interval from first signs/noticeable breast cancer symptoms to diagnosis and initiation of treatment was 2.4 months (IQR: 1.1 to 7.1 months) overall and 5.5 months (IQR: 2.5-9.3 months) for the breast cancer patients who postponed seeking medical care after first symptom recognition.[26] Since some breast cancer-related symptoms such as pain, tenderness, and lumps are non-cancer specific, patients and their health consultants/professionals might have failed to consider breast cancer as a possible diagnosis.[243] Antibiotics, herbal or alternative medical treatment might have been used among women residing in rural areas when symptoms first appear, which may have altered patients' gut microbiota. In Vietnam, antibiotics can be purchased over the counter; The most commonly purchased antibiotics include the broad spectrum of penicillin, such as amoxicillin and ampicillin and first-generation cephalosporins, such as cephalexin.[244] Furthermore, we speculate that physical and psychological stresses experienced by

patients during the delay period may altered patients' dietary intakes[207] leading to changes in the gut microbial community.

We recruited newly diagnosed breast cancer patients in different settings from inpatient surgical units and chemotherapy inpatient and outpatient units of two major cancer hospitals. At the time of stool collection, over half of the study participants had undergone breast cancer surgery and were prescribed prophylactic antibiotics as a routine protocol to prevent infections after breast cancer surgery. As a result, patients with pre-breast cancer treatment and non-antibiotic exposure accounted for only 45% of our participants, which is a limitation of our study design. Nevertheless, our results confirmed the findings from many studies that GI microbiome profiles among cancer patients changed considerably after breast cancer surgery and prophylaxis antibiotics before chemotherapy.[160-166,168,169,189] We found a significantly lower GI microbial richness among patients whose stool samples were collected following breast cancer surgery and prophylactic antibiotics. The GI microbial richness appears to diminish and then recover after three weeks from breast cancer surgery. The changes in the GI microbiome might attenuate or distort the associations of GI microbial richness and composition as well as individual GI microbial taxa with demographic and clinical factors. This may explain why we found only a few associations between GI microbial richness and composition with cancer stages, breast cancer subtypes, ER, PR status, HER2 expression, and other factors such as BMI levels in the analysis of all participants. Nevertheless, clinical characteristics were not associated with gut microbiome in the analyses stratified by stool sample collection time.

This is the first and largest study evaluating the association between GI microbiome and clinical and demographic factors among Vietnamese breast cancer patients to date. Stool collection, transportation, and storage of stool samples following a standard protocol to minimize errors. In addition, the shotgun metagenomic sequencing performed in our study not only provided enhanced taxonomic resolution but made it possible to assess the functional potentials of GI microbiota in further studies. We also used a human bacterial genome from the UHGG collection as the reference, a massive sequence catalog containing the information of more than 4,600 species, with 71% lacking a cultured

representative.[231-233] This allowed us to estimate the prevalence and abundance of species or genes with enhanced resolution and accuracy.

Several limitations in our study should be considered when interpreting our findings. Although we carefully adjusted for a variety of covariates, residual confounding from other unknown or unmeasured confounders cannot be excluded when evaluating evaluated the associations of the GI microbiome with non-clinical and clinical factors. As aforementioned, our study is limited by over half of the study participants had undergone breast cancer surgery and were prescribed prophylactic antibiotics to prevent infections after breast cancer surgery before their stool samples were collected. The statistical power of the study, thus, was compromised. In addition, some of the observed associations might be caused by antibiotics use. True associations might have been missed. Last but not least, our findings may not be generalizable to the GI microbiome profile of all Vietnamese breast cancer patients, particularly those treated in other settings in Vietnam.

In conclusion, we found a significantly lower GI microbial richness and composition among breast cancer patients who received breast cancer surgery and followed by antibiotic treatment compared with the GI microbiome of patients before breast cancer surgery and neoadjuvant chemotherapy. In addition, serious delay in diagnosis was significantly associated with lower GI microbial richness and composition and decreased abundance of *Roseburia hominis* and members of the class *Clostridia*, particularly among patients whose GI microbiome was assessed pre-breast cancer surgery. We also found several GI microbial taxa that were significantly associated with age, income, residence, cancer stages, breast cancer subtypes, ER and HER2 status, and BMI levels.

# CHAPTER 4

SPECIFIC AIM 3

**Pre-chemotherapy GI microbiome and chemotherapy-induced toxicity**

To evaluate the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients.

## 1. Methods

### 1.1. Parent Study

This study included 396 newly diagnosed Vietnamese breast cancer patients who have received neoadjuvant or adjuvant chemotherapy during breast cancer treatment in Aim 1. Details of breast cancer case recruitment in Aim 1 have been described.

### 1.2. Population Selection

To be included in this analysis, newly diagnosed breast cancer patients in this study had provided stool samples at baseline and received neoadjuvant or adjuvant chemotherapy at Vietnam National Cancer Hospital (K Hospitals) and Hanoi Oncology hospital during the follow-up period. We excluded participants who did not donate stool samples (n=91) at the baseline survey. In addition, stool samples from 4 participants were excluded due to low DNA yields (n=4). Finally, A total of 301 participants was included for Aim 3 (**Figure 12**).

### 1.3. Outcome Assessment

Four chemotherapy-induced toxicities including the combined hematological toxicity, the combined GI toxicity, neutropenia, and nausea/vomiting as defined in Aim 1 are the primary outcomes for Aim 3. Briefly, the outcomes of the study are the highest grade of toxicities reached during the first-line chemotherapy treatment until the first day of radiotherapy for patients who received chemotherapy and/or sequential radiotherapy, or during the first-line treatment of chemotherapy and 90 follow-up days after the treatment for patients who received only chemotherapy without radiotherapy. The combined

hematological toxicity refers to having any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia, while the combined GI toxicity was combined to include the four symptoms: nausea, vomiting, diarrhea, constipation, or stomatitis. For evaluating association analyses, the combined and specific toxicities were grouped as dichotomous variables (i.e., grade ≥3 vs. grade <3).

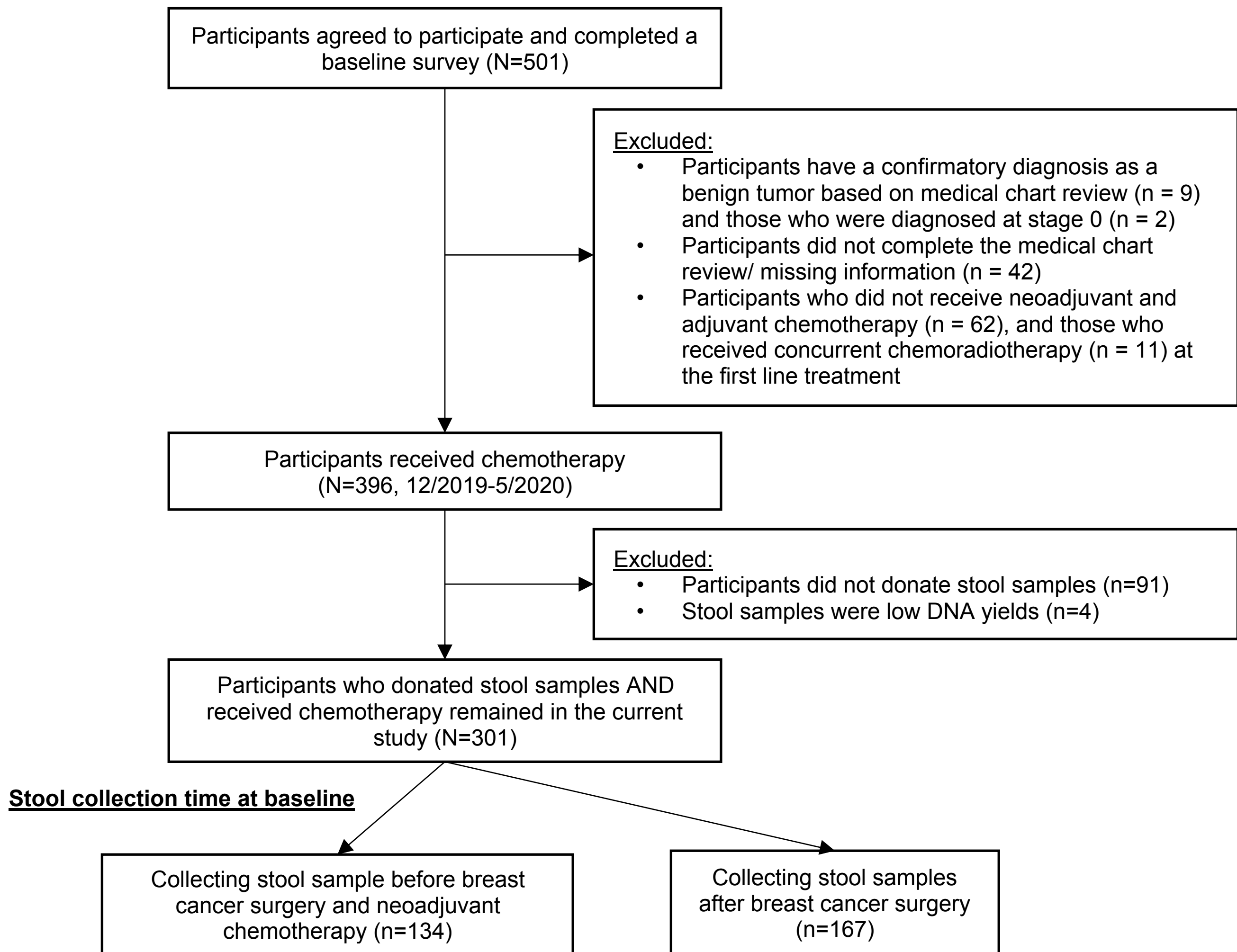



**Figure 12:** Flow diagram of study subject inclusion criteria for Aim 3

1.4. Stool Collection Time

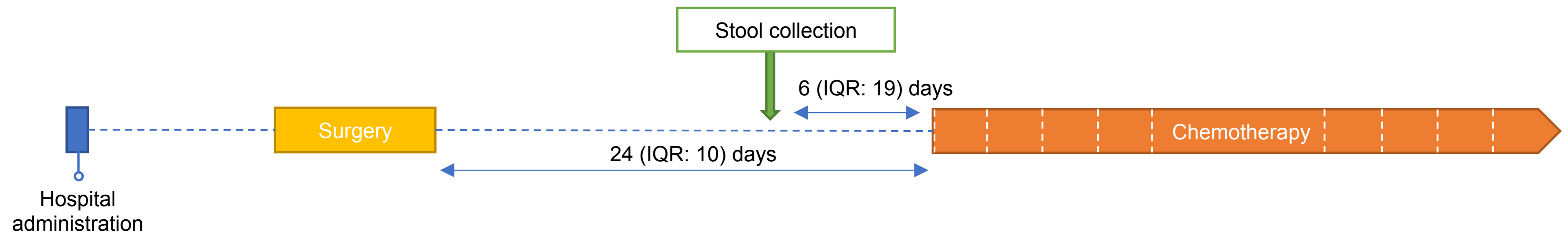
In this study, 167 stool samples (55.5%) were collected after a breast cancer surgery, whereas 134 stool samples (44.5%) were collected before breast cancer surgery or before the initiation of chemotherapy among patients who received neoadjuvant chemotherapy and patients who received adjuvant chemotherapy without breast cancer surgery. The median time interval from the day of stool collection to the first date of the following treatment was approximately 6 (IQR: 9) days among the 167

patients whose stool samples were collected after breast cancer surgery and before chemotherapy and 6 (IQR: 5) days among the 93 patients whose stool samples were collected before breast cancer surgery, respectively. The median time interval from the day of stool collection to the first date of chemotherapy was approximately 7 (IQR: 10) days among 41 patients who received neoadjuvant chemotherapy and those who received adjuvant chemotherapy without breast cancer surgery (**Figure 13**).

### 1.5. Microbiome Profiling

DNA extraction, library preparation, sequencing, and data processing for 301 stool samples were conducted in one batch. Detailed descriptions are given in Aim 2. Sequencing reads were subjected to quality trimming via Trimmomatic v0.39 and Bowtie2 v2.3.0 was used to remove human reads.[229,230] Taxonomic profiling was conducted using Kraken v2.1.1 and Bracken v2.6 to estimate the absolute abundance of microbial taxa with human bacterial genomes from the UHGG collection as the reference.[231-233] The HUMAnN2 algorithm (v0.11.1) was utilized to perform functional profiling of the GI microbiome using clean reads and the UniRef90 complete protein database as a reference. We estimated the relative abundance of GI microbial metabolic pathways.[245]

**167 patients whose stool samples were collected after breast cancer surgery**




**134 patients whose stool samples were collected before breast cancer surgery & neoadjuvant chemotherapy**

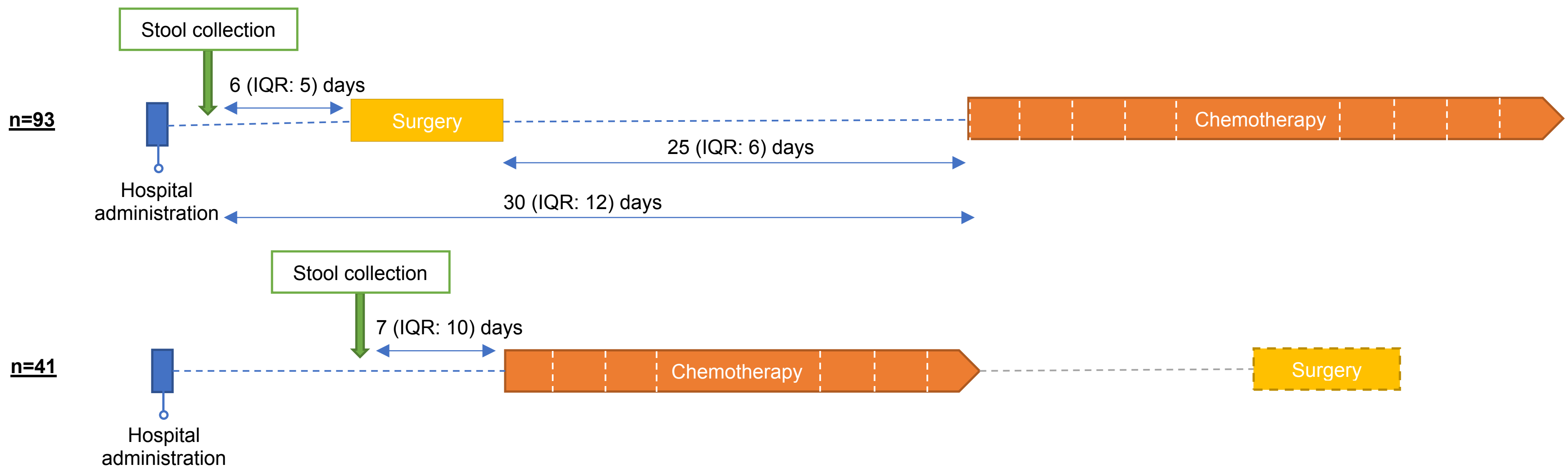



**Figure 13:** Timeline of stool collection

## 1.6. Covariate Assessment

Based on the literature review and the results of Aim 1, we selected a set of covariates, including non-clinical and clinical factors, described in **Table 37** as potential confounders to be adjusted in assessing the association between chemotherapy-induced toxicities and pre-chemotherapy GI microbiome.

**Table 37**: Summary of breast cancer patients' sociodemographic and clinical characteristics (Aim 3)

| | All eligible participants (N = 301) | | | All eligible participants (N = 301) | |
|---|---|---|---|---|---|
| | n | % | | n | % |
| **Age group** | | | **Pre-existing hematological condition** [b] | | |
| < 40 | 44 | 14.6 | No | 218 | 72.4 |
| 40-49 | 121 | 40.2 | Yes | 83 | 27.6 |
| 50-59 | 101 | 33.6 | **Pre-existing nephrological condition** [c] | | |
| 60+ | 35 | 11.6 | No | 240 | 79.7 |
| **Education levels** | | | Yes | 61 | 20.3 |
| Primary school | 48 | 16.0 | **Pre-existing hepatological condition** [d] | | |
| Middle school | 133 | 44.2 | No | 251 | 83.4 |
| High school | 68 | 22.6 | Yes | 50 | 16.6 |
| College or higher | 52 | 17.3 | **Breast cancer subtypes** | | |
| **Income levels** | | | Luminal/HER2-negative | 120 | 39.9 |
| Low (T1) | 109 | 36.2 | Luminal/HER2-positive | 73 | 24.2 |
| Middle (T2) | 94 | 31.2 | HER2 enriched | 68 | 22.6 |
| High (T3) | 98 | 32.6 | Triple-negative | 40 | 13.3 |
| **Residence** | | | **TNM stage** | | |
| Urban area | 111 | 36.9 | Stage I | 59 | 19.6 |
| Rural area | 190 | 63.1 | Stage II | 168 | 55.8 |
| **Menopausal status** | | | Stage III-IV | 74 | 24.6 |
| Pre-menopausal | 174 | 57.8 | **Diagnosis delay** | | |
| Post-menopausal | 127 | 42.2 | No delay | 212 | 53.5 |
| **BMI levels (**kg/m$^2$) | | | Moderate delay | 114 | 28.8 |
| Underweight (<18.5) | 29 | 9.6 | Serious delay | 70 | 17.7 |
| Normal weight (18.5-22.9) | 193 | 64.1 | **Sequential anthracycline and taxane** | | |
| Overweight (23-24.9) | 54 | 17.9 | No | 89 | 29.6 |
| Obese (≥25) | 25 | 8.3 | Yes | 212 | 70.4 |
| **Comorbidity** [a] | | | **Dose-dense chemotherapy** | | |
| No | 252 | 83.7 | No | 264 | 87.7 |
| Yes | 49 | 16.3 | Yes | 37 | 12.2 |

[a] Having diagnosis of specific comorbidities including diabetes mellitus, hypertension, hyperlipidemia, coronary heart disease (CHD), stroke, myocardial infarction, arthritis, lupus, and another chronic disease at the enrollment; [b] Having at least one of hematological symptoms (grade≥1) including anemia, neutropenia, lymphopenia, and thrombocytopenia before chemotherapy 120 days; [c] Having at least one of nephrological symptoms (grade≥1) including high creatinine, proteinuria, and hematuria before chemotherapy 120 days; [d] Having at least one of hepatological symptoms (grade≥1) including high bilirubin, SGOT, and SGPT before chemotherapy 120 days.

The mean age of the 301 study participants was 48.8 years at the time of diagnosis and treatment. Approximately 36.9% of patients lived in rural areas, and 39.9% of cases had attained a high school or college or higher education. The percentages of underweight (BMI <18.5 kg/m$^2$), overweight (BMI: 23-24.9 kg/m$^2$), and obese (BMI ≥25 kg/m$^2$) in our breast cancer patients were 9.6%, 17.9%, and 8.3%, respectively. Comorbidity was reported by approximately 16.3% of patients. Prior to chemotherapy, 27.6%, 20.3%, and 16.6% of breast cancer patients had pre-existing hematological, nephrological, or hepatological condition. Over half (55.8%) of participants were diagnosed at disease stage II, while19.6% were diagnosed at stage I, and 24.6% were diagnosed at stage III or later. The percentage of breast cancer patients who experienced moderate (4-8 months) and serious delays (≥ 9 months) in diagnosis was 28.8% and 17.7%, respectively. Sequential anthracycline and taxane treatment were the most common chemotherapy (70.4%) regimens, and approximately 12.2% of participants received dose-dense chemotherapy (**Table 37**).

1.7. Statistical Analysis

To evaluate the associations between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients, we first evaluated overall microbial richness (alpha diversity) and composition (beta diversity) with chemotherapy-induced toxicities. As mentioned in Aim 2, both alpha and beta diversity estimates might be affected by sequencing depth,[236] so we first rarefied the species level absolute abundance, i.e., read counts, of every sample to the minimum number of clean reads (n= 3,578,947) among 301 samples, using the R function *vegan::rarefy.*[237] Then, alpha diversity and beta diversity were calculated based on the rarefied species level absolute abundance data using the R functions *vegan::diversity* and *vegan::vegdist*, respectively.[237] In our study, alpha diversity was measured by the Chao1 index, Shannon index, and inversed Simpson diversity index. Tertile distributions of each alpha diversity index were used to categorize the measurement, with the lowest tertile serving as the reference group. Associations for alpha diversity indexes with chemotherapy-induced toxicities, including the combined hematological toxicity, combined GI toxicity, neutropenia, and nausea/vomiting, were evaluated by multivariable logistic analysis. ORs and respective 95% CIs were calculated in models adjusted for potential confounders. Covariates adjusted included age groups, income levels, residence, menopausal status, body mass index (continuous, kg/m$^2$), comorbidity, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis

delay, TNM cancer stage, breast cancer subtype, sequential anthracycline and taxane (yes/no), dose-dense chemotherapy (yes/no), fiber intake, fat intake, carbohydrate intake, physical activity (MET/wk) and stool sample collection time. Tests for trend were conducted using the median values of each tertile. We performed sub-analyses with stratification by stool collection time. Beta diversity was measured by Bray-Curtis dissimilarity matrix (Bray-Curtis), weighted UniFrac distance matrix (wUniFrac), and unweighted UniFrac distance (uwUniFrac). The PERMANOVA test was implemented to assess whether there was a difference regarding GI microbial composition according to chemotherapy-induced toxicities.[238] P values from PERMANOVA tests were produced in models adjusted for the aforementioned potential confounders and 999 permutation using the R functions *vegan.*[237] All statistical analyses were performed at two-sided tests, and associations with $P<0.05$ were considered statistically significant.

Second, we evaluated the associations of pre-chemotherapy GI microbial taxa with chemotherapy-induced toxicities. Logistic regression model was conducted to evaluate associations between chemotherapy-induced toxicities and the clr-transformed taxa abundance of common taxa and rare taxa, which have been described in Aim 2. We limited our analysis for rare taxa to those present in 10-50% of samples. ORs and corresponding 95% CI per SD increase in clr-transformed absolute abundance of microbial taxa were calculated in models adjusted for the aforementioned confounders. Analyses were conducted for all 301 breast cancer participants and stratified by stool collection time. False discovery rate (FDR) was calculated at each taxonomic level separately by common and rare taxa for overall analyses or stratified analyses to account for multiple testing. Association with FDR-corrected p-value ($P_{FDR}$) of <0.1 was considered statistically significant.

In addition, we evaluated the associations of GI microbial metabolic pathways with chemotherapy-induced toxicities via multivariable logistic regression analyses. We limited our analysis of GI microbial metabolic pathways to those present in >10% of samples. Arcsine square root (asr) transformation was used to normalize the relative abundance of microbial metabolic pathways. ORs and corresponding 95% CIs per SD increase in asr-transformed relative abundance of metabolic pathways were calculated in models adjusted for the confounders as mentioned above. For multiple testing on association with FDR-corrected p-value ($P_{FDR}$) of <0.1 was considered statistically significant. However, all FDR-corrected p-values were >0.1 in the analysis for

metabolic pathways. Thus, all associations with unadjusted *P*<0.05 were presented. All statistical analyses were performed using R version 3.6.3.

1.8. Statistical Power Calculation

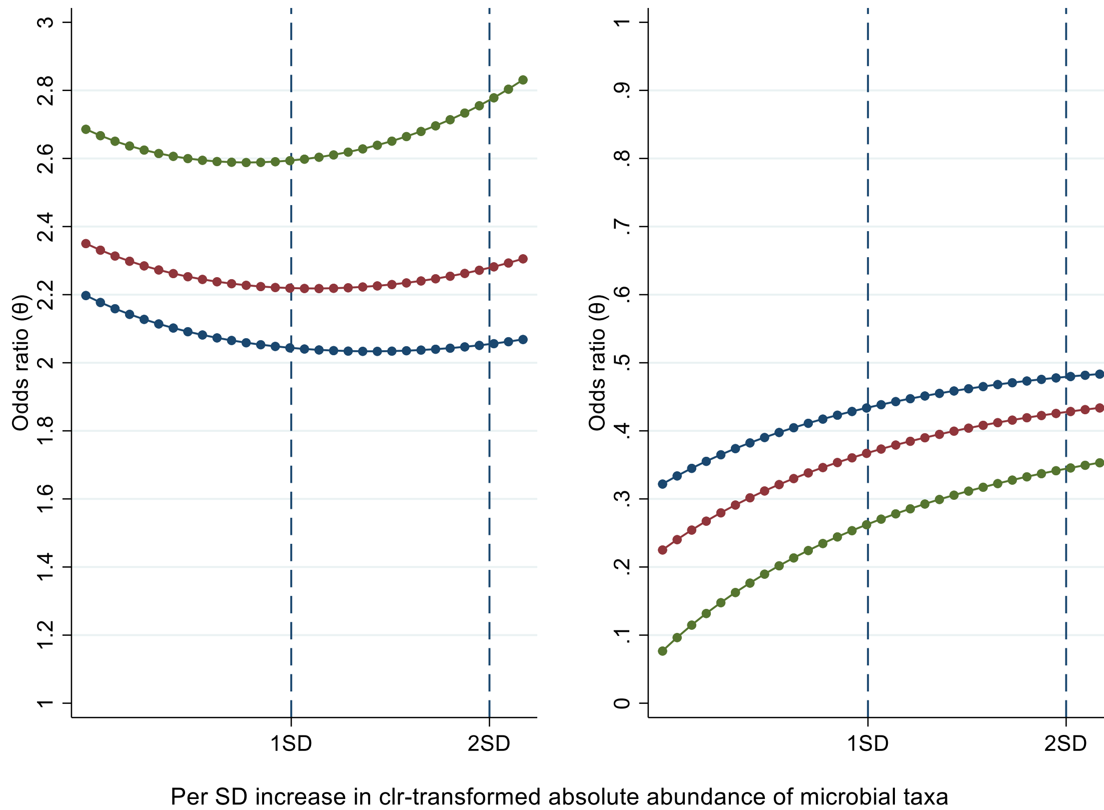



**Figure 14:** Estimated OR per SD increase in clr-transformed absolute abundance of microbial taxa by the prevalence of grade ≥ 3 chemotherapy-induced toxicity

(Green, red, and blue curves represent15%, 25% and 40% severe toxicity at 85% power)

Based on Aim 1 results, 15-40% of our population have had severe (grade ≥ 3) chemotherapy-induced toxicities among 301 breast cancer patients. The minimum detectable OR per SD increase in clr-transformed absolute abundance of microbial taxa ranges from 2.04 to 2.59 for OR >1, and the maximum detectable OR per SD increase ranges from 0.26 to 0.43 for OR<1 for severe chemotherapy-induced toxicities with a prevalence of 15% and 40% at 85% power and a 0.05 two-sided significance level. For example, our sample size of 301 with 39% severe combined hematological toxicity was able to detect an OR per SD increase in clr-transformed absolute abundance of microbial taxa greater than 2.05 or less than 0.43. (**Figure 14**). This calculation can be applied for OR per SD increase in asr-transformed relative abundance of metabolic pathways.

## 2. Results

**Table 38:** Incidence of ≥ 3 grade chemotherapy-induced toxicities by stool collection time

| | | | Stool collection time | | |
|---|---|---|---|---|---|
| | | Overall | Collection before breast cancer surgery & neoadjuvant chemotherapy | Collection after breast cancer surgery | |
| | | N=301 | N=134 | N=167 | *p* |
| | Grade | n (%) | n (%) | n (%) | *value* |
| **Hematological toxicity** | | | | | |
| Neutropenia | ≥ 3 | 92 (30.6) | 38 (28.4) | 54 (32.3) | 0.46 |
| Combined toxicity [a] | ≥ 3 | 118 (39.2) | 54 (40.3) | 64 (38.3) | 0.73 |
| **GI toxicity** | | | | | |
| Nausea/vomiting | ≥ 3 | 32 (10.6) | 15 (11.2) | 17 (10.2) | 0.78 |
| Combined toxicity [b] | ≥ 3 | 40 (13.3) | 19 (14.2) | 21 (12.6) | 0.68 |

[a] Combined hematological toxicity refers to having any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[b] Combined GI toxicity refers to having any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.

Among 301 breast cancer patients who donated a stool sample and received chemotherapy, the incidence rates of grade ≥ 3 combined hematological toxicity and combined GI toxicity were 39.2% and 13.3%, respectively. Neutropenia was the most common chemotherapy-induced hematological toxicities, with 30.6% of patients experiencing grade ≥ 3. Meanwhile, nausea/vomiting was the most common GI toxicities, with 10.6% of patients experiencing grade ≥ 3 among breast cancer patients. No significant differences were observed for all hematological and GI toxicities by stool collection time (**Table 38**).

**Association of GI microbial richness and composition with chemotherapy-induced toxicity**

After quality control and rarefaction, a total of 4206 OTUs was identified among 301 breast cancer patients, based on which alpha diversity indexes (Chao1, Shannon, and inverse Simpson indexes) and beta dissimilarity matrices (Bray-Curtis dissimilarity, weighted UniFrac distance, and unweighted UniFrac distance) were calculated.

Multivariable analyses showed that a high alpha diversity was significantly and inversely associated with the risk of grade ≥3 combined hematological toxicity and grade ≥3 neutropenia. The highest tertiles of Chao 1 and Shannon indexes were associated with significantly low OR of grade ≥3 combined hematological

toxicity ($OR_{T3vsT1}$ for Chao1 index: 0.47 (0.24-0.92); $P_{trend}$=0.025 and $OR_{T3vsT1}$ for Shannon index: 0.49 (0.24-0.98); $P_{trend}$=0.04). Chao1 index was also significantly associated with a reduced risk of grade ≥3 neutropenia. The adjusted ORs and 95% CIs for tertiles 2-3 vs. tertile 1 were 0.74 (0.38-1.46) and 0.38 (0.18-0.78), and $P_{trend}$= 0.009. No significant association with grade ≥3 neutropenia was found for the Shannon and inverse Simpson indexes.

Additional analyses stratified by stool collection time showed the highest tertile of Chao 1 was significantly associated with the risk of grade ≥3 combined hematological toxicity among patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, with $OR_{T3vsT1}$ and 95% CIs of 0.26 (0.07-0.97); $P_{trend}$=0.016. Furthermore, in this group of patients, all three alpha diversity was significantly and inversely associated with the risk of grade ≥3 neutropenia. Similar association patterns were among patients with stool samples collected after breast cancer surgery although none of point estimates reached statistical significance (**Tables 39-40**).

None of the alpha diversity indexes were not significantly associated with the risk of grade ≥3 combined GI toxicity. The adjusted ORs and 95% CIs for tertiles 2-3 vs. tertile 1 in the overall analyses for Chao1 index were: 1.61 (0.61-4.22), 2.15 (0.83-5.60) and $P_{trend}$ =0.12; Shannon index: 1.07 (0.41-2.79), 1.73 (0.68-4.40), and $P_{trend}$ =0.24; and inverse Simpson index: 0.68 (0.26-1.73), 1.32 (0.57-3.08), and $P_{trend}$ =0.36. Likewise, there were no significant associations between grade ≥3 nausea/vomiting and alpha diversity indexes, except for the Chao1 index in the overall analysis. A high Chao1 index was significantly associated with an increased risk of nausea/vomiting. The adjusted ORs and 95% CIs for tertiles 2-3 vs. tertile 1 were 3.99 (1.08-14.69) and 6.59 (1.78-24.39), and $P_{trend}$ = 0.005. This association did not appear to vary by time of stool sample collection (**Tables 41-42**).

**Table 39**: Association of grade ≥3 combined hematological toxicity with alpha diversity indexes by stool collection time

| | Stool collection time | | | | | |
|---|---|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | | **Overall** | |
| | N=134 | | N=167 | | N=301 | |
| | *No.* of grade ≥3/ grade <3 | *OR (95%CI)*[1] | *No.* of grade ≥3/ grade <3 | *OR (95%CI)*[2] | *No.* of grade ≥3/ grade <3 | *OR (95%CI)*[3] |
| **Chao1 index** | | | | | | |
| T1 | 13/ 18 | 1.00 | 31/ 39 | 1.00 | 44/ 57 | 1.00 |
| T2 | 23/ 25 | 1.46 (0.43-4.92) | 16/ 36 | 0.78 (0.30-2.06) | 39/ 61 | 0.81 (0.42-1.55) |
| T3 | 18/ 37 | 0.26 (0.07-0.97) | 17/ 28 | 0.56 (0.21-1.49) | 35/ 65 | 0.47 (0.24-0.92) |
| P for trend | | 0.016 | | 0.25 | | 0.025 |
| **Shannon Index** | | | | | | |
| T1 | 12/ 16 | 1.00 | 30/ 43 | 1.00 | 42/ 59 | 1.00 |
| T2 | 26/ 25 | 1.47 (0.42-5.13) | 17/ 32 | 0.79 (0.3-2.04) | 43/ 57 | 1.07 (0.55-2.06) |
| T3 | 16/ 39 | 0.27 (0.07-1.08) | 17/ 28 | 0.70 (0.26-1.85) | 33/ 67 | 0.49 (0.24-0.98) |
| P for trend | | 0.02 | | 0.45 | | 0.04 |
| **Inverse Simpson index** | | | | | | |
| T1 | 16/ 20 | 1.00 | 25/ 40 | 1.00 | 41/ 60 | 1.00 |
| T2 | 21/ 23 | 2.14 (0.67-6.84) | 25/ 31 | 1.53 (0.61-3.84) | 46/ 54 | 1.41 (0.74-2.67) |
| T3 | 17/ 37 | 0.38 (0.11-1.33) | 14/ 32 | 0.47 (0.16-1.41) | 31/ 69 | 0.52 (0.26-1.03) |
| P for trend | | 0.027 | | 0.12 | | 0.018 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Table 40**: Association of grade ≥3 neutropenia with alpha diversity indexes by stool collection time

| | **Stool collection time** | | | | | |
|---|---|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | | **Overall** | |
| | N=134 | | N=167 | | N=301 | |
| | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[1] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[2] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[3] |
| **Chao1 index** | | | | | | |
| T1 | 9/ 22 | 1.00 | 28/ 42 | 1.00 | 37/ 64 | 1.00 |
| T2 | 18/ 30 | 1.23 (0.26-5.75) | 13/ 39 | 0.69 (0.25-1.92) | 31/ 69 | 0.74 (0.38-1.46) |
| T3 | 11/ 44 | 0.06 (0.01-0.40) | 13/ 32 | 0.41 (0.14-1.14) | 24/ 76 | 0.39 (0.19-0.79) |
| P for trend | | 0.002 | | 0.08 | | 0.009 |
| **Shannon Index** | | | | | | |
| T1 | 8/ 20 | 1.00 | 26/ 47 | 1.00 | 34/ 67 | 1.00 |
| T2 | 19/ 32 | 1.05 (0.21-5.11) | 14/ 35 | 0.68 (0.25-1.84) | 33/ 67 | 1.05 (0.53-2.08) |
| T3 | 11/ 44 | 0.09 (0.01-0.61) | 14/ 31 | 0.64 (0.23-1.77) | 25/ 75 | 0.51 (0.24-1.05) |
| P for trend | | 0.007 | | 0.35 | | 0.07 |
| **Inverse Simpson index** | | | | | | |
| T1 | 11/ 25 | 1.00 | 23/ 42 | 1.00 | 34/ 67 | 1.00 |
| T2 | 14/ 30 | 1.21 (0.30-4.87) | 19/ 37 | 1.14 (0.45-2.92) | 33/67 | 1.02 (0.53-1.97) |
| T3 | 13/ 41 | 0.19 (0.04-0.93) | 12/ 34 | 0.44 (0.14-1.35) | 25/75 | 0.53 (0.26-1.09) |
| P for trend | | 0.02 | | 0.12 | | 0.05 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Table 41**: Association of grade ≥3 combined GI toxicity with alpha diversity indexes by stool collection time

| | Stool collection time | | | | Overall | |
|---|---|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | | | |
| | N=134 | | N=167 | | N=301 | |
| | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[1] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[2] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[3] |
| **Chao1 index** | | | | | | |
| T1 | 3/ 28 | 1.00 | 6/ 64 | 1.00 | 9/ 92 | 1.00 |
| T2 | 9/ 39 | 2.61 (0.44-15.46) | 5/ 47 | 2.96 (0.46-18.86) | 14/ 86 | 1.61 (0.61-4.22) |
| T3 | 7/ 48 | 1.87 (0.28-12.32) | 10/ 35 | 1.24 (0.16-9.41) | 17/ 83 | 2.15 (0.83-5.60) |
| P for trend | | 0.62 | | 0.95 | | 0.12 |
| **Shannon Index** | | | | | | |
| T1 | 4/ 24 | 1.00 | 7/ 66 | 1.00 | 11/ 90 | 1.00 |
| T2 | 8/ 43 | 0.96 (0.14-6.48) | 5/ 44 | 0.85 (0.12-6.11) | 13/ 87 | 1.07 (0.41-2.79) |
| T3 | 7/ 48 | 1.24 (0.20-7.82) | 9/ 36 | 0.88 (0.12-6.19) | 16/ 84 | 1.73 (0.68-4.40) |
| P for trend | | 0.77 | | 0.92 | | 0.24 |
| **Inverse Simpson index** | | | | | | |
| T1 | 7/ 29 | 1.00 | 8/ 57 | 1.00 | 15/ 86 | 1.00 |
| T2 | 5/ 39 | 0.75 (0.15-3.75) | 4/ 52 | 0.66 (0.12-3.68) | 9/ 91 | 0.68 (0.27-1.73) |
| T3 | 7/ 47 | 0.86 (0.18-4.12) | 9/ 37 | 0.88 (0.17-4.52) | 16/ 84 | 1.32 (0.57-3.08) |
| P for trend | | 0.82 | | 0.87 | | 0.36 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

**Table 42:** Association of grade ≥3 nausea/vomiting with alpha diversity indexes by stool collection time

| | Stool collection time | | | | Overall | |
|---|---|---|---|---|---|---|
| | Collection before breast cancer surgery & neoadjuvant chemotherapy | | Collection after breast cancer surgery | | | |
| | N=134 | | N=167 | | N=301 | |
| | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[1] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[2] | *No*. of grade ≥3/ grade <3 | *OR (95%CI)*[3] |
| **Chao1 index** | | | | | | |
| T1 | 1/ 30 | 1.00 | 3/ 67 | 1.00 | 4/ 97 | 1.00 |
| T2 | 8/ 40 | 16.59 (0.62-443) | 4/ 48 | 19.38 (0.62-606) | 12/ 88 | 3.99 (1.08-14.69) |
| T3 | 6/ 49 | 14.38 (0.46-445) | 10/ 35 | 13.85 (0.39-495) | 16/ 84 | 6.59 (1.78-24.39) |
| P for trend | | 0.19 | | 0.22 | | 0.005 |
| **Shannon Index** | | | | | | |
| T1 | 3/ 25 | 1.00 | 5/ 68 | 1.00 | 8/ 93 | 1.00 |
| T2 | 6/ 45 | 0.78 (0.07-8.50) | 3/ 46 | 0.77 (0.07-8.50) | 9/ 91 | 1.03 (0.32-3.32) |
| T3 | 6/ 49 | 1.63 (0.17-15.36) | 9/ 36 | 1.55 (0.16-15.07) | 15/ 85 | 2.63 (0.90-7.70) |
| P for trend | | 0.56 | | 0.59 | | 0.06 |
| **Inverse Simpson index** | | | | | | |
| T1 | 5/ 31 | 1.00 | 6/ 59 | 1.00 | 11/ 90 | 1.00 |
| T2 | 3/ 41 | 0.61 (0.08-4.70) | 2/ 54 | 0.65 (0.08-5.21) | 5/ 95 | 0.5 (0.15-1.67) |
| T3 | 7/ 47 | 1.7 (0.25-11.36) | 9/ 37 | 1.9 (0.27-13.29) | 16/ 84 | 2.17 (0.84-5.65) |
| P for trend | | 0.53 | | 0.46 | | 0.04 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

The overall analysis showed that 0.4% of variations in the weighted UniFracs distance matrix were associated with the experience of grade ≥ 3 combined GI toxicity among 301 patients ($P_{wUniFrac}$=0.003). In addition, 0.8% of variations in beta-diversity were associated with the experience of grade ≥ 3 nausea/vomiting among 301 patients ($P_{Bray\text{-}Curtis}$=0.005; $P_{uwUniFrac}$=0.009; and $P_{wUniFrac}$=0.008). A significant difference was consistently found for the Bray-Curtis dissimilarity matrix and unweighted UniFracs distance matrix with 1.1-1.2% variations ($P_{Bray\text{-}Curtis}$=0.033; $P_{uwUniFrac}$=0.028) among patients with stool samples collected after breast cancer surgery. However, no differences were found in pre-chemotherapy beta diversity between grade <3 and grade ≥3 chemotherapy-induced toxicities, including the combined hematological toxicity, the combined GI

toxicity, and neutropenia, in the overall analysis or stratified analyses by stool collection time (all *p for PERMANOVA* >0.05) (**Table 43**).

**Table 43**: PERMANOVA test difference of pre-chemotherapy beta diversity between chemotherapy-induced toxicities (grade <3 vs. grade ≥3)

| | Bray-Curtis dissimilarity matrix | | Unweighted UniFracs distance matrix | | Weighted UniFracs distance matrix | |
|---|---|---|---|---|---|---|
| | *$R^2$(p for PERMANOVA test)* | | *$R^2$(p for PERMANOVA test)* | | *$R^2$(p for PERMANOVA test)* | |
| | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| **Overall (N=301)[1]** | | | | | | |
| Combine hematological toxicity | 0.4% (0.19) | 0.4% (0.17) | 0.4% (0.23) | 0.4% (0.21) | 0.5% (0.13) | 0.5% (0.11) |
| Neutropenia | 0.4% (0.26) | 0.4% (0.21) | 0.5% (0.07) | 0.5% (0.06) | 0.4% (0.20) | 0.4% (0.24) |
| Combined GI toxicity | 0.5% (0.12) | 0.5% (0.11) | 0.4% (0.12) | 0.4% (0.12) | 0.4% (0.03) | 0.4% (0.03) |
| Nausea/vomiting | 0.8% (0.007) | 0.8% (0.005) | 0.8% (0.01) | 0.8% (0.009) | 0.8% (0.01) | 0.8% (0.008) |
| **Stool collection before breast cancer surgery and neoadjuvant chemotherapy (N=134)[2]** | | | | | | |
| Combine hematological toxicity | 0.8% (0.30) | 0.8% (0.25) | 1.0% (0.12) | 1.0% (0.11) | 0.8% (0.40) | 0.8% (0.32) |
| Neutropenia | 0.8% (0.33) | 0.8% (0.28) | 1.1% (0.06) | 1.1% (0.05) | 0.8% (0.39) | 0.8% (0.33) |
| Combined GI toxicity | 0.8% (0.31) | 0.8% (0.27) | 0.7% (0.44) | 0.7% (0.39) | 0.8% (0.36) | 0.8% (0.30) |
| Nausea/vomiting | 1.0% (0.17) | 1.0% (0.15) | 0.9% (0.21) | 0.9% (0.19) | 0.9% (0.27) | 0.9% (0.22) |
| **Stool collection after breast cancer surgery (N=167)[3]** | | | | | | |
| Combine hematological toxicity | 0.7% (0.24) | 0.7% (0.22) | 0.7% (0.29) | 0.7% (0.30) | 0.7% (0.23) | 0.7% (0.22) |
| Neutropenia | 0.7% (0.29) | 0.7% (0.28) | 0.7% (0.20) | 0.7% (0.20) | 0.7% (0.23) | 0.7% (0.22) |
| Combined GI toxicity | 0.6% (0.39) | 0.6% (0.39) | 0.7% (0.19) | 0.7% (0.19) | 0.4% (0.76) | 0.4% (0.75) |
| Nausea/vomiting | 1.1% (0.034) | 1.1% (0.033) | 1.2% (0.026) | 1.2% (0.028) | 0.9% (0.12) | 0.9% (0.12) |

Model 1 was adjusted for age group, TNM cancer stage, sequential anthracycline and taxane, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[1] Model 2 was the multivariable model 1 with additional adjustment for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, and breast cancer subtypes.

[2] Model 2 was the multivariable model 1 with additional adjustment for income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, and breast cancer subtypes.

[3] Model 2 was the multivariable model 1 with additional adjustment for the number of days from stool collection time to breast cancer surgery (days), income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, and breast cancer subtypes.

**Association of GI microbial taxa with chemotherapy-induced toxicity**

Among 301 stool samples, a total of 21 phyla, 29 classes, 77 orders, 244 families, 1278 genera, and 4206 species were identified and estimated absolute abundance. After excluding rare taxa with a prevalence of < 10% in 301 samples, a total of 17 phyla, 23 classes, 53 orders, 135 families, 628 genera, and 1,954 species were included to evaluate the associations of GI microbial taxa with chemotherapy-induced toxicities.

**Table 44** presents significant associations of GI microbial taxa with the combined hematological toxicity after FDR correction ($P_{FDR} < 0.1$). We found that two of the three rare taxa were associated with a reduced risk of grade ≥3 combined hematological toxicity, including two families *UBA1255* (belonging to the class *Clostridia* within *phylum Firmicutes A*), and *CAG-312* (belonging to the order *Opitutales* within phylum *Verrucomicrobiota)* among patients whose stool samples were collected before surgery and neoadjuvant chemotherapy. The abundance of family UBA1255 and *CAG-312* were inversely associated with the risk of grade ≥3 combined hematological toxicity, with OR of 0.50 (95% CI: 0.29-0.84) and 0.52 (0.32-0.84); and all $P_{FDR}$=0.092, respectively. Analysis for patients with stool samples collected before surgery and neoadjuvant chemotherapy also showed that the family *CAG-826* (belonging to phylum *Firmicutes*) was associated with an increased risk of grade ≥3 combined hematological toxicity. The abundance of family *CAG-826 was* associated with a higher risk of grade ≥3 combined hematological toxicity, with OR of 2.00 (95% CI: 1.22-3.28) and $P_{FDR}$=0.092. A significant association ($P_{FDR} < 0.1$) of the family *CAG-312* with grade ≥3 combined hematological toxicity was observed for grade ≥3 neutropenia (OR =0.33, 95% CI: 0.16-0.68; $P_{FDR}$=0.081). No significant associations after FDR correction ($P_{FDR} < 0.1$) with the combined hematological toxicity for gut microbial taxa were observed in the overall analysis and the analysis for patients with stool samples collected after breast cancer surgery.

In the overall analysis, we found a significant association with a reduced risk of grade ≥3 neutropenia for the family *Synergistaceae* (OR=0.59, 95% CI: 0.42-0.83; $P_{FDR}$=*0.075*). Many significant associations between grade ≥3 neutropenia and GI microbial taxa after FDR correction were observed, almost all among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. We found that 47 common taxa were associated with the risk of grade ≥3 neutropenia in

this sub-population. In the phylum *Bacteroidota*, the family *Bacteroidaceae* and its four species were associated with an increased risk of grade ≥3 neutropenia, including *Bacteroides A mediterraneensis* and three *MGYG-HGUT species (03163, 02275,* and *01240)*. The family *Bacteroidaceae* was associated with an increased risk of grade ≥3 neutropenia (OR=4.39, 95% CI: 1.80-10.71; *P*=0.001; and $P_{FDR}$=0.057), which was driven by species *Bacteroides A mediterraneensis* (OR= 3.44, 95% CI: 1.60-7.40; *P*=0.002 and $P_{FDR}$=0.094). Likewise, in phylum *Firmicutes A*, the abundance of family *Sporanaerobacteraceae* (i.e., a member of the order *Tissierellales*) and 11 species were positively associated with the risk of grade ≥3 neutropenia. The adjusted ORs and 95% CIs for the family *Sporanaerobacteraceae* were 3.28 (1.49-7.23), *P*=0.003 and $P_{FDR}$=0.057, which was driven by the species *Sporanaerobacter acetigenes* (OR=2.94, 95% CI: 1.39-6.18; *P*=0.005, and $P_{FDR}$=0.094). Among ten species belonging to the family *Lachnospiraceae* (e.g., *Blautia hansenii, Blautia producta, Dorea scindens* and *Tyzzerella sp000411335),* the species *Blautia sp003287895,* showed the strongest association, with an OR of 5.04 (95% CI: 2.11-12.04), *P*=$2.69 \times 10^{-4}$; $P_{FDR}$=0.094. Moreover, species *MGYG-HGUT-00794* (within phylum *Firmicutes C*) were associated with increased grade ≥3 neutropenia (OR=2.43, 95% CI: 1.31-4.51, *P*=0.005, and $P_{FDR}$=0.094). Furthermore, in phylum *Fusobacteriota*, we found a significantly increased association with grade ≥3 neutropenia for the family *Fusobacteriaceae* (OR=2.67, 95% CI: 1.35-5.25, *P*=0.005; and $P_{FDR}$=0.061), which was driven by their species *MGYG-HGUT-03919* (OR=2.52, 95% CI: 1.35-4.72, *P*=0.004; and $P_{FDR}$=0.094).

Conversely, we also found that two families, *QALW01* (within the order *Christensenellales*) and *CAG-272* (within the order *Oscillospirales*), and 25 species (e.g., *Coprococcus eutactus, Dorea scindens, Eubacterium E hallii A, Eubacterium G ventriosum, Intestinimonas butyriciproducens, Faecalibacterium prausnitzii J, and Ruminococcus D bicirculans*) belonging to the phylum *Firmicutes A* were significantly associated with a reduced risk of grade ≥3 neutropenia among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. The adjusted ORs and 95% CIs for the family *QALW01* were 0.40 (0.22-0.74), *P*=0.003, and $P_{FDR}$=0.057 and for the family *CAG-272* were 0.43 (0.23-0.80), *P*=0.008, and $P_{FDR}$=0.084. A high abundance of *Faecalibacterium prausnitzii J* was associated with a reduced risk of grade ≥3 neutropenia, with an OR of 0.37 (95% CI: 0.19-0.73), *P*=0.004, and $P_{FDR}$=0.094. Moreover, a significant inverse association with grade ≥3 neutropenia was

observed for the family *Victivallaceae* (within phylum *Verrucomicrobiota*) with an OR of 0.53 (95% CI: 0.35-0.81); $P$=0.003; and $P_{FDR}$=0.061. Finally, no significant associations between GI microbiota taxa with grade ≥3 neutropenia were observed among patients whose stool samples were collected after breast cancer surgery. Similar association patterns with the risk of grade ≥3 neutropenia were observed for 25 of the 47 taxa mention above, which had significant associations with grade ≥3 neutropenia among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy (**Table 45**).

**Table 44**: Association of grade ≥3 combined hematological toxicity with pre-chemotherapy GI microbial taxa by stool collection time

| | | | Stool collection time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)[1] | | | Collection after breast cancer surgery (N= 194)[2] | | | Overall (N=301)[3] | | |
| *Taxonomy* | RA, median (%) | Pre (%) | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Class *Clostridia* | | | | | | | | | | | |
| Family *UBA1255* ┼ | 0.003¥ | 38.9 | 0.50 (0.29-0.84) | 0.009 | 0.092 | 0.85 (0.56-1.29) | 0.445 | 0.889 | 0.68 (0.51-0.91) | 0.008 | 0.132 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Family *CAG-826* ┼ | 0.003¥ | 20.3 | 2.00 (1.22-3.28) | 0.006 | 0.092 | 0.95 (0.63-1.43) | 0.802 | 0.907 | 1.25 (0.95-1.65) | 0.105 | 0.375 |
| Phylum *Verrucomicrobiota* | | | | | | | | | | | |
| Family *CAG-312* ┼↓ | 0.004¥ | 16.6 | 0.52 (0.32-0.84) | 0.008 | 0.092 | 1.52 (1.02-2.25) | 0.038 | 0.368 | 0.91 (0.70-1.18) | 0.468 | 0.833 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}<0.1$ is considered statistically significant.
↓ : significantly decreased association with stool collection time ($P_{FDR}<0.1$)
↑: significantly increased association with stool collection time ($P_{FDR}<0.1$)
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 45:** Association of grade ≥3 neutropenia with pre-chemotherapy GI microbial taxa by stool collection time

| | | | **Stool collection time** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)[1] | | | Collection after breast cancer surgery (N= 194)[2] | | | **Overall** (N=301)[3] | | |
| *Taxonomy* | RA, median (%) | Pre (%) | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] |
| Phylum *Bacteroidota* | | | | | | | | | | | |
| Family *Bacteroidaceae* | 51.180 | 100.0 | 4.39 (1.8-10.71) | 0.001 | 0.057 | 1.43 (0.92-2.22) | 0.116 | 0.509 | 1.54 (1.11-2.14) | 0.010 | 0.203 |
| Species *Bacteroides A mediterraneensis* ↑ | 51.180 | 100.0 | 3.44 (1.60-7.40) | 0.002 | 0.094 | 1.30 (0.85-2.01) | 0.228 | 0.938 | 1.58 (1.16-2.17) | 0.004 | 0.406 |
| Species *MGYG-HGUT-03163* ↑ | 0.016 | 92.7 | 3.09 (1.60-5.97) | 0.001 | 0.094 | 0.87 (0.57-1.33) | 0.522 | 0.942 | 1.31 (0.98-1.75) | 0.072 | 0.511 |
| Species *MGYG-HGUT-02275* | 0.010 | 92.7 | 3.03 (1.55-5.91) | 0.001 | 0.094 | 0.93 (0.61-1.42) | 0.730 | 0.981 | 1.31 (0.99-1.73) | 0.056 | 0.496 |
| Species *MGYG-HGUT-01240* ↓ | 0.021 | 97.0 | 2.90 (1.41-5.94) | 0.004 | 0.094 | 1.13 (0.73-1.75) | 0.596 | 0.972 | 1.54 (1.13-2.08) | 0.005 | 0.406 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Class *Clostridia* | | | | | | | | | | | |
| Order *Christensenellales* | | | | | | | | | | | |
| Family *QALW01* | 0.001 | 50.5 | 0.40 (0.22-0.74) | 0.003 | 0.057 | 0.99 (0.66-1.49) | 0.966 | 0.981 | 0.8 (0.60-1.07) | 0.137 | 0.405 |
| Order *Lachnospirales* | | | | | | | | | | | |
| Family *Lachnospiraceae* | | | | | | | | | | | |
| Species *MGYG-HGUT-04359* | 0.002 | 68.4 | 3.68 (1.64-8.24) | 0.002 | 0.094 | 1.11 (0.74-1.66) | 0.627 | 0.972 | 1.30 (0.98-1.74) | 0.073 | 0.511 |
| Species *Blautia_hansenii* ↑ | 0.011 | 92.4 | 3.07 (1.53-6.14) | 0.002 | 0.094 | 1.03 (0.69-1.54) | 0.897 | 0.990 | 1.21 (0.90-1.61) | 0.208 | 0.614 |
| Species *Blautia_producta* ↑ | 0.002 | 74.1 | 2.59 (1.33-5.04) | 0.005 | 0.094 | 0.81 (0.51-1.29) | 0.376 | 0.942 | 1.01 (0.75-1.35) | 0.963 | 0.979 |
| Species *Blautia_sp003287895* ↑ | 0.002 | 62.1 | 5.04 (2.11-12.04) | $2.69 \times 10^{-4}$ | 0.094 | 1.10 (0.74-1.62) | 0.651 | 0.981 | 1.31 (0.99-1.74) | 0.060 | 0.496 |
| Species *MGYG-HGUT-00913* ↑ | 0.002 | 61.5 | 2.89 (1.39-5.98) | 0.004 | 0.094 | 1.06 (0.72-1.57) | 0.754 | 0.981 | 1.19 (0.90-1.59) | 0.227 | 0.640 |
| Species *MGYG-HGUT-02598* ↑ | 0.021 | 97.0 | 2.71 (1.35-5.44) | 0.005 | 0.094 | 0.95 (0.62-1.45) | 0.820 | 0.982 | 1.04 (0.78-1.39) | 0.794 | 0.930 |
| Species *CAG-127 sp900319515*↓ | 0.002 | 64.1 | 0.27 (0.12-0.61) | 0.002 | 0.094 | 0.95 (0.60-1.51) | 0.831 | 0.982 | 0.66 (0.47-0.92) | 0.015 | 0.468 |
| Species *CAG-95_sp000438155* | 0.002 | 58.1 | 0.30 (0.14-0.63) | 0.001 | 0.094 | 1.39 (0.94-2.05) | 0.102 | 0.867 | 0.93 (0.70-1.23) | 0.597 | 0.861 |
| Species *Coprococcus_eutactus* ↓ | 0.003 | 72.8 | 0.31 (0.14-0.68) | 0.004 | 0.094 | 0.97 (0.65-1.47) | 0.898 | 0.990 | 0.70 (0.52-0.96) | 0.025 | 0.468 |
| Species *Coprococcus eutactus A* ↓ | 0.001 | 51.2 | 0.30 (0.13-0.67) | 0.003 | 0.094 | 1.41 (0.91-2.20) | 0.127 | 0.900 | 0.92 (0.69-1.25) | 0.605 | 0.867 |
| Species *Dorea_scindens* ↑ | 0.004 | 88.4 | 3.44 (1.53-7.71) | 0.003 | 0.094 | 0.86 (0.56-1.32) | 0.491 | 0.942 | 1.20 (0.90-1.61) | 0.221 | 0.630 |
| Species *Eubacterium E hallii A* ↓ | 0.009 | 71.4 | 0.25 (0.11-0.55) | 0.001 | 0.094 | 0.98 (0.63-1.52) | 0.925 | 0.993 | 0.69 (0.51-0.95) | 0.024 | 0.468 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Species *Eubacterium G sp000434315* ↓ | 0.001 | 51.8 | 0.29 (0.13-0.68) | 0.004 | 0.094 | 1.39 (0.94-2.05) | 0.098 | 0.857 | 0.92 (0.69-1.23) | 0.576 | 0.850 |
| Species *Eubacterium G ventriosum* ↓ | 0.002 | 57.1 | 0.37 (0.19-0.71) | 0.003 | 0.094 | 1.38 (0.90-2.12) | 0.140 | 0.900 | 0.84 (0.62-1.12) | 0.237 | 0.647 |
| Species *MGYG-HGUT-03065* ↑ | 0.012 | 90.0 | 2.84 (1.43-5.61) | 0.003 | 0.094 | 1.22 (0.83-1.81) | 0.308 | 0.938 | 1.41 (1.05-1.88) | 0.021 | 0.468 |
| Species *Marvinbryantia sp900066075* ↓ | 0.005 | 82.1 | 0.32 (0.15-0.69) | 0.004 | 0.094 | 1.79 (1.14-2.80) | 0.011 | 0.652 | 1.04 (0.78-1.39) | 0.785 | 0.930 |
| Species *MGYG-HGUT-00574* | 0.017 | 91.0 | 0.39 (0.21-0.74) | 0.004 | 0.094 | 1.35 (0.84-2.16) | 0.211 | 0.938 | 0.99 (0.74-1.32) | 0.935 | 0.973 |
| Species *MGYG-HGUT-00859* ↑ | 0.001 | 55.5 | 3.61 (1.75-7.45) | 0.001 | 0.094 | 0.89 (0.59-1.34) | 0.590 | 0.972 | 1.14 (0.86-1.50) | 0.373 | 0.740 |
| Species *Tyzzerella sp000411335* | 0.002 | 70.4 | 2.27 (1.29-3.99) | 0.005 | 0.094 | 0.98 (0.64-1.51) | 0.934 | 0.995 | 1.18 (0.90-1.56) | 0.238 | 0.647 |
| Species *MGYG-HGUT-00137* ↓ | 0.012 | 89.0 | 0.23 (0.10-0.54) | 0.001 | 0.094 | 1.44 (0.91-2.28) | 0.117 | 0.890 | 0.87 (0.66-1.15) | 0.329 | 0.722 |
| Order *Oscillospirales* | | | | | | | | | | | |
| Family *CAG-272* | 0.013 | 93.0 | 0.43 (0.23-0.80) | 0.008 | 0.084 | 0.67 (0.43-1.05) | 0.078 | 0.455 | 0.66 (0.49-0.90) | 0.008 | 0.203 |
| Family *Oscillospiraceae* | | | | | | | | | | | |
| Species *CAG-110_sp003525905* | 0.002 | 54.5 | 0.36 (0.18-0.71) | 0.003 | 0.094 | 0.93 (0.61-1.41) | 0.733 | 0.981 | 0.82 (0.62-1.09) | 0.178 | 0.590 |
| Species *MGYG-HGUT-02115* | 0.003 | 68.8 | 0.32 (0.16-0.65) | 0.002 | 0.094 | 0.99 (0.64-1.54) | 0.963 | 0.995 | 0.80 (0.59-1.08) | 0.141 | 0.544 |
| Species *CAG-83_sp000435975* ↓ | 0.004 | 64.5 | 0.32 (0.15-0.69) | 0.003 | 0.094 | 0.98 (0.64-1.49) | 0.920 | 0.993 | 0.81 (0.60-1.08) | 0.150 | 0.552 |
| Species *MGYG-HGUT-00713* ↓ | 0.001 | 50.2 | 0.30 (0.13-0.65) | 0.003 | 0.094 | 0.82 (0.54-1.25) | 0.353 | 0.942 | 0.76 (0.56-1.02) | 0.068 | 0.511 |
| Species *MGYG-HGUT-02143* | 0.002 | 60.1 | 0.28 (0.12-0.66) | 0.004 | 0.094 | 1.12 (0.72-1.73) | 0.616 | 0.972 | 0.78 (0.57-1.06) | 0.115 | 0.523 |
| Species *MGYG-HGUT-02724* | 0.002 | 58.1 | 0.34 (0.16-0.71) | 0.004 | 0.094 | 0.94 (0.61-1.45) | 0.770 | 0.981 | 0.84 (0.63-1.12) | 0.238 | 0.647 |
| Species *Intestinimonas butyriciproducens* ↑ | 0.004 | 75.1 | 0.39 (0.21-0.74) | 0.004 | 0.094 | 1.01 (0.64-1.58) | 0.977 | 0.995 | 0.80 (0.59-1.08) | 0.152 | 0.555 |
| Species *MGYG-HGUT-03668* | 0.001 | 51.2 | 0.38 (0.19-0.75) | 0.005 | 0.094 | 0.68 (0.43-1.08) | 0.101 | 0.862 | 0.73 (0.54-0.98) | 0.037 | 0.468 |
| Family *Ruminococcaceae* | | | | | | | | | | | |
| Species *Faecalibacterium prausnitzii J* | 0.093 | 90.7 | 0.37 (0.19-0.73) | 0.004 | 0.094 | 1.11 (0.73-1.69) | 0.626 | 0.972 | 0.82 (0.62-1.10) | 0.186 | 0.602 |
| Species *MGYG-HGUT-01157* ↓ | 0.005 | 70.4 | 0.31 (0.14-0.67) | 0.003 | 0.094 | 1.02 (0.66-1.58) | 0.914 | 0.993 | 0.73 (0.54-0.99) | 0.044 | 0.468 |
| Species *MGYG-HGUT-01627* ↓ | 0.015 | 74.4 | 0.34 (0.17-0.69) | 0.003 | 0.094 | 1.33 (0.85-2.08) | 0.217 | 0.938 | 0.83 (0.61-1.11) | 0.210 | 0.615 |
| Species *MGYG-HGUT-00605* | 0.002 | 60.1 | 0.26 (0.11-0.64) | 0.003 | 0.094 | 0.91 (0.59-1.41) | 0.679 | 0.981 | 0.79 (0.58-1.06) | 0.115 | 0.523 |
| Species *Ruminococcus D bicirculans* ↓ | 0.001 | 51.2 | 0.36 (0.18-0.73) | 0.005 | 0.094 | 0.87 (0.56-1.35) | 0.536 | 0.947 | 0.71 (0.53-0.96) | 0.027 | 0.468 |
| Species *MGYG-HGUT-02708* | 0.007 | 86.4 | 0.39 (0.20-0.75) | 0.005 | 0.094 | 0.96 (0.61-1.50) | 0.847 | 0.982 | 0.66 (0.49-0.88) | 0.005 | 0.406 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Order *Tissierellales* | | | | | | | | | | | |
| Family *Sporanaerobacteraceae* ↑ | 0.009 | 85.0 | 3.28 (1.49-7.23) | 0.003 | 0.057 | 0.87 (0.57-1.33) | 0.524 | 0.788 | 1.18 (0.88-1.57) | 0.275 | 0.625 |
| Species | | | | | | | | | | | |
| *Sporanaerobacter_acetigenes* | 0.009 | 84.1 | 2.94 (1.39-6.18) | 0.005 | 0.094 | 0.89 (0.59-1.35) | 0.594 | 0.972 | 1.23 (0.92-1.66) | 0.168 | 0.582 |
| Phylum *Firmicutes C* | | | | | | | | | | | |
| Species *MGYG-HGUT-00794* ↑ | 0.001 | 51.5 | 2.43 (1.31-4.51) | 0.005 | 0.094 | 1.11 (0.73-1.69) | 0.620 | 0.972 | 1.33 (0.99-1.80) | 0.058 | 0.496 |
| Phylum *Fusobacteriota* | | | | | | | | | | | |
| Family *Fusobacteriaceae* | 0.054 | 99.3 | 2.67 (1.35-5.25) | 0.005 | 0.061 | 0.95 (0.61-1.47) | 0.812 | 0.921 | 1.24 (0.92-1.66) | 0.153 | 0.421 |
| Species *MGYG-HGUT-03919* | 0.001 | 57.1 | 2.52 (1.35-4.72) | 0.004 | 0.094 | 1.16 (0.77-1.76) | 0.467 | 0.942 | 1.58 (1.18-2.12) | 0.002 | 0.251 |
| Phylum *Synergistota* | | | | | | | | | | | |
| Family *Synergistaceae*┼ | 0.006 | 20.6 ¥ | 0.58 (0.29-1.14) | 0.113 | 0.300 | 0.52 (0.30-0.91) | 0.021 | 0.332 | 0.59 (0.42-0.83) | 0.002 | 0.075 |
| Phylum *Verrucomicrobiota* | | | | | | | | | | | |
| Order *Victivallales* | | | | | | | | | | | |
| Family *Victivallaceae* | 0.005 | 78.7 | 0.53 (0.35-0.81) | 0.003 | 0.061 | 0.83 (0.63-1.11) | 0.207 | 0.668 | 0.77 (0.63-0.94) | 0.009 | 0.215 |
| Order *Opitutales* | | | | | | | | | | | |
| Family *CAG-312* | 0.004 | 16.6 | 0.33 (0.16-0.68) | 0.003 | 0.081 | 1.28 (0.86-1.91) | 0.224 | 0.717 | 0.85 (0.63-1.13) | 0.256 | 0.512 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥ Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}<0.1$ is considered statistically significant.

↓ : significantly decreased association with stool collection time ($P_{FDR}<0.1$)

↑: significantly increased association with stool collection time ($P_{FDR}<0.1$)

ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

Our study found no significant association between GI microbial taxa and grade ≥3 combined GI toxicity after FDR correction ($P_{FDR} < 0.1$). **Tables 46-48** show microbial taxa with a significant association ($P_{FDR} < 0.1$) with grade ≥3 nausea/vomiting. Most of the significant associations between grade ≥3 nausea/vomiting and GI microbial taxa belong to one family, including 47 genera, and 197 species among patients whose stool samples were collected after breast cancer surgery. Approximately 83% of these 245 taxa, including 41 genera, and 157 species, all belonging to phylum *Firmicutes A*, were significantly associated with an increased risk of grade ≥3 nausea/vomiting. **Tables 47 and 48** showed that the respective top 20 common taxa and rare taxa at the species levels had a significant association with grade ≥3 nausea/vomiting among patients with stool samples collected after breast cancer surgery. Among common taxa, species *CAG-83 sp000435975* showed the strongest association with OR and 95% CI of 6.38 (2.27-19.91), $P$=4.32x$10^{-4}$ and $P_{FDR}$=0.078. In addition, a high abundance of three species (belonging to phylum *Firmicutes A*) was associated with an increased risk of grade ≥3 nausea/vomiting, with OR and 95% CI of 3.42 (1.60-7.31) for *Coprococcus eutactus*, 6.64 (1.99-22.12) for *Ruminiclostridium E siraeum*, *and* 6.77 (2.06-22.18) for *Ruminiclostridium C sp000435295* (all $P$=0.002), respectively. Meanwhile, three species, *Odoribacter laneus* and *MGYG-HGUT-00222* (belonging to phylum *Bacteroidota)* and *Massiliomicrobiota timonensis* (belonging to phylum *Firmicutes),* were significantly associated with a reduced risk of grade ≥3 nausea/vomiting, with OR and 95% CI of 0.14 (0.04-0.48), 0.21 (0.08-0.56), and 0.19 (0.07-0.54); all $P$=0.002, respectively. Among rare taxa, the abundances of three *MGYG-HGUT* species (including *02250, 02661*, and *03293*) showed the strongest associations with OR and 95% CI of 6.68 (2.44-18.30) for *MGYG-HGUT-02250* ($P$=2.18x$10^{-4}$), 4.29 (1.97-9.35) for *MGYG-HGUT-02661* ($P$=2.45x$10^{-4}$), and 7.15 (2.37-21.59) for *MGYG-HGUT-03293* ($P$=4.84x$10^{-4}$), respectively (all $P_{FDR}$<0.1). In addition, species *Anaeromassilibacillus sp001305115* was positively associated with the risk of grade ≥3 nausea/vomiting, with OR of 11.34 (95% CI: 2.65-48.58); $P$=0.001 and $P_{FDR}$=0.064.

Last but not least, in the analysis for patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy, we found the genus *Ezakiella* was inversely associated with the risk of grade ≥3 nausea/vomiting, with an OR of 0.21 (95% CI: 0.07-0.63), $P$=0.005 and $P_{FDR}$=0.087.

**Table 46**: Association of grade ≥3 nausea/vomiting with gut microbiome taxa by stool collection time at order to genus levels

| | | | **Stool collection time** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)[1] | | | Collection after breast cancer surgery (N= 194)[2] | | | **Overall** (N=301)[3] | | |
| *Taxonomy* | RA, median (%) | Pre (%) | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] |
| Phylum *Bacteroidota* | | | | | | | | | | | |
| Genus *MGYG-HGUT-03221* ↑ | 0.006 | 84.4 | 0.55 (0.21-1.39) | 0.204 | 0.680 | 0.29 (0.12-0.70) | 0.006 | 0.087 | 0.54 (0.34-0.86) | 0.010 | 0.188 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Class *Clostridia* | | | | | | | | | | | |
| Order *4C28d-15* | | | | | | | | | | | |
| Genus *QALS01*ⱡ | 0.007¥ | 29.6 | 1.03 (0.49-2.17) | 0.935 | 0.975 | 3.82 (1.74-8.40) | 0.001 | 0.071 | 1.69 (1.16-2.46) | 0.006 | 0.126 |
| Genus *UBA7597*ⱡ | 0.025¥ | 21.3 | 0.95 (0.4-2.24) | 0.898 | 0.961 | 3.11 (1.42-6.77) | 0.004 | 0.081 | 1.49 (1.03-2.15) | 0.034 | 0.286 |
| Order *Christensenellales* | | | | | | | | | | | |
| Family *CAG-138* | 0.002 | 60.8 | 1.12 (0.51-2.45) | 0.782 | 0.939 | 6.28 (2.12-18.59) | 0.001 | 0.060 | 1.68 (1.13-2.48) | 0.010 | 0.200 |
| Genus *MGYG-HGUT-00718*ⱡ | 0.003 | 27.2 | 0.78 (0.31-1.98) | 0.605 | 0.824 | 7.70 (2.21-26.81) | 0.001 | 0.071 | 1.76 (1.17-2.65) | 0.006 | 0.126 |
| Genus *PeH17*ⱡ | 0.037¥ | 30.2 | 1.02 (0.42-2.49) | 0.963 | 0.975 | 2.87 (1.40-5.88) | 0.004 | 0.081 | 1.57 (1.07-2.31) | 0.021 | 0.237 |
| Genus *Phil1*ⱡ | 0.005¥ | 28.6 | 1.21 (0.55-2.65) | 0.635 | 0.837 | 5.06 (1.96-13.07) | 0.001 | 0.071 | 1.70 (1.17-2.49) | 0.006 | 0.126 |
| Family *CAG-74* | | | | | | | | | | | |
| Genus *MGYG-HGUT-01103* | 0.002 | 59.5 | 1.32 (0.58-3.02) | 0.509 | 0.786 | 8.94 (2.62-30.56) | $4.75 \times 10^{-4}$ | 0.087 | 1.89 (1.21-2.97) | 0.005 | 0.131 |
| Genus *MGYG-HGUT-01823*ⱡ | 0.005¥ | 15.9 | 1.38 (0.60-3.19) | 0.447 | 0.751 | 2.50 (1.31-4.80) | 0.006 | 0.085 | 1.44 (1.01-2.04) | 0.041 | 0.297 |
| Genus *MGYG-HGUT-02098*ⱡ | 0.004¥ | 36.9 | 2.79 (1.13-6.87) | 0.026 | 0.445 | 3.34 (1.49-7.49) | 0.003 | 0.081 | 1.78 (1.19-2.65) | 0.005 | 0.126 |
| Genus *MGYG-HGUT-03224*ⱡ | 0.013¥ | 48.2 | 2.25 (0.89-5.67) | 0.087 | 0.535 | 3.85 (1.55-9.54) | 0.004 | 0.081 | 1.63 (1.07-2.49) | 0.024 | 0.245 |
| Genus *MGYG-HGUT-03417*ⱡ | 0.001¥ | 52.8 | 0.58 (0.25-1.34) | 0.202 | 0.680 | 3.69 (1.41-9.67) | 0.008 | 0.092 | 1.27 (0.82-1.96) | 0.276 | 0.503 |
| Genus *MGYG-HGUT-03875*ⱡ | 0.003¥ | 40.2 | 1.93 (0.89-4.18) | 0.094 | 0.535 | 5.00 (1.84-13.62) | 0.002 | 0.071 | 1.72 (1.13-2.60) | 0.011 | 0.161 |
| Genus *MGYG-HGUT-04088*ⱡ | 0.004¥ | 35.2 | 1.04 (0.44-2.44) | 0.935 | 0.975 | 4.32 (1.55-12.03) | 0.005 | 0.081 | 1.22 (0.82-1.80) | 0.328 | 0.670 |
| Genus *UBA11524* ⱡ↓ | 0.024¥ | 36.2 | 1.96 (0.86-4.45) | 0.108 | 0.557 | 2.83 (1.31-6.11) | 0.008 | 0.098 | 1.61 (1.08-2.38) | 0.019 | 0.219 |
| Genus *MGYG-HGUT-02681*ⱡ | 0.003¥ | 36.2 | 0.73 (0.33-1.64) | 0.444 | 0.751 | 2.89 (1.33-6.31) | 0.008 | 0.097 | 1.21 (0.79-1.85) | 0.375 | 0.683 |
| Family *UBA1750*ⱡ | | | | | | | | | | | |
| Genus *UBA7102*ⱡ | 0.004¥ | 42.2 | 0.98 (0.45-2.13) | 0.962 | 0.975 | 8.69 (2.31-32.70) | 0.001 | 0.071 | 1.74 (1.14-2.66) | 0.010 | 0.150 |
| Family *Lachnospiraceae* | | | | | | | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Genus *CAG-791*† | 0.002¥ | 22.6 | 0.68 (0.25-1.85) | 0.450 | 0.751 | 2.66 (1.31-5.37) | 0.007 | 0.090 | 1.16 (0.79-1.71) | 0.457 | 0.746 |
| Genus *Coprococcus* ↓ | 0.017 | 90.0 | 1.44 (0.63-3.31) | 0.385 | 0.743 | 5.07 (1.97-13.07) | 0.001 | 0.087 | 2.01 (1.28-3.14) | 0.002 | 0.110 |
| Genus *Lachnoclostridium A* ↑ | 0.002 | 76.7 | 0.48 (0.20-1.14) | 0.096 | 0.639 | 0.27 (0.11-0.68) | 0.006 | 0.087 | 0.56 (0.35-0.90) | 0.017 | 0.188 |
| Genus *MGYG-HGUT-00202* ↑ | 0.002 | 65.4 | 1.74 (0.73-4.16) | 0.213 | 0.680 | 3.17 (1.40-7.19) | 0.006 | 0.087 | 1.67 (1.10-2.54) | 0.017 | 0.188 |
| Genus *MGYG-HGUT-03366*† | 0.002¥ | 15.6 | 2.39 (1.00-5.74) | 0.051 | 0.506 | 3.12 (1.45-6.71) | 0.004 | 0.081 | 1.84 (1.25-2.70) | 0.002 | 0.126 |
| Genus *MGYG-HGUT-04169* | 0.009 | 91.7 | 0.96 (0.44-2.12) | 0.925 | 0.972 | 0.24 (0.09-0.64) | 0.004 | 0.087 | 0.45 (0.27-0.77) | 0.003 | 0.110 |
| Order *Monoglobales* | | | | | | | | | | | |
| Genus *MGYG-HGUT-00495*† | 0.002¥ | 18.3 | 1.33 (0.61-2.91) | 0.471 | 0.759 | 6.70 (1.77-25.31) | 0.005 | 0.081 | 1.56 (1.09-2.22) | 0.014 | 0.180 |
| Order *Oscillospirales* | | | | | | | | | | | |
| Family *Acutalibacteraceae* | | | | | | | | | | | |
| Genus *CAG-177* ↓ | 0.002 | 58.8 | 1.11 (0.51-2.40) | 0.793 | 0.924 | 2.65 (1.30-5.38) | 0.007 | 0.092 | 1.49 (1.01-2.20) | 0.044 | 0.255 |
| Genus *CAG-180* | 0.005 | 62.5 | 1.06 (0.48-2.33) | 0.879 | 0.955 | 3.39 (1.52-7.58) | 0.003 | 0.087 | 1.35 (0.90-2.01) | 0.145 | 0.369 |
| Genus *Marseille-P4683* | 0.001 | 55.1 | 0.96 (0.43-2.12) | 0.912 | 0.972 | 3.36 (1.41-8.02) | 0.006 | 0.087 | 1.56 (1.04-2.36) | 0.034 | 0.236 |
| Genus *MGYG-HGUT-04279*† | 0.004¥ | 21.3 | 1.76 (0.77-4.01) | 0.183 | 0.591 | 3.48 (1.63-7.44) | 0.001 | 0.071 | 1.63 (1.14-2.34) | 0.008 | 0.133 |
| Genus *RUG806*† | 0.003 | 39.9 | 1.72 (0.80-3.72) | 0.168 | 0.583 | 3.11 (1.40-6.94) | 0.006 | 0.085 | 1.82 (1.18-2.82) | 0.007 | 0.126 |
| Genus *UBA737* ↓ | 0.002 | 52.8 | 0.68 (0.31-1.52) | 0.351 | 0.736 | 2.90 (1.30-6.46) | 0.009 | 0.099 | 1.13 (0.75-1.71) | 0.552 | 0.719 |
| Family *CAG-272* | | | | | | | | | | | |
| Genus *MGYG-HGUT-00516*† | 0.002¥ | 11.0 | 0.41 (0.11-1.50) | 0.178 | 0.587 | 2.11 (1.22-3.62) | 0.007 | 0.092 | 1.30 (0.92-1.83) | 0.135 | 0.494 |
| Genus *MGYG-HGUT-02872*† | 0.010¥ | 23.6 | 1.56 (0.70-3.49) | 0.280 | 0.645 | 3.47 (1.50-8.01) | 0.004 | 0.081 | 1.60 (1.10-2.33) | 0.013 | 0.180 |
| Family *Oscillospiraceae* | | | | | | | | | | | |
| Genus *CAG-110* ↓ | 0.037 | 92.0 | 1.72 (0.67-4.41) | 0.263 | 0.736 | 3.84 (1.46-10.08) | 0.006 | 0.087 | 1.96 (1.16-3.31) | 0.012 | 0.188 |
| Genus *CAG-170* ↓ | 0.006 | 65.8 | 1.83 (0.76-4.44) | 0.180 | 0.669 | 4.77 (1.84-12.40) | 0.001 | 0.087 | 1.97 (1.25-3.10) | 0.003 | 0.110 |
| Genus *CAG-83* ↓ | 0.075 | 83.4 | 1.82 (0.70-4.73) | 0.222 | 0.680 | 5.65 (1.74-18.37) | 0.004 | 0.087 | 1.88 (1.13-3.12) | 0.015 | 0.188 |
| Genus *EO* ↓ | 0.039 | 93.0 | 1.53 (0.65-3.60) | 0.335 | 0.736 | 3.02 (1.33-6.85) | 0.008 | 0.092 | 1.60 (1.02-2.50) | 0.041 | 0.255 |
| Genus *F23-B02* | 0.004 | 67.8 | 0.76 (0.35-1.65) | 0.486 | 0.782 | 3.44 (1.47-8.05) | 0.004 | 0.087 | 1.34 (0.89-2.03) | 0.164 | 0.398 |
| Genus *MGYG-HGUT-02213* | 0.001 | 51.2 | 1.24 (0.57-2.70) | 0.582 | 0.819 | 3.48 (1.42-8.48) | 0.006 | 0.087 | 1.61 (1.06-2.45) | 0.027 | 0.219 |
| Genus *Oscillibacter* | 0.148 | 99.0 | 1.70 (0.64-4.49) | 0.288 | 0.736 | 3.78 (1.39-10.28) | 0.009 | 0.099 | 2.03 (1.15-3.60) | 0.015 | 0.188 |
| Genus *Ruminiclostridium C* ↓ | 0.014 | 72.1 | 2.78 (0.97-7.96) | 0.058 | 0.587 | 5.22 (1.70-16.09) | 0.004 | 0.087 | 2.12 (1.24-3.60) | 0.006 | 0.131 |
| Genus *UBA1777* | 0.004 | 71.8 | 0.91 (0.41-2.02) | 0.822 | 0.932 | 5.44 (1.94-15.27) | 0.001 | 0.087 | 1.59 (1.03-2.47) | 0.038 | 0.241 |
| Genus *UBA738*† | 0.007¥ | 34.2 | 1.25 (0.55-2.80) | 0.594 | 0.822 | 4.20 (1.75-10.07) | 0.001 | 0.071 | 1.76 (1.19-2.60) | 0.004 | 0.126 |
| Family *Ruminococcaceae* | | | | | | | | | | | |
| Genus *Angelakisella* ↓ | 0.002 | 55.5 | 1.30 (0.60-2.79) | 0.503 | 0.782 | 2.92 (1.36-6.28) | 0.006 | 0.087 | 1.53 (1.04-2.25) | 0.032 | 0.236 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Genus *Ruminiclostridium E* | 0.003 | 69.4 | 2.77 (1.19-6.43) | 0.018 | 0.473 | 6.53 (1.96-21.82) | 0.002 | 0.087 | 2.24 (1.50-3.35) | $8.03 \times 10^{-5}$ | 0.022 |
| Genus *Ruminococcus C* ↓ | 0.013 | 80.1 | 2.04 (0.80-5.16) | 0.133 | 0.654 | 3.11 (1.34-7.21) | 0.008 | 0.092 | 1.54 (0.99-2.40) | 0.054 | 0.265 |
| Family *UBA644 A* | | | | | | | | | | | |
| Genus *MGYG-HGUT-01818*┼ | 0.004¥ | 21.3 | 2.02 (0.85-4.79) | 0.110 | 0.557 | 3.58 (1.61-7.94) | 0.002 | 0.071 | 1.92 (1.33-2.78) | 0.001 | 0.066 |
| Order *Tissierellales* ↑ | | | | | | | | | | | |
| Genus *Ezakiella* ↑ | 0.002 | 60.5 | 0.42 (0.17-1.04) | 0.060 | 0.590 | 0.21 (0.07-0.63) | 0.005 | 0.087 | 0.42 (0.25-0.70) | 0.001 | 0.110 |
| Order *UBA1212* | | | | | | | | | | | |
| Genus *MGYG-HGUT-00373*┼ | 0.002¥ | 22.3 | 0.36 (0.11-1.12) | 0.077 | 0.512 | 4.22 (1.70-10.5) | 0.002 | 0.071 | 1.22 (0.81-1.84) | 0.332 | 0.670 |
| Phylum *Firmicutes* | | | | | | | | | | | |
| Order *Erysipelotrichales* | | | | | | | | | | | |
| Genus *Massiliomicrobiota* ↑ | 0.009 | 90.7 | 0.90 (0.39-2.10) | 0.808 | 0.928 | 0.18 (0.06-0.52) | 0.002 | 0.087 | 0.49 (0.31-0.79) | 0.004 | 0.110 |
| Genus *CAG-884* | 0.001 | 53.2 | 1.29 (0.55-3.01) | 0.556 | 0.815 | 0.16 (0.05-0.55) | 0.003 | 0.087 | 0.57 (0.35-0.94) | 0.028 | 0.219 |
| Phylum *Synergistota* | | | | | | | | | | | |
| Genus *Cloacibacillus* ┼↓ | 0.004¥ | 16.6 | 1.85 (0.76-4.53) | 0.176 | 0.587 | 1.96 (1.15-3.36) | 0.014 | 0.122 | 1.86 (1.29-2.69) | 0.001 | 0.088 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[4] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$<0.1 is considered statistically significant.

↓ : significantly decreased association with stool collection time ($P_{FDR}$<0.1)

↑: significantly increased association with stool collection time ($P_{FDR}$<0.1)

ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 47:** Association of grade ≥3 nausea/vomiting with common microbiome taxa by stool collection time at species levels (Top 20)

| | | | Stool collection time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA, median | Pre | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)[1] | | | Collection after breast cancer surgery (N= 194)[2] | | | **Overall** (N=301)[3] | | |
| *Taxonomy* | (%) | (%) | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] | *OR (95%CI)*[1] | *P* | $P_{FDR}$[4] |
| Phylum *Bacteroidota* | | | | | | | | | | | |
| Species *Odoribacter laneus* ↑ | 0.002 | 60.8 | 0.10 (0.02-0.56) | 0.010 | 0.632 | 0.14 (0.04-0.48) | 0.002 | 0.078 | 0.36 (0.20-0.66) | 0.001 | 0.099 |
| Species *MGYG-HGUT-00222* ↑ | 0.002 | 68.4 | 0.29 (0.08-1.02) | 0.053 | 0.632 | 0.21 (0.08-0.56) | 0.002 | 0.078 | 0.39 (0.23-0.68) | 0.001 | 0.099 |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Species *MGYG-HGUT-01103* | 0.002 | 59.5 | 1.29 (0.57-2.94) | 0.538 | 0.855 | 9.58 (2.64-34.75) | 0.001 | 0.078 | 1.86 (1.20-2.90) | 0.006 | 0.123 |
| Species *MGYG-HGUT-02738* | 0.003 | 84.7 | 0.44 (0.17-1.13) | 0.088 | 0.632 | 0.27 (0.11-0.63) | 0.003 | 0.078 | 0.41 (0.25-0.68) | 0.001 | 0.099 |
| Species *Coprococcus eutactus* ↓ | 0.003 | 72.8 | 1.64 (0.76-3.54) | 0.212 | 0.752 | 3.42 (1.60-7.31) | 0.002 | 0.078 | 1.91 (1.29-2.83) | 0.001 | 0.099 |
| Species *CAG-110_sp000434635* | 0.001 | 50.8 | 1.18 (0.56-2.48) | 0.665 | 0.889 | 4.81 (1.92-12.07) | 0.001 | 0.078 | 1.88 (1.25-2.83) | 0.002 | 0.099 |
| Species *CAG-110_sp003525905* | 0.002 | 54.5 | 1.30 (0.58-2.94) | 0.526 | 0.846 | 3.29 (1.59-6.83) | 0.001 | 0.078 | 1.63 (1.08-2.47) | 0.020 | 0.187 |
| Species *MGYG-HGUT-00741* | 0.001 | 50.2 | 1.08 (0.50-2.31) | 0.850 | 0.946 | 3.41 (1.57-7.40) | 0.002 | 0.078 | 1.61 (1.08-2.42) | 0.021 | 0.187 |
| Species *CAG-83 sp000435555* | 0.003 | 60.8 | 1.32 (0.60-2.90) | 0.495 | 0.832 | 4.43 (1.80-10.91) | 0.001 | 0.078 | 1.74 (1.17-2.60) | 0.007 | 0.131 |
| Species *CAG-83 sp000435975* ↓ | 0.004 | 64.5 | 1.17 (0.51-2.67) | 0.705 | 0.907 | 6.38 (2.27-17.91) | $4.32 \times 10^{-4}$ | 0.078 | 1.97 (1.27-3.05) | 0.002 | 0.099 |
| Species *CAG-83 sp001916855* ↓ | 0.003 | 57.1 | 1.41 (0.69-2.86) | 0.342 | 0.782 | 4.35 (1.80-10.53) | 0.001 | 0.078 | 1.77 (1.17-2.67) | 0.006 | 0.131 |
| Species *CAG-83 sp003539495* | 0.001 | 52.2 | 1.48 (0.70-3.13) | 0.301 | 0.773 | 3.52 (1.60-7.74) | 0.002 | 0.078 | 1.66 (1.12-2.46) | 0.011 | 0.166 |
| Species *MGYG-HGUT-00713* ↓ | 0.001 | 50.2 | 1.57 (0.66-3.71) | 0.305 | 0.773 | 4.26 (1.74-10.44) | 0.002 | 0.078 | 1.91 (1.26-2.91) | 0.002 | 0.099 |
| Species *EO sp900317525* | 0.002 | 57.5 | 1.44 (0.67-3.10) | 0.349 | 0.785 | 3.25 (1.56-6.77) | 0.002 | 0.078 | 1.84 (1.24-2.74) | 0.003 | 0.099 |
| Species *MGYG-HGUT-00837* ↓ | 0.003 | 66.1 | 1.54 (0.67-3.53) | 0.310 | 0.773 | 7.99 (2.32-27.45) | 0.001 | 0.078 | 2.38 (1.44-3.92) | 0.001 | 0.099 |
| Species *Ruminiclostridium C sp000435295* | 0.002 | 65.8 | 1.74 (0.77-3.89) | 0.181 | 0.716 | 6.77 (2.06-22.18) | 0.002 | 0.078 | 2.23 (1.39-3.58) | 0.001 | 0.099 |
| Species *MGYG-HGUT-03668* | 0.001 | 51.2 | 1.53 (0.70-3.36) | 0.290 | 0.771 | 4.43 (1.78-11.01) | 0.001 | 0.078 | 1.87 (1.24-2.83) | 0.003 | 0.099 |
| Species *Ruminiclostridium E siraeum* | 0.002 | 63.5 | 3.12 (1.31-7.44) | 0.010 | 0.632 | 6.64 (1.99-22.12) | 0.002 | 0.078 | 2.36 (1.59-3.51) | $2.21 \times 10^{-5}$ | 0.016 |
| Species *MGYG-HGUT-04016* | 0.002 | 54.5 | 0.93 (0.41-2.16) | 0.874 | 0.952 | 4.57 (1.81-11.50) | 0.001 | 0.078 | 1.54 (1.00-2.37) | 0.051 | 0.240 |
| Phylum *Firmicutes* | | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Species *Massiliomicrobiota timonensis* ↑ | 0.008 | 90.0 | 0.81 (0.34-1.89) | 0.617 | 0.874 | 0.19 (0.07-0.54) | 0.002 | 0.078 | 0.50 (0.31-0.79) | 0.003 | 0.099 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level. $P_{FDR}$ <0.1 is considered statistically significant. Tests were conducted for common taxa, including 12 phyla, 14 classes, 35 orders, 70 families, 276 genera, and 757 species.
↓ : significantly decreased association with stool collection time ($P_{FDR}$ <0.1)
↑: significantly increased association with stool collection time ($P_{FDR}$ <0.1)
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 48**: Association of grade ≥3 nausea/vomiting with rare microbiome taxa by stool collection time at species levels (Top 20)

| | | | Stool collection time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | RA, media[¥] | Pre | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=162)[1] | | | Collection after breast cancer surgery (N= 194) [2] | | | **Overall** (N=301)[3] | | |
| *Taxonomy* | (%) | (%) | *OR (95%CI)[1]* | *P* | $P_{FDR}$[4] | *OR (95%CI)[1]* | *P* | $P_{FDR}$[4] | *OR (95%CI)[1]* | *P* | $P_{FDR}$[4] |
| Phylum *Firmicutes A* | | | | | | | | | | | |
| Species *MGYG-HGUT-00718* | 0.003 | 27.2 | 0.68 (0.27-1.73) | 0.421 | 0.776 | 6.74 (2.26-20.08) | 0.001 | 0.064 | 1.75 (1.17-2.61) | 0.006 | 0.192 |
| Species *Phil1 sp001940855* | 0.003 | 28.2 | 1.22 (0.55-2.68) | 0.624 | 0.873 | 4.82 (1.91-12.19) | 0.001 | 0.064 | 1.67 (1.15-2.42) | 0.007 | 0.197 |
| Species *MGYG-HGUT-03875* | 0.003 | 40.2 | 1.68 (0.79-3.55) | 0.176 | 0.655 | 5.44 (1.94-15.30) | 0.001 | 0.064 | 1.73 (1.14-2.61) | 0.010 | 0.221 |
| Species *MGYG-HGUT-02099* | 0.004 | 40.2 | 0.96 (0.45-2.02) | 0.906 | 0.967 | 3.73 (1.69-8.22) | 0.001 | 0.064 | 1.62 (1.07-2.44) | 0.021 | 0.287 |
| Species *MGYG-HGUT-03366* | 0.002 | 15.6 | 2.44 (1.02-5.84) | 0.045 | 0.596 | 3.85 (1.70-8.73) | 0.001 | 0.064 | 1.91 (1.30-2.80) | 0.001 | 0.124 |
| Species *Anaeromassilibacillus sp001305115* | 0.002 | 25.2 | 0.37 (0.12-1.15) | 0.087 | 0.623 | 11.34 (2.65-48.58) | 0.001 | 0.064 | 1.49 (1.02-2.17) | 0.041 | 0.339 |
| Species *MGYG-HGUT-04279* | 0.004 | 21.3 | 1.74 (0.76-4.01) | 0.191 | 0.656 | 3.70 (1.70-8.05) | 0.001 | 0.064 | 1.67 (1.17-2.39) | 0.004 | 0.192 |
| Species *MGYG-HGUT-00715* | 0.003 | 32.2 | 1.04 (0.50-2.18) | 0.917 | 0.969 | 4.67 (1.85-11.80) | 0.001 | 0.064 | 1.61 (1.07-2.42) | 0.023 | 0.287 |
| Species *MGYG-HGUT-02045* | 0.004 | 24.3 | 0.71 (0.30-1.65) | 0.423 | 0.776 | 4.13 (1.85-9.23) | 0.001 | 0.064 | 1.46 (0.99-2.16) | 0.059 | 0.395 |
| Species *MGYG-HGUT-02250* | 0.003 | 11.0 | 0.11 (0.01-1.19) | 0.069 | 0.596 | 6.68 (2.44-18.30) | $2.18 \times 10^{-4}$ | 0.064 | 1.34 (0.93-1.93) | 0.115 | 0.451 |
| Species *MGYG-HGUT-03293* ↓ | 0.008 | 35.5 | 1.24 (0.56-2.73) | 0.602 | 0.864 | 7.15 (2.37-21.59) | $4.84 \times 10^{-4}$ | 0.064 | 1.77 (1.17-2.67) | 0.006 | 0.192 |
| Species *MGYG-HGUT-01632* ↓ | 0.006 | 39.9 | 3.68 (1.30-10.42) | 0.014 | 0.596 | 7.37 (2.29-23.72) | 0.001 | 0.064 | 2.53 (1.59-4.03) | $9.44 \times 10^{-5}$ | 0.058 |
| Species *CAG-83 sp003487665* | 0.008 | 41.9 | 1.28 (0.57-2.84) | 0.549 | 0.830 | 4.33 (1.89-9.94) | 0.001 | 0.064 | 1.73 (1.17-2.56) | 0.006 | 0.192 |
| Species *MGYG-HGUT-04028* | 0.004 | 30.9 | 0.59 (0.24-1.47) | 0.258 | 0.671 | 4.49 (1.87-10.79) | 0.001 | 0.064 | 1.29 (0.87-1.92) | 0.200 | 0.554 |
| Species *F23-B02 sp003533405* | 0.003 | 23.9 | 0.90 (0.42-1.93) | 0.788 | 0.938 | 5.38 (1.99-14.58) | 0.001 | 0.064 | 1.53 (1.06-2.20) | 0.023 | 0.287 |
| Species *MGYG-HGUT-02716* | 0.003 | 20.6 | 1.29 (0.53-3.16) | 0.571 | 0.852 | 3.50 (1.65-7.43) | 0.001 | 0.064 | 1.75 (1.21-2.55) | 0.003 | 0.167 |
| Species *MGYG-HGUT-00474* ↓ | 0.004 | 35.2 | 1.33 (0.66-2.68) | 0.428 | 0.776 | 5.60 (1.97-15.92) | 0.001 | 0.064 | 1.76 (1.17-2.64) | 0.006 | 0.192 |
| Species *UBA738 sp003522945* | 0.010 | 27.9 | 1.35 (0.63-2.91) | 0.444 | 0.783 | 3.84 (1.70-8.71) | 0.001 | 0.064 | 1.79 (1.23-2.61) | 0.002 | 0.167 |
| Species *MGYG-HGUT-02661* | 0.002 | 14.0 | 0.77 (0.35-1.69) | 0.513 | 0.816 | 4.29 (1.97-9.35) | $2.45 \times 10^{-4}$ | 0.064 | 1.57 (1.12-2.21) | 0.009 | 0.221 |
| Species *MGYG-HGUT-03494* | 0.003 | 11.0 | 2.81 (1.25-6.32) | 0.012 | 0.596 | 3.31 (1.60-6.83) | 0.001 | 0.064 | 1.96 (1.38-2.78) | $1.82 \times 10^{-4}$ | 0.073 |

[¥] Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).

[3] Models were adjusted for stool collection time (after breast cancer surgery vs. Before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[4] FDR were calculated at each taxonomic level. $P_{FDR}$ <0.1 is considered statistically significant. Test were conducted for rare taxa, including 5 phyla, 9 classes, 17 orders, 67 families, 370 genera, and 1232 species.
↓ : significantly decreased association with stool collection time ($P_{FDR}$ <0.1)
↑: significantly increased association with stool collection time ($P_{FDR}$ <0.1)
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.
.

**Association of GI microbial metabolic pathways with chemotherapy-induced toxicity**

Among 301 stool samples, a total of 549 GI microbial metabolic pathways was identified and their relative abundance were estimated. After excluding GI microbial metabolic pathways that present in <10% of samples, a total of 419 microbial metabolic pathways were included to the evaluation of the associations between GI microbial metabolic pathways and chemotherapy-induced toxicities.

**Tables 49-50** present significant associations ($P < 0.05$) between GI microbial metabolic pathways and grade ≥3 combined hematological toxicity and neutropenia, all had $P_{FDR}$>0.1. Among a total of 419 pathways investigated, 34 metabolic pathways showed significant associations with the grade ≥3 combined hematological toxicity in the overall analysis. In analyses stratified by stool collection time, 28 of 53 significant associations with the risk of grade ≥3 combined hematological toxicity were confined to patients with stool collected before breast cancer surgery with a non-significant association abbey in the same direction observed in patients with stool samples after breast cancer surgery (**Table 49**). Among patients with stool samples collected before surgery and chemotherapy, seven pathways showed strong associations with a reduced risk of grade ≥3 combined hematological toxicity. Three of the seven pathways are related to pyrimidine deoxyribonucleotides de novo biosynthesis II and the superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis, and were significantly associated with grade ≥3 combined hematological toxicity with OR of 0.43 (95% CI: 0.26-0.73; $P$=1.48×10$^{-3}$) for PWY-7187, 0.46 (95% CI: 0.27-0.78; $P$=3.97×10$^{-3}$) for PWY-7211 and 0.47 (95% CI: 0.28-0.79; $P$=4.52×10$^{-3}$) for PWY0-166, respectively. Two pathways of peptidoglycan biosynthesis IV and II were associated with the risk of grade ≥3 combined hematological toxicity, with OR of 0.48 (95% CI: 0.30-0.77) for PWY-6471 ($P$=2.33×10$^{-3}$) and 0.69 (95% CI: 0.55-0.87) for PWY-5265 ($P$<0.01), respectively. In addition, methanogenesis from acetate (METH-ACETATE-PWY) and lactose and galactose degradation I (LACTOSECAT-PWY) showed significantly inverse associations with grade ≥3 combined hematological toxicity, with OR of 0.47 (95% CI:0.28-0.76; $P$=2.47×10$^{-3}$) and 0.46 (95% CI:0.28-0.76; $P$=2.29x10$^{-3}$), respectively. Moreover, three pathways which are related to the superpathway of polyamine biosynthesis I, the sucrose degradation IV, and phosphatidate metabolism showed significant associations with an increased risk of grade ≥3 combined hematological toxicity, with OR of 1.80 (95% CI: 1.10-2.94; $P$=0.02) for POLYAMSYN-PWY, 1.74 (95% CI: 1.07-2.83; $P$=0.02) for

PWY-5384 and 1.69 (95% CI: 1.03-2.75; $P$=0.04) for PWY-7039, respectively. The significant associations of these above-mentioned pathways also were found in the overall analyses, except for PWY-7211. The pathways PWY-6470 related to peptidoglycan biosynthesis V (β-lactam resistance) showed the strongest association with decreased risk of grade ≥3 combined hematological toxicity in the overall analysis, with OR of 0.66 (95% CI: 0.50-0.87; $P$=3.00×10$^{-3}$) (**Table 49**).

Furthermore, four metabolic pathways, including methylerythritol phosphate pathway II (PWY-7560), starch biosynthesis (PWY-622), isoprene biosynthesis I (PWY-6270), and creatinine degradation I (CRNFORCAT-PWY), were associated with increased risk of grade ≥3 combined hematological toxicity in the analysis for patients whose stool samples were collected after breast cancer surgery, which were consistently observed in the overall analysis. There were eight significant associations found only among patients with stool samples collected after surgery, but all the eight metabolic pathways had a non-significant association among patients with stool samples before breast cancer surgery. Five of them were significantly associated with an increased risk of grade ≥3 combined hematological toxicity. The five metabolic pathways, including NONMEVIPP-PWY, PWY-5751, PWY-6270, PWY66-391, and PWY-7288, had different association patterns compared with that among patients with stool collected before breast cancer surgery (**Table 49**).

The metabolic pathways, including PWY-7187, PWY0-166, PWY-6471, PWY-5265, METH-ACETATE-PWY, LACTOSECAT-PWY, PWY-6470, POLYAMSYN-PWY, and PWY-7039, showed similarly significant association patterns with the risk of grade ≥3 neutropenia in the overall analysis and the analysis for restricted to patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. In addition, the pathways PWY-6470 related to peptidoglycan biosynthesis V (β-lactam resistance) also showed a strong association with decreased risk of grade ≥3 neutropenia in the overall analysis (OR=0.61, 95% CI: 0.46-0.82; $P$ =9.82×10$^{-4}$), and the analysis for patients with stool collected before breast cancer surgery and neoadjuvant chemotherapy (OR=0.40, 95% CI: 0.22-0.72; $P$ =2.27×10$^{-3}$). In the analysis of the latter subgroup, the CMP-pseudaminate biosynthesis (PWY-6143) was associated with decreased risk of grade ≥3 neutropenia, with OR of 0.14 (95% CI: 0.04-0.49); and $P$ =2.07×10$^{-3}$ (**Table 50**).

**Table 49:** Association of grade ≥3 combined hematological toxicity with metabolic pathways by stool collection time

| | | Stool collecting time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| ALLANTOINDEG-PWY | Superpathway of allantoin degradation in yeast | 1.44 (0.91-2.26) | 0.12 | 1.23 (0.84-1.80) | 0.28 | 1.35 (1.04-1.77) | 0.02 |
| ANAGLYCOLYSIS-PWY | Glycolysis III (from glucose) | 0.56 (0.34-0.93) | 0.02 | 0.92 (0.64-1.33) | 0.67 | 0.76 (0.58-1.00) | 0.05 |
| ARG+POLYAMINE-SYN | Superpathway of arginine and polyamine biosynthesis | 1.76 (1.06-2.90) | 0.03 | 1.02 (0.72-1.46) | 0.90 | 1.27 (0.97-1.66) | 0.08 |
| ARGSYNBSUB-PWY | L-arginine biosynthesis II (acetyl cycle) | 0.63 (0.41-0.98) | 0.04 | 1.11 (0.76-1.60) | 0.59 | 0.88 (0.68-1.14) | 0.33 |
| ARGSYN-PWY | L-arginine biosynthesis I (via L-ornithine) | 0.61 (0.39-0.96) | 0.03 | 1.09 (0.76-1.58) | 0.64 | 0.86 (0.66-1.12) | 0.26 |
| ARO-PWY | Chorismate biosynthesis I | 0.66 (0.42-1.04) | 0.08 | 0.87 (0.60-1.26) | 0.46 | 0.76 (0.58-0.99) | 0.05 |
| CALVIN-PWY | Calvin-Benson-Bassham cycle | 0.56 (0.36-0.89) | 0.01 | 0.82 (0.57-1.18) | 0.30 | 0.72 (0.55-0.94) | 0.02 |
| CENTFERM-PWY | Pyruvate fermentation to butanoate | 0.58 (0.36-0.92) | 0.02 | 0.90 (0.61-1.32) | 0.58 | 0.75 (0.57-0.98) | 0.04 |
| COBALSYN-PWY | Adenosylcobalamin salvage from cobinamide I | 0.79 (0.52-1.21) | 0.29 | 0.70 (0.47-1.03) | 0.07 | 0.70 (0.53-0.93) | 0.01 |
| CRNFORCAT-PWY | Creatinine degradation I | 1.08 (0.68-1.71) | 0.74 | 1.60 (1.01-2.53) | 0.05 | 1.37 (1.04-1.80) | 0.02 |
| DAPLYSINESYN-PWY | L-lysine biosynthesis I | 1.43 (0.89-2.29) | 0.14 | 1.31 (0.91-1.90) | 0.15 | 1.34 (1.03-1.75) | 0.03 |
| DENOVOPURINE2-PWY | Superpathway of purine nucleotides de novo biosynthesis II | 0.57 (0.33-0.97) | 0.04 | 0.76 (0.52-1.09) | 0.13 | 0.70 (0.53-0.93) | 0.01 |
| GLUDEG-I-PWY | GABA shunt | 1.68 (1.06-2.65) | 0.03 | 0.88 (0.61-1.27) | 0.50 | 1.21 (0.93-1.58) | 0.15 |
| GLYCOGENSYNTH-PWY | Glycogen biosynthesis I (from ADP-D-Glucose) | 0.61 (0.38-0.96) | 0.03 | 1.18 (0.82-1.71) | 0.38 | 0.85 (0.65-1.11) | 0.23 |
| HCAMHPDEG-PWY | 3-phenylpropanoate and 3-(3-Hydroxyphenyl) propanoate degradation to 2-oxopent-4-enoate | 1.64 (1.03-2.62) | 0.04 | 0.93 (0.64-1.34) | 0.68 | 1.18 (0.91-1.54) | 0.21 |
| HEXITOLDEGSUPER-PWY | Superpathway of hexitol degradation (bacteria) | 1.64 (1.02-2.63) | 0.04 | 0.92 (0.65-1.31) | 0.64 | 1.17 (0.90-1.52) | 0.25 |
| LACTOSECAT-PWY | Lactose and galactose degradation I | 0.46 (0.28-0.76) | $2.29x10^{-3}$ | 0.98 (0.67-1.43) | 0.91 | 0.74 (0.56-0.99) | 0.04 |
| METH-ACETATE-PWY | Methanogenesis from acetate | 0.47 (0.28-0.76) | $2.47x10^{-3}$ | 0.97 (0.66-1.42) | 0.88 | 0.69 (0.52-0.92) | 0.01 |
| NONMEVIPP-PWY | Methylerythritol phosphate pathway I | 0.94 (0.62-1.44) | 0.79 | 1.58 (1.07-2.31) | 0.02 | 1.28 (0.99-1.66) | 0.06 |
| P108-PWY | Pyruvate fermentation to propanoate I | 0.64 (0.40-1.00) | 0.05 | 0.70 (0.47-1.04) | 0.08 | 0.74 (0.57-0.97) | 0.03 |
| P162-PWY | L-glutamate degradation V (via hydroxyglutarate) | 0.78 (0.50-1.24) | 0.30 | 0.77 (0.53-1.12) | 0.17 | 0.74 (0.56-0.97) | 0.03 |
| P461-PWY | Hexitol fermentation to lactate, formate, ethanol and acetate | 1.74 (1.08-2.80) | 0.02 | 1.04 (0.73-1.49) | 0.82 | 1.28 (0.98-1.67) | 0.07 |
| P4-PWY | Superpathway of L-lysine, L-Threonine and L-methionine biosynthesis I | 1.58 (0.98-2.55) | 0.06 | 1.19 (0.83-1.71) | 0.35 | 1.32 (1.01-1.72) | 0.04 |
| POLYAMSYN-PWY | Superpathway of polyamine biosynthesis I | 1.80 (1.10-2.94) | 0.02 | 1.06 (0.74-1.52) | 0.73 | 1.31 (1.00-1.72) | 0.05 |
| PROPFERM-PWY | L-alanine fermentation to propanoate and acetate | 0.96 (0.60-1.55) | 0.88 | 0.62 (0.39-0.99) | 0.05 | 0.74 (0.55-1.00) | 0.05 |

| | | Stool collecting time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167) [2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY0-1261 | Anhydromuropeptides recycling | 0.84 (0.55-1.26) | 0.39 | 0.65 (0.45-0.96) | 0.03 | 0.76 (0.59-0.98) | 0.04 |
| PWY0-1296 | Purine ribonucleosides degradation | 0.65 (0.42-1.00) | 0.05 | 0.88 (0.61-1.26) | 0.49 | 0.78 (0.60-1.02) | 0.07 |
| PWY0-1415 | Superpathway of heme biosynthesis from uroporphyrinogen-III | 1.59 (1.00-2.53) | 0.05 | 0.86 (0.60-1.24) | 0.41 | 1.13 (0.87-1.46) | 0.37 |
| PWY0-162 | Superpathway of pyrimidine ribonucleotides de novo biosynthesis | 0.62 (0.38-0.99) | 0.05 | 0.99 (0.70-1.42) | 0.97 | 0.88 (0.68-1.14) | 0.32 |
| PWY0-166 | Superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis (E. coli) | 0.47 (0.28-0.79) | $4.52 \times 10^{-3}$ | 0.91 (0.63-1.32) | 0.63 | 0.76 (0.58-1.00) | 0.05 |
| PWY-1269 | CMP-3-deoxy-D-manno-octulosonate biosynthesis I | 1.42 (0.90-2.25) | 0.13 | 1.30 (0.88-1.93) | 0.19 | 1.35 (1.03-1.77) | 0.03 |
| PWY-241 | C4 photosynthetic carbon assimilation cycle, NADP-ME type | 1.69 (1.04-2.74) | 0.03 | 0.94 (0.65-1.36) | 0.74 | 1.22 (0.93-1.59) | 0.15 |
| PWY3O-355 | Stearate biosynthesis III (fungi) | 1.72 (1.08-2.71) | 0.02 | 0.78 (0.52-1.16) | 0.22 | 1.20 (0.92-1.56) | 0.19 |
| PWY-4981 | L-proline biosynthesis II (from arginine) | 0.63 (0.39-1.00) | 0.05 | 0.75 (0.52-1.08) | 0.13 | 0.76 (0.58-0.99) | 0.04 |
| PWY-5022 | 4-aminobutanoate degradation V | 1.71 (1.08-2.71) | 0.02 | 0.85 (0.59-1.22) | 0.38 | 1.18 (0.91-1.54) | 0.21 |
| PWY-5088 | L-glutamate degradation VIII (to propanoate) | 1.88 (1.15-3.08) | 0.01 | 1.15 (0.80-1.66) | 0.45 | 1.28 (0.99-1.66) | 0.06 |
| PWY-5100 | Pyruvate fermentation to acetate and lactate II | 0.48 (0.29-0.79) | $3.83 \times 10^{-3}$ | 1.14 (0.78-1.65) | 0.50 | 0.78 (0.59-1.03) | 0.07 |
| PWY-5265 | Peptidoglycan biosynthesis II (staphylococci) | 0.39 (0.20-0.76) | 0.01 | 0.86 (0.60-1.24) | 0.43 | 0.69 (0.52-0.92) | 0.01 |
| PWY-5384 | Sucrose degradation IV (sucrose phosphorylase) | 1.74 (1.07-2.83) | 0.02 | 1.20 (0.84-1.73) | 0.32 | 1.34 (1.03-1.74) | 0.03 |
| PWY-5656 | Mannosylglycerate biosynthesis I | 1.62 (1.04-2.54) | 0.03 | 0.94 (0.66-1.34) | 0.72 | 1.21 (0.94-1.57) | 0.14 |
| PWY-5723 | Rubisco shunt | 1.68 (1.04-2.73) | 0.04 | 0.91 (0.63-1.31) | 0.62 | 1.20 (0.92-1.56) | 0.19 |
| PWY-5751 | Phenylethanol biosynthesis | 0.77 (0.43-1.36) | 0.36 | 1.55 (1.06-2.27) | 0.02 | 1.16 (0.91-1.50) | 0.23 |
| PWY-622 | Starch biosynthesis | 1.06 (0.69-1.62) | 0.79 | 1.67 (1.13-2.46) | 0.01 | 1.39 (1.07-1.82) | 0.01 |
| PWY-6270 | Isoprene biosynthesis I | 0.93 (0.61-1.42) | 0.74 | 1.62 (1.10-2.38) | 0.01 | 1.29 (1.00-1.67) | 0.05 |
| PWY-6285 | Superpathway of fatty acids biosynthesis (E. coli) | 1.60 (1.00-2.57) | 0.05 | 0.86 (0.60-1.23) | 0.41 | 1.18 (0.90-1.54) | 0.22 |
| PWY-6317 | Galactose degradation I (Leloir pathway) | 0.65 (0.41-1.01) | 0.06 | 0.79 (0.55-1.14) | 0.21 | 0.72 (0.55-0.95) | 0.02 |
| PWY-6435 | 4-hydroxybenzoate biosynthesis V | 0.90 (0.58-1.39) | 0.63 | 1.71 (1.14-2.56) | 0.01 | 1.17 (0.90-1.51) | 0.24 |
| PWY-6470 | Peptidoglycan biosynthesis V (β-lactam resistance) | 0.57 (0.36-0.90) | 0.02 | 0.73 (0.50-1.06) | 0.10 | 0.66 (0.50-0.87) | $3.00 \times 10^{-3}$ |
| PWY-6471 | Peptidoglycan biosynthesis IV (Enterococcus faecium) | 0.48 (0.30-0.77) | $2.33 \times 10^{-3}$ | 0.91 (0.63-1.32) | 0.63 | 0.73 (0.56-0.95) | 0.02 |
| PWY-6531 | Mannitol cycle | 1.71 (1.08-2.69) | 0.02 | 1.02 (0.71-1.47) | 0.91 | 1.25 (0.96-1.63) | 0.09 |

| | | Stool collecting time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167) [2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-6545 | Pyrimidine deoxyribonucleotides de novo biosynthesis III | 0.59 (0.37-0.94) | 0.03 | 1.07 (0.74-1.54) | 0.74 | 0.90 (0.69-1.18) | 0.45 |
| PWY-6590 | Superpathway of Clostridium acetobutylicum acidogenic fermentation | 0.57 (0.36-0.92) | 0.02 | 0.90 (0.61-1.32) | 0.58 | 0.75 (0.57-0.98) | 0.03 |
| PWY-6595 | Superpathway of guanosine nucleotides degradation (plants) | 0.48 (0.28-0.83) | 0.01 | 1.06 (0.74-1.51) | 0.76 | 0.87 (0.66-1.14) | 0.30 |
| PWY66-391 | Fatty acid β-oxidation VI (peroxisome) | 1.22 (0.78-1.90) | 0.38 | 0.60 (0.41-0.89) | 0.01 | 0.79 (0.60-1.03) | 0.08 |
| PWY66-399 | Gluconeogenesis III | 0.62 (0.39-0.97) | 0.04 | 0.71 (0.49-1.04) | 0.08 | 0.72 (0.55-0.93) | 0.01 |
| PWY66-422 | D-galactose degradation V (Leloir pathway) | 0.62 (0.40-0.98) | 0.04 | 1.00 (0.70-1.43) | 0.99 | 0.81 (0.62-1.06) | 0.12 |
| PWY-6690 | cinnamate and 3-Hydroxycinnamate degradation to 2-oxopent-4-enoate | 1.64 (1.03-2.62) | 0.04 | 0.93 (0.64-1.34) | 0.68 | 1.18 (0.91-1.54) | 0.21 |
| PWY-6803 | Phosphatidylcholine acyl editing | 1.61 (1.02-2.55) | 0.04 | 0.93 (0.65-1.34) | 0.71 | 1.21 (0.93-1.58) | 0.16 |
| PWY-7039 | Phosphatidate metabolism, as a signaling molecule | 1.69 (1.03-2.75) | 0.04 | 1.14 (0.80-1.64) | 0.46 | 1.34 (1.03-1.74) | 0.03 |
| PWY-7046 | 4-coumarate degradation (anaerobic) | 1.60 (1.01-2.55) | 0.05 | 1.10 (0.77-1.57) | 0.59 | 1.24 (0.95-1.62) | 0.11 |
| PWY-7117 | C4 photosynthetic carbon assimilation cycle, PEPCK type | 1.61 (1.01-2.59) | 0.05 | 1.06 (0.74-1.52) | 0.77 | 1.26 (0.97-1.64) | 0.09 |
| PWY-7187 | Pyrimidine deoxyribonucleotides de novo biosynthesis II | 0.43 (0.26-0.73) | $1.48 \times 10^{-3}$ | 0.91 (0.63-1.31) | 0.62 | 0.73 (0.56-0.95) | 0.02 |
| PWY-7196 | Superpathway of pyrimidine ribonucleosides salvage | 0.52 (0.33-0.81) | $3.96 \times 10^{-3}$ | 1.21 (0.84-1.75) | 0.31 | 0.85 (0.66-1.10) | 0.21 |
| PWY-7197 | Pyrimidine deoxyribonucleotide phosphorylation | 0.55 (0.33-0.90) | 0.02 | 0.83 (0.58-1.20) | 0.32 | 0.77 (0.58-1.00) | 0.05 |
| PWY-7208 | Superpathway of pyrimidine nucleobases salvage | 0.60 (0.36-0.99) | 0.05 | 0.74 (0.52-1.06) | 0.10 | 0.71 (0.54-0.93) | 0.01 |
| PWY-7211 | Superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis | 0.46 (0.27-0.78) | $3.97 \times 10^{-3}$ | 1.19 (0.83-1.71) | 0.34 | 0.89 (0.69-1.15) | 0.37 |
| PWY-724 | Superpathway of L-lysine, L-threonine and L-methionine biosynthesis II | 0.50 (0.30-0.83) | 0.01 | 1.02 (0.71-1.47) | 0.90 | 0.78 (0.60-1.01) | 0.06 |
| PWY-7282 | 4-amino-2-methyl-5-Phosphomethylpyrimidine biosynthesis (yeast) | 0.73 (0.46-1.17) | 0.20 | 0.68 (0.46-0.99) | 0.04 | 0.73 (0.56-0.96) | 0.02 |
| PWY-7288 | Fatty acid β-oxidation (peroxisome, yeast) | 1.21 (0.77-1.88) | 0.41 | 0.58 (0.39-0.86) | 0.01 | 0.77 (0.59-1.00) | 0.05 |
| PWY-7383 | Anaerobic energy metabolism (invertebrates, cytosol) | 0.66 (0.42-1.03) | 0.07 | 0.74 (0.51-1.08) | 0.12 | 0.76 (0.58-0.99) | 0.04 |
| PWY-7385 | 1,3-propanediol biosynthesis (engineered) | 1.71 (1.03-2.82) | 0.04 | 1.02 (0.71-1.47) | 0.92 | 1.25 (0.96-1.64) | 0.10 |

| Pathways | Function | Stool collecting time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167) [2] | | **Overall (N = 301)[3]** | |
| | | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-7400 | L-arginine biosynthesis IV (archaebacteria) | 0.61 (0.39-0.96) | 0.03 | 1.08 (0.75-1.56) | 0.68 | 0.85 (0.66-1.11) | 0.24 |
| PWY-7560 | Methylerythritol phosphate pathway II | 1.08 (0.71-1.66) | 0.71 | 1.67 (1.13-2.46) | 0.01 | 1.40 (1.08-1.82) | 0.01 |
| PWY-841 | Superpathway of purine nucleotides de novo biosynthesis I | 0.60 (0.35-1.02) | 0.06 | 0.73 (0.50-1.05) | 0.09 | 0.70 (0.53-0.93) | 0.01 |
| PYRIDNUCSAL-PWY | NAD salvage pathway I | 1.63 (1.01-2.65) | 0.05 | 0.93 (0.64-1.35) | 0.70 | 1.18 (0.91-1.54) | 0.21 |
| SO4ASSIM-PWY | Sulfate reduction I (assimilatory) | 1.69 (1.07-2.67) | 0.03 | 0.85 (0.59-1.22) | 0.37 | 1.19 (0.92-1.56) | 0.19 |
| SULFATE-CYS-PWY | Superpathway of sulfate Assimilation and cysteine biosynthesis | 1.63 (1.03-2.60) | 0.04 | 0.80 (0.56-1.15) | 0.24 | 1.14 (0.88-1.49) | 0.32 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 50**: Association of grade ≥3 neutropenia with metabolic pathways by stool collection time

| | | Stool collection time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| 3-HYDROXYPHENYLACETATE-DEGRADATION-PWY | 4-hydroxyphenylacetate degradation | 1.58 (0.89-2.79) | 0.12 | 1.11 (0.75-1.66) | 0.60 | 1.33 (1.00-1.78) | 0.05 |
| ALLANTOINDEG-PWY | Superpathway of allantoin degradation in yeast | 1.37 (0.83-2.27) | 0.22 | 1.20 (0.82-1.77) | 0.35 | 1.37 (1.05-1.80) | 0.02 |
| ANAEROFRUCAT-PWY | Homolactic fermentation | 0.73 (0.42-1.28) | 0.28 | 0.63 (0.41-0.95) | 0.03 | 0.76 (0.57-1.01) | 0.06 |
| ARGSYNBSUB-PWY | L-arginine biosynthesis II (acetyl cycle) | 0.51 (0.29-0.90) | 0.02 | 1.10 (0.75-1.62) | 0.63 | 0.87 (0.66-1.14) | 0.32 |
| ARGSYN-PWY | L-arginine biosynthesis I (via L-ornithine) | 0.50 (0.28-0.87) | 0.02 | 1.10 (0.75-1.63) | 0.62 | 0.86 (0.65-1.13) | 0.28 |
| CENTFERM-PWY | Pyruvate fermentation to butanoate | 0.48 (0.27-0.86) | 0.01 | 0.74 (0.49-1.12) | 0.15 | 0.68 (0.51-0.91) | 0.01 |
| COBALSYN-PWY | Adenosylcobalamin salvage from Cobinamide I | 0.81 (0.49-1.34) | 0.41 | 0.73 (0.49-1.09) | 0.12 | 0.75 (0.57-0.99) | 0.04 |
| CRNFORCAT-PWY | Creatinine degradation I | 1.29 (0.74-2.26) | 0.36 | 1.43 (0.93-2.20) | 0.10 | 1.50 (1.13-2.00) | 0.01 |
| FUC-RHAMCAT-PWY | Superpathway of fucose and rhamnose degradation | 1.88 (1.04-3.42) | 0.04 | 0.91 (0.63-1.32) | 0.63 | 1.18 (0.89-1.56) | 0.24 |
| GLUCONEO-PWY | Gluconeogenesis I | 2.32 (1.14-4.74) | 0.02 | 0.69 (0.46-1.02) | 0.06 | 1.08 (0.82-1.42) | 0.58 |
| GLUDEG-I-PWY | GABA shunt | 1.94 (1.11-3.37) | 0.02 | 0.90 (0.62-1.32) | 0.59 | 1.24 (0.94-1.64) | 0.13 |
| GLYCOLYSIS-E-D | Superpathway of glycolysis and Entner-Doudoroff | 2.14 (1.15-3.95) | 0.02 | 0.70 (0.47-1.03) | 0.07 | 1.03 (0.79-1.35) | 0.83 |
| LACTOSECAT-PWY | Lactose and galactose degradation I | 0.32 (0.16-0.63) | $1.07 \times 10^{-3}$ | 0.90 (0.61-1.32) | 0.58 | 0.68 (0.50-0.93) | 0.02 |
| METH-ACETATE-PWY | methanogenesis from acetate | 0.42 (0.23-0.78) | 0.01 | 1.09 (0.73-1.63) | 0.67 | 0.74 (0.55-0.99) | 0.04 |
| NAD-BIOSYNTHESIS-II | NAD salvage pathway II | 1.76 (1.01-3.06) | 0.05 | 1.09 (0.73-1.60) | 0.68 | 1.28 (0.97-1.71) | 0.08 |
| P108-PWY | Pyruvate fermentation to propanoate I | 0.70 (0.41-1.20) | 0.20 | 0.60 (0.39-0.92) | 0.02 | 0.74 (0.56-0.98) | 0.03 |
| P162-PWY | L-glutamate degradation V (via hydroxyglutarate) | 0.77 (0.43-1.39) | 0.39 | 0.63 (0.42-0.96) | 0.03 | 0.65 (0.48-0.88) | 0.01 |
| P164-PWY | Purine nucleobases degradation I (anaerobic) | 0.78 (0.45-1.33) | 0.36 | 0.71 (0.48-1.05) | 0.08 | 0.75 (0.57-0.99) | 0.04 |
| POLYAMSYN-PWY | Superpathway of polyamine biosynthesis I | 1.73 (0.98-3.06) | 0.06 | 1.08 (0.75-1.58) | 0.67 | 1.34 (1.01-1.78) | 0.04 |
| PWY0-1297 | Superpathway of purine deoxyribonucleosides degradation | 1.34 (0.80-2.25) | 0.27 | 0.65 (0.43-0.96) | 0.03 | 0.94 (0.71-1.23) | 0.63 |
| PWY0-166 | Superpathway of pyrimidine deoxyribonucleotides de novo biosynthesis (E. coli) | 0.50 (0.26-0.96) | 0.04 | 1.01 (0.70-1.48) | 0.94 | 0.85 (0.64-1.12) | 0.25 |
| PWY0-881 | Superpathway of fatty acid biosynthesis I (E. coli) | 1.59 (0.90-2.82) | 0.11 | 1.07 (0.72-1.60) | 0.74 | 1.36 (1.03-1.81) | 0.03 |
| PWY-1269 | CMP-3-deoxy-D-manno-octulosonate biosynthesis I | 1.77 (0.95-3.28) | 0.07 | 1.50 (0.99-2.29) | 0.06 | 1.49 (1.10-2.00) | 0.01 |
| PWY-241 | C4 photosynthetic carbon assimilation cycle, NADP-ME type | 1.95 (1.08-3.51) | 0.03 | 0.83 (0.56-1.21) | 0.33 | 1.23 (0.93-1.63) | 0.16 |
| PWY3O-355 | Stearate biosynthesis III (fungi) | 1.77 (1.01-3.08) | 0.04 | 0.77 (0.50-1.17) | 0.22 | 1.20 (0.90-1.60) | 0.20 |
| PWY-4722 | creatinine degradation II | 0.71 (0.40-1.28) | 0.25 | 0.75 (0.48-1.19) | 0.22 | 0.70 (0.50-0.96) | 0.03 |
| PWY-5022 | 4-aminobutanoate degradation V | 2.05 (1.17-3.61) | 0.01 | 0.88 (0.61-1.28) | 0.51 | 1.23 (0.93-1.63) | 0.15 |
| PWY-5083 | NAD/NADH phosphorylation and dephosphorylation | 1.77 (1.01-3.09) | 0.05 | 0.98 (0.68-1.42) | 0.92 | 1.24 (0.94-1.63) | 0.13 |
| PWY-5088 | L-glutamate degradation VIII (to propanoate) | 1.78 (1.01-3.16) | 0.05 | 1.07 (0.73-1.57) | 0.72 | 1.17 (0.90-1.54) | 0.24 |
| PWY-5100 | Pyruvate fermentation to acetate and lactate II | 0.48 (0.26-0.90) | 0.02 | 0.96 (0.65-1.41) | 0.82 | 0.76 (0.56-1.02) | 0.07 |
| PWY-5265 | Peptidoglycan biosynthesis II (staphylococci) | 0.44 (0.20-0.93) | 0.03 | 0.81 (0.55-1.19) | 0.29 | 0.72 (0.52-0.97) | 0.03 |
| PWY-6143 | CMP-pseudaminate biosynthesis | 0.14 (0.04-0.49) | $2.07 \times 10^{-3}$ | 0.89 (0.61-1.28) | 0.52 | 0.74 (0.55-1.00) | 0.05 |
| PWY-6147 | 6-hydroxymethyl-dihydropterin diphosphate biosynthesis I | 1.07 (0.61-1.85) | 0.82 | 1.57 (1.05-2.34) | 0.03 | 1.33 (1.00-1.77) | 0.05 |
| PWY-6470 | Peptidoglycan biosynthesis V (β-lactam resistance) | 0.40 (0.22-0.72) | $2.27 \times 10^{-3}$ | 0.72 (0.49-1.06) | 0.09 | 0.61 (0.46-0.82) | $9.82 \times 10^{-4}$ |

| *Pathways* | Function | Stool collection time: Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Stool collection time: Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
|---|---|---|---|---|---|---|---|
| | | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-6471 | Peptidoglycan biosynthesis IV (Enterococcus faecium) | 0.34 (0.19-0.64) | $6.43x10^{-4}$ | 0.90 (0.61-1.32) | 0.59 | 0.70 (0.53-0.92) | 0.01 |
| PWY-6531 | mannitol cycle | 1.73 (1.00-2.98) | 0.05 | 0.94 (0.64-1.38) | 0.74 | 1.18 (0.90-1.56) | 0.23 |
| PWY-6590 | Superpathway of Clostridium acetobutylicum acidogenic fermentation | 0.49 (0.27-0.87) | 0.01 | 0.74 (0.49-1.12) | 0.15 | 0.69 (0.52-0.91) | 0.01 |
| PWY-6595 | Superpathway of guanosine nucleotides degradation (plants) | 0.43 (0.23-0.82) | 0.01 | 0.97 (0.67-1.40) | 0.87 | 0.87 (0.66-1.15) | 0.34 |
| PWY66-391 | Fatty acid β-oxidation VI (peroxisome) | 1.30 (0.79-2.14) | 0.30 | 0.67 (0.45-0.99) | 0.05 | 0.80 (0.60-1.05) | 0.11 |
| PWY66-399 | Gluconeogenesis III | 0.45 (0.25-0.81) | 0.01 | 0.71 (0.48-1.05) | 0.08 | 0.68 (0.51-0.89) | 0.01 |
| PWY66-409 | Superpathway of purine nucleotide salvage | 1.00 (0.59-1.67) | 0.99 | 0.54 (0.36-0.80) | $2.22x10^{-3}$ | 0.79 (0.61-1.03) | 0.09 |
| PWY-6737 | Starch degradation V | 0.58 (0.34-1.00) | 0.05 | 1.01 (0.68-1.49) | 0.98 | 0.81 (0.61-1.06) | 0.13 |
| PWY-7039 | Phosphatidate metabolism, as a signaling molecule | 1.80 (1.05-3.08) | 0.03 | 1.17 (0.81-1.70) | 0.40 | 1.34 (1.02-1.76) | 0.03 |
| PWY-7117 | C4 photosynthetic carbon assimilation cycle, PEPCK type | 1.81 (1.04-3.18) | 0.04 | 0.93 (0.63-1.35) | 0.69 | 1.23 (0.94-1.62) | 0.13 |
| PWY-7187 | Pyrimidine deoxyribonucleotides de novo biosynthesis II | 0.51 (0.28-0.93) | 0.03 | 1.00 (0.68-1.46) | 0.99 | 0.82 (0.62-1.08) | 0.16 |
| PWY-7196 | Superpathway of pyrimidine ribonucleosides salvage | 0.52 (0.30-0.89) | 0.02 | 1.18 (0.80-1.73) | 0.41 | 0.87 (0.67-1.14) | 0.32 |
| PWY-7282 | 4-amino-2-methyl-5-Phosphomethylpyrimidine biosynthesis (yeast) | 1.08 (0.60-1.93) | 0.80 | 0.64 (0.43-0.94) | 0.02 | 0.81 (0.62-1.07) | 0.13 |
| PWY-7288 | Fatty acid β-oxidation (peroxisome, yeast) | 1.31 (0.80-2.17) | 0.28 | 0.66 (0.44-0.98) | 0.04 | 0.79 (0.60-1.04) | 0.10 |
| PWY-7316 | dTDP-N-acetylviosamine biosynthesis | 1.25 (0.73-2.12) | 0.41 | 0.63 (0.39-1.00) | 0.05 | 0.87 (0.65-1.16) | 0.35 |
| PWY-7383 | anaerobic energy metabolism (invertebrates, cytosol) | 0.43 (0.24-0.78) | 0.01 | 0.76 (0.52-1.12) | 0.17 | 0.71 (0.54-0.94) | 0.02 |
| PWY-7385 | 1,3-propanediol biosynthesis (engineered) | 2.08 (1.12-3.85) | 0.02 | 0.93 (0.63-1.36) | 0.71 | 1.28 (0.96-1.70) | 0.09 |
| PWY-7400 | L-arginine biosynthesis IV (archaebacteria) | 0.50 (0.28-0.87) | 0.02 | 1.10 (0.75-1.62) | 0.64 | 0.86 (0.65-1.13) | 0.27 |
| PWY-7539 | 6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia) | 1.04 (0.60-1.80) | 0.89 | 1.56 (1.05-2.33) | 0.03 | 1.32 (0.99-1.76) | 0.06 |
| RHAMCAT-PWY | L-rhamnose degradation I | 2.01 (1.11-3.64) | 0.02 | 0.83 (0.56-1.23) | 0.35 | 1.08 (0.81-1.42) | 0.61 |
| THREOCAT-PWY | Superpathway of L-threonine metabolism | 1.86 (1.04-3.34) | 0.04 | 1.04 (0.70-1.54) | 0.86 | 1.36 (1.02-1.82) | 0.04 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Tables 51-52** present significant associations (P < 0.05) between GI microbial metabolic pathways and the grade ≥3 combined GI toxicity and nausea/vomiting. Among a total of 419 metabolic pathways investigated, 19 pathways showed significant associations with the grade ≥3 combined GI toxicity in the overall analysis. In stratified analysis by stool collection time, 30 and 10 metabolic pathways were significantly associated with the risk of grade ≥3 combined GI toxicity among patients with stool collected after breast cancer surgery and those who collected stool samples before breast cancer surgery and neoadjuvant chemotherapy, respectively (**Table 51**)

Eight of 30 metabolic pathways showed significant associations with an increased risk of grade ≥3 combined GI toxicity among patients with stool samples collected after breast cancer surgery. Two metabolic pathways that are related to reductive TCA cycle I and II had a significantly high risk of grade ≥3 combined GI toxicity, with ORs of 2.26 (95% CI: 1.22-4.19) for P23-PWY (*P*=0.01) and 2.05 (95% CI: 1.13-3.72) for PWY-5392 (*P*=0.02), respectively. Among patients who collected stool before breast cancer surgery and neoadjuvant chemotherapy, seven of 13 metabolic pathways were significantly associated with an increased risk of grade ≥3 combined GI toxicity. The 1,4-dihydroxy-6-naphthoate biosynthesis II (PWY-7371) showed the strongest association with a high risk of grade ≥3 combined GI toxicity, with ORs of 3.56 (95% CI: 1.56-8.16); $P = 2.64\times10^{-3}$. Three metabolic pathways that are related to phosphatidylglycerol biosynthesis I and II (PWY4FS-7 and PWY4FS-8) and superpathway of phospholipid biosynthesis I (PHOSLIPSYN-PWY) were positively associated with the risk of grade ≥3 combined hematological toxicity, with ORs of 2.27 (95% CI: 1.08-4.78) for PWY4FS-7 (*P* =0.03), 2.27 (95% CI: 1.08-4.78) for PWY4FS-8 (*P* =0.03) and 2.90 (95% CI: 1.30-6.46) for PHOSLIPSYN-PWY (*P* <0.01) respectively. Conversely, the lactose and galactose degradation I (LACTOSECAT-PWY) showed reduced associations with grade ≥3 combined GI toxicity in the analysis for patients whose stool samples were collected before breast cancer surgery and neoadjuvant chemotherapy, with OR of 0.41 (95% CI:0.18-0.94); and *P*=0.03. The significant associations of five pathways (PWY4FS-7, PWY4FS-8, PHOSLIPSYN-PWY, LACTOSECAT-PWY, and P23-PWY) also were found in the overall analyses.

The significant associations with an increased risk of grade ≥3 combined GI toxicity for the two metabolic pathways, including P23-PWY and PHOSLIPSYN-PWY, were also observed with the risk of

grade ≥3 nausea/vomiting in the overall analysis and the stratified analysis by stool collection time. In the overall analysis, the reductive TCA cycle I (P23-PWY) and the superpathway of phospholipid biosynthesis I (PHOSLIPSYN-PWY) were significantly associated with an increased risk of grade ≥3 nausea/vomiting, with ORs of 2.21 (95% CI: 1.41-3.46; $P = 5.53 \times 10^{-4}$) and 2.07 (95% CI: 1.27-3.37; $P = 3.57 \times 10^{-3}$), respectively. Two pathways that are related to phosphatidylglycerol biosynthesis I and II (PWY4FS-7 and PWY4FS-8) were positively associated with the risk of grade ≥3 nausea/vomiting in the analysis for patients with stool samples collected after breast cancer surgery (the same ORs and 95% CI: 3.15 (1.31-7.61); $P = 0.01$) and in the overall analysis (the same ORs and 95% CI: 1.94 (1.23-3.07); $P = 4.47 \times 10^{-3}$). Furthermore, only one pathway PWY-4981, which is related to the L-proline biosynthesis II (from arginine) had a significant association with a reduced risk of grade ≥3 nausea/vomiting in both overall analysis and stratified analysis by stool collection time, with an OR of 0.43 (95% CI: 0.27-0.70); and $P = 7.17 \times 10^{-4}$ (**Table 52**).

**Table 51:** Association of grade ≥3 combined GI toxicity with metabolic pathways by stool collection time

| | | Stool collection time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| BIOTIN-BIOSYNTHESIS-PWY | Biotin biosynthesis I | 0.79 (0.43-1.45) | 0.45 | 0.51 (0.28-0.92) | 0.03 | 0.69 (0.48-1.00) | 0.05 |
| CRNFORCAT-PWY | Creatinine degradation I | 0.86 (0.37-2.01) | 0.73 | 2.03 (1.18-3.48) | 0.01 | 1.21 (0.85-1.71) | 0.28 |
| DAPLYSINESYN-PWY | L-lysine biosynthesis I | 0.61 (0.31-1.21) | 0.16 | 0.48 (0.26-0.91) | 0.02 | 0.61 (0.41-0.9) | 0.01 |
| ENTBACSYN-PWY | Enterobactin biosynthesis | 0.93 (0.49-1.76) | 0.83 | 0.53 (0.28-0.99) | 0.05 | 0.75 (0.51-1.09) | 0.13 |
| FASYN-ELONG-PWY | Fatty acid elongation -- saturated | 0.77 (0.42-1.43) | 0.41 | 0.47 (0.25-0.86) | 0.01 | 0.67 (0.47-0.97) | 0.03 |
| FASYN-INITIAL-PWY | superpathway of fatty acid Biosynthesis initiation (E. coli) | 0.80 (0.42-1.51) | 0.49 | 0.45 (0.23-0.86) | 0.02 | 0.67 (0.46-0.98) | 0.04 |
| LACTOSECAT-PWY | Lactose and galactose degradation I | 0.41 (0.18-0.94) | 0.03 | 0.96 (0.55-1.67) | 0.89 | 0.62 (0.39-0.98) | 0.04 |
| P23-PWY | Reductive TCA cycle I | 1.92 (0.88-4.20) | 0.10 | 2.26 (1.22-4.19) | 0.01 | 1.63 (1.11-2.40) | 0.01 |
| P461-PWY | Hexitol fermentation to lactate, formate, ethanol and acetate | 0.81 (0.40-1.65) | 0.57 | 0.48 (0.24-1.00) | 0.05 | 0.72 (0.48-1.09) | 0.12 |
| PHOSLIPSYN-PWY | Superpathway of phospholipid biosynthesis I (bacteria) | 2.90 (1.30-6.46) | 0.01 | 1.46 (0.85-2.53) | 0.17 | 1.57 (1.05-2.35) | 0.03 |
| PPGPPMET-PWY | PpGpp biosynthesis | 0.45 (0.21-0.97) | 0.04 | 1.09 (0.62-1.92) | 0.77 | 0.79 (0.53-1.17) | 0.24 |
| PWY0-162 | Superpathway of pyrimidine Ribonucleotides de novo biosynthesis | 0.46 (0.22-0.95) | 0.04 | 1.18 (0.67-2.07) | 0.56 | 0.96 (0.65-1.40) | 0.82 |
| PWY0-862 | (5Z)-dodec-5-enoate biosynthesis | 0.77 (0.41-1.44) | 0.41 | 0.46 (0.24-0.85) | 0.01 | 0.66 (0.46-0.96) | 0.03 |
| PWY-4981 | L-proline biosynthesis II (from arginine) | 0.49 (0.24-1.00) | 0.05 | 0.55 (0.30-1.03) | 0.06 | 0.64 (0.43-0.94) | 0.02 |
| PWY4FS-7 | Phosphatidylglycerol biosynthesis I (plastidic) | 2.27 (1.08-4.78) | 0.03 | 1.51 (0.88-2.60) | 0.13 | 1.48 (1.01-2.17) | 0.04 |
| PWY4FS-8 | Phosphatidylglycerol biosynthesis II (non-plastidic) | 2.27 (1.08-4.78) | 0.03 | 1.51 (0.88-2.61) | 0.13 | 1.48 (1.01-2.17) | 0.04 |
| PWY-5005 | Biotin biosynthesis II | 3.01 (1.26-7.19) | 0.01 | 0.76 (0.43-1.34) | 0.35 | 1.19 (0.84-1.68) | 0.34 |
| PWY-5177 | Glutaryl-CoA degradation | 0.99 (0.51-1.92) | 0.97 | 0.56 (0.31-1.01) | 0.05 | 0.60 (0.41-0.87) | 0.01 |
| PWY-5392 | Reductive TCA cycle II | 1.07 (0.50-2.31) | 0.86 | 2.05 (1.13-3.72) | 0.02 | 1.36 (0.94-1.98) | 0.11 |
| PWY-5676 | acetyl-CoA fermentation to butanoate II | 0.79 (0.43-1.47) | 0.46 | 0.48 (0.25-0.91) | 0.02 | 0.65 (0.44-0.96) | 0.03 |
| PWY-5850 | Superpathway of menaquinol-6 biosynthesis I | 0.94 (0.48-1.84) | 0.86 | 0.42 (0.22-0.81) | 0.01 | 0.67 (0.45-0.98) | 0.04 |
| PWY-5860 | Superpathway of Demethylmenaquinol-6 biosynthesis I | 0.99 (0.51-1.93) | 0.97 | 0.43 (0.22-0.84) | 0.01 | 0.69 (0.47-1.03) | 0.07 |
| PWY-5863 | Superpathway of phylloquinol biosynthesis | 1.01 (0.53-1.91) | 0.99 | 0.52 (0.27-0.98) | 0.04 | 0.78 (0.53-1.15) | 0.21 |
| PWY-5896 | Superpathway of menaquinol-10 biosynthesis | 0.94 (0.48-1.84) | 0.86 | 0.42 (0.22-0.81) | 0.01 | 0.67 (0.45-0.98) | 0.04 |
| PWY-5971 | Palmitate biosynthesis II (bacteria and plants) | 0.79 (0.42-1.51) | 0.48 | 0.56 (0.32-1.00) | 0.05 | 0.73 (0.50-1.05) | 0.09 |
| PWY-5989 | Stearate biosynthesis II (bacteria and plants) | 0.77 (0.41-1.45) | 0.42 | 0.46 (0.24-0.87) | 0.02 | 0.66 (0.45-0.95) | 0.03 |
| PWY5F9-12 | Biphenyl degradation | 0.90 (0.29-2.83) | 0.86 | 1.73 (1.05-2.86) | 0.03 | 1.16 (0.82-1.63) | 0.41 |

| Pathways | Function | Stool collection time: Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Stool collection time: Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
|---|---|---|---|---|---|---|---|
| | | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-6263 | Superpathway of menaquinol-8 biosynthesis II | 2.74 (1.17-6.40) | 0.02 | 0.91 (0.53-1.55) | 0.72 | 1.32 (0.92-1.89) | 0.13 |
| PWY-6282 | Palmitoleate biosynthesis I (from (5Z)-dodec-5-enoate) | 0.77 (0.41-1.45) | 0.42 | 0.46 (0.24-0.86) | 0.01 | 0.67 (0.46-0.96) | 0.03 |
| PWY-6471 | Peptidoglycan biosynthesis IV (Enterococcus faecium) | 0.65 (0.33-1.28) | 0.21 | 1.98 (1.13-3.47) | 0.02 | 1.16 (0.80-1.70) | 0.43 |
| PWY-6519 | 8-amino-7-oxononanoate biosynthesis I | 0.79 (0.43-1.47) | 0.46 | 0.48 (0.26-0.89) | 0.02 | 0.68 (0.47-0.99) | 0.04 |
| PWY66-399 | Gluconeogenesis III | 0.80 (0.41-1.58) | 0.52 | 2.32 (1.19-4.53) | 0.01 | 1.35 (0.91-2.00) | 0.14 |
| PWY66-400 | Glycolysis VI (metazoan) | 0.49 (0.24-0.99) | 0.05 | 0.66 (0.35-1.24) | 0.20 | 0.70 (0.48-1.02) | 0.06 |
| PWY-6737 | Starch degradation V | 0.95 (0.49-1.83) | 0.87 | 2.01 (1.05-3.84) | 0.03 | 1.26 (0.85-1.87) | 0.26 |
| PWY-7039 | Phosphatidate metabolism, as a signaling molecule | 1.99 (1.01-3.92) | 0.05 | 0.89 (0.48-1.68) | 0.73 | 1.12 (0.80-1.58) | 0.51 |
| PWY-7228 | Superpathway of guanosine nucleotides de novo biosynthesis I | 1.03 (0.52-2.04) | 0.94 | 0.54 (0.30-0.97) | 0.04 | 0.78 (0.53-1.16) | 0.22 |
| PWY-7328 | Superpathway of UDP-glucose-derived O-antigen building blocks biosynthesis | 0.75 (0.38-1.46) | 0.39 | 0.49 (0.25-0.95) | 0.04 | 0.68 (0.45-1.01) | 0.06 |
| PWY-7371 | 1,4-dihydroxy-6-naphthoate biosynthesis II | 3.56 (1.56-8.16) | $2.64 \times 10^{-3}$ | 0.82 (0.47-1.43) | 0.48 | 1.20 (0.86-1.67) | 0.29 |
| PWY-7383 | Anaerobic energy metabolism (invertebrates, cytosol) | 0.79 (0.40-1.55) | 0.49 | 2.40 (1.29-4.45) | 0.01 | 1.39 (0.95-2.02) | 0.09 |
| PWY-7388 | Octanoyl-[acyl-carrier protein] biosynthesis (mitochondria, yeast) | 0.84 (0.45-1.58) | 0.59 | 0.51 (0.27-0.98) | 0.04 | 0.74 (0.50-1.09) | 0.12 |
| PWY-7664 | Oleate biosynthesis IV (anaerobic) | 0.77 (0.42-1.43) | 0.41 | 0.46 (0.25-0.86) | 0.01 | 0.67 (0.46-0.96) | 0.03 |
| PWYG-321 | Mycolate biosynthesis | 0.85 (0.45-1.58) | 0.60 | 0.49 (0.26-0.92) | 0.03 | 0.73 (0.51-1.05) | 0.09 |
| RUMP-PWY | Formaldehyde oxidation I | 0.38 (0.16-0.88) | 0.02 | 1.15 (0.66-1.99) | 0.63 | 0.86 (0.58-1.25) | 0.42 |
| SO4ASSIM-PWY | Sulfate reduction I (assimilatory) | 1.11 (0.54-2.27) | 0.78 | 0.54 (0.29-1.00) | 0.05 | 0.73 (0.49-1.08) | 0.11 |
| TEICHOICACID-PWY | Teichoic acid (poly-glycerol) biosynthesis | 0.36 (0.14-0.90) | 0.03 | 1.21 (0.70-2.08) | 0.49 | 0.85 (0.58-1.25) | 0.41 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 52:** Association of grade ≥3 combined nausea/vomiting with metabolic pathways by stool collection time

| | | Stool collection time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167)[2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| ARGSYNBSUB-PWY | L-arginine biosynthesis II (acetyl cycle) | 1.31 (0.59-2.89) | 0.51 | 3.36 (1.36-8.30) | 0.01 | 1.43 (0.91-2.25) | 0.12 |
| ARGSYN-PWY | L-arginine biosynthesis I (via L-ornithine) | 1.28 (0.58-2.81) | 0.54 | 3.02 (1.26-7.23) | 0.01 | 1.37 (0.88-2.15) | 0.16 |
| ARO-PWY | Chorismate biosynthesis I | 1.24 (0.57-2.73) | 0.59 | 2.60 (1.05-6.47) | 0.04 | 1.31 (0.83-2.07) | 0.24 |
| BIOTIN-BIOSYNTHESIS-PWY | Biotin biosynthesis I | 0.69 (0.35-1.37) | 0.29 | 0.29 (0.12-0.68) | $4.28x10^{-3}$ | 0.60 (0.39-0.91) | 0.02 |
| CALVIN-PWY | Calvin-Benson-Bassham cycle | 1.39 (0.62-3.08) | 0.42 | 2.48 (1.04-5.90) | 0.04 | 1.35 (0.89-2.06) | 0.16 |
| CRNFORCAT-PWY | creatinine degradation I | 0.63 (0.19-2.04) | 0.44 | 2.59 (1.27-5.25) | 0.01 | 1.23 (0.84-1.79) | 0.28 |
| FASYN-ELONG-PWY | Fatty acid elongation -- saturated | 0.72 (0.36-1.41) | 0.33 | 0.31 (0.13-0.71) | 0.01 | 0.62 (0.41-0.93) | 0.02 |
| FASYN-INITIAL-PWY | Superpathway of fatty acid biosynthesis initiation (E. coli) | 0.68 (0.33-1.39) | 0.29 | 0.30 (0.13-0.73) | 0.01 | 0.60 (0.39-0.93) | 0.02 |
| FUCCAT-PWY | Fucose degradation | 0.79 (0.37-1.68) | 0.54 | 2.08 (1.00-4.32) | 0.05 | 1.19 (0.77-1.83) | 0.44 |
| GLUTORN-PWY | L-ornithine biosynthesis | 1.73 (0.73-4.08) | 0.21 | 3.12 (1.25-7.81) | 0.02 | 1.64 (1.02-2.64) | 0.04 |
| GLYCOGENSYNTH-PWY | Glycogen biosynthesis I (from ADP-D-Glucose) | 1.07 (0.52-2.22) | 0.85 | 2.09 (1.01-4.33) | 0.05 | 1.28 (0.84-1.95) | 0.25 |
| METH-ACETATE-PWY | Methanogenesis from acetate | 0.78 (0.34-1.81) | 0.57 | 2.53 (1.07-5.97) | 0.03 | 1.12 (0.69-1.80) | 0.65 |
| OANTIGEN-PWY | O-antigen building blocks biosynthesis (E. coli) | 1.82 (0.79-4.16) | 0.16 | 2.14 (1.05-4.38) | 0.04 | 1.48 (0.97-2.25) | 0.07 |
| P108-PWY | Pyruvate fermentation to propanoate I | 0.71 (0.33-1.51) | 0.37 | 2.21 (1.04-4.69) | 0.04 | 1.19 (0.77-1.83) | 0.43 |
| P23-PWY | Reductive TCA cycle I | 3.91 (1.31-11.67) | 0.01 | 3.32 (1.45-7.60) | $4.44x10^{-3}$ | 2.21 (1.41-3.46) | $5.53x10^{-4}$ |
| P42-PWY | Incomplete reductive TCA cycle | 2.60 (0.84-8.03) | 0.10 | 2.21 (1.06-4.61) | 0.03 | 1.70 (1.09-2.66) | 0.02 |
| PHOSLIPSYN-PWY | Superpathway of phospholipid biosynthesis I (bacteria) | 3.07 (1.21-7.79) | 0.02 | 2.64 (1.11-6.27) | 0.03 | 2.07 (1.27-3.37) | $3.57x10^{-3}$ |
| POLYAMINSYN3-PWY | Superpathway of polyamine biosynthesis II | 2.39 (0.85-6.71) | 0.10 | 2.23 (1.11-4.46) | 0.02 | 1.79 (1.13-2.84) | 0.01 |
| POLYISOPRENSYN-PWY | Polyisoprenoid biosynthesis (E. coli) | 0.69 (0.32-1.49) | 0.34 | 0.44 (0.20-0.97) | 0.04 | 0.70 (0.45-1.08) | 0.10 |
| PPGPPMET-PWY | PpGpp biosynthesis | 0.28 (0.10-0.80) | 0.02 | 1.53 (0.82-2.87) | 0.18 | 0.87 (0.57-1.33) | 0.53 |
| PWY0-1296 | purine ribonucleosides degradation | 1.41 (0.64-3.10) | 0.39 | 2.17 (1.02-4.61) | 0.04 | 1.28 (0.83-1.96) | 0.26 |
| PWY0-162 | Superpathway of pyrimidine ribonucleotides de novo biosynthesis | 0.44 (0.20-0.99) | 0.05 | 1.31 (0.66-2.63) | 0.44 | 0.96 (0.62-1.47) | 0.84 |
| PWY0-862 | (5Z)-dodec-5-enoate biosynthesis | 0.71 (0.35-1.43) | 0.34 | 0.31 (0.13-0.71) | 0.01 | 0.61 (0.40-0.93) | 0.02 |
| PWY-4321 | L-glutamate degradation IV | 0.57 (0.19-1.67) | 0.31 | 0.73 (0.30-1.79) | 0.50 | 0.57 (0.33-0.99) | 0.04 |
| PWY-4981 | L-proline biosynthesis II (from arginine) | 0.14 (0.03-0.57) | 0.01 | 0.31 (0.12-0.83) | 0.02 | 0.43 (0.27-0.70) | $7.17x10^{-4}$ |
| PWY4FS-7 | Phosphatidylglycerol biosynthesis I (plastidic) | 2.35 (0.98-5.63) | 0.06 | 3.15 (1.30-7.60) | 0.01 | 1.94 (1.23-3.07) | $4.47x10^{-3}$ |
| PWY4FS-8 | Phosphatidylglycerol biosynthesis II (non-plastidic) | 2.35 (0.98-5.63) | 0.06 | 3.15 (1.31-7.61) | 0.01 | 1.94 (1.23-3.07) | $4.47x10^{-3}$ |
| PWY-5005 | biotin biosynthesis II | 4.40 (1.48-13.13) | 0.01 | 0.55 (0.25-1.20) | 0.13 | 1.16 (0.78-1.73) | 0.47 |
| PWY-5154 | L-arginine biosynthesis III (via N-acetyl-L-citrulline) | 1.79 (0.79-4.04) | 0.16 | 2.48 (1.11-5.54) | 0.03 | 1.52 (0.98-2.36) | 0.06 |

| Pathways | Function | Stool collection time: Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167) [2] | | **Overall (N = 301)[3]** | |
|---|---|---|---|---|---|---|---|
| | | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-5392 | Reductive TCA cycle II | 1.54 (0.62-3.82) | 0.35 | 2.66 (1.28-5.55) | 0.01 | 1.66 (1.09-2.53) | 0.02 |
| PWY-5505 | L-glutamate and L-glutamine biosynthesis | 2.24 (0.75-6.67) | 0.15 | 2.18 (1.05-4.52) | 0.04 | 1.56 (0.99-2.45) | 0.06 |
| PWY-5654 | 2-amino-3-carboxymuconate semialdehyde degradation to 2-oxopentenoate | 1.42 (0.54-3.72) | 0.47 | 2.90 (1.06-7.91) | 0.04 | 1.06 (0.66-1.70) | 0.82 |
| PWY-5850 | Superpathway of menaquinol-6 biosynthesis I | 0.75 (0.34-1.66) | 0.47 | 0.42 (0.19-0.92) | 0.03 | 0.66 (0.43-1.02) | 0.06 |
| PWY-5860 | Superpathway of demethylmenaquinol-6 biosynthesis I | 0.76 (0.34-1.70) | 0.51 | 0.44 (0.20-0.97) | 0.04 | 0.68 (0.44-1.06) | 0.09 |
| PWY-5896 | Superpathway of menaquinol-10 biosynthesis | 0.75 (0.34-1.66) | 0.47 | 0.42 (0.19-0.92) | 0.03 | 0.66 (0.43-1.02) | 0.06 |
| PWY-5971 | Palmitate biosynthesis II (bacteria and plants) | 0.65 (0.31-1.35) | 0.25 | 0.43 (0.21-0.89) | 0.02 | 0.65 (0.43-0.98) | 0.04 |
| PWY-5989 | Stearate biosynthesis II (bacteria and plants) | 0.70 (0.34-1.41) | 0.32 | 0.30 (0.13-0.72) | 0.01 | 0.59 (0.39-0.91) | 0.02 |
| PWY5F9-12 | Biphenyl degradation | 0.63 (0.12-3.22) | 0.58 | 2.23 (1.16-4.29) | 0.02 | 1.19 (0.82-1.73) | 0.37 |
| PWY-6125 | Superpathway of guanosine nucleotides de novo biosynthesis II | 1.00 (0.44-2.28) | 1.00 | 0.32 (0.15-0.71) | 0.01 | 0.59 (0.37-0.93) | 0.02 |
| PWY-6263 | Superpathway of menaquinol-8 biosynthesis II | 5.92 (1.45-24.17) | 0.01 | 1.05 (0.53-2.04) | 0.90 | 1.58 (1.04-2.39) | 0.03 |
| PWY-6282 | Palmitoleate biosynthesis I (from (5Z)-dodec-5-enoate) | 0.71 (0.35-1.44) | 0.35 | 0.31 (0.13-0.72) | 0.01 | 0.61 (0.40-0.93) | 0.02 |
| PWY-6317 | Galactose degradation I (Leloir pathway) | 1.30 (0.59-2.87) | 0.51 | 2.51 (1.15-5.45) | 0.02 | 1.46 (0.94-2.29) | 0.09 |
| PWY-6471 | Peptidoglycan biosynthesis IV (Enterococcus faecium) | 0.87 (0.38-2.01) | 0.75 | 2.86 (1.31-6.22) | 0.01 | 1.51 (0.97-2.34) | 0.07 |
| PWY-6519 | 8-amino-7-oxononanoate biosynthesis I | 0.69 (0.35-1.39) | 0.30 | 0.29 (0.12-0.69) | $4.92 \times 10^{-3}$ | 0.60 (0.40-0.92) | 0.02 |
| PWY-6527 | Stachyose degradation | 1.14 (0.52-2.54) | 0.74 | 2.58 (1.19-5.57) | 0.02 | 1.43 (0.92-2.23) | 0.11 |
| PWY-6549 | L-glutamine biosynthesis III | 1.29 (0.53-3.19) | 0.57 | 2.28 (1.01-5.14) | 0.05 | 1.18 (0.74-1.86) | 0.49 |
| PWY66-391 | Fatty acid β-oxidation VI (peroxisome) | 0.64 (0.24-1.67) | 0.36 | 0.38 (0.16-0.90) | 0.03 | 0.75 (0.49-1.16) | 0.20 |
| PWY66-399 | Gluconeogenesis III | 1.56 (0.63-3.85) | 0.34 | 2.55 (1.10-5.92) | 0.03 | 1.61 (1.01-2.55) | 0.04 |
| PWY66-400 | Glycolysis VI (metazoan) | 0.35 (0.14-0.89) | 0.03 | 0.82 (0.39-1.73) | 0.61 | 0.69 (0.46-1.06) | 0.09 |
| PWY66-422 | D-galactose degradation V (Leloir pathway) | 1.17 (0.54-2.55) | 0.69 | 2.26 (1.05-4.87) | 0.04 | 1.42 (0.91-2.22) | 0.12 |
| PWY-6737 | Starch degradation V | 1.29 (0.58-2.85) | 0.53 | 2.17 (1.01-4.68) | 0.05 | 1.41 (0.89-2.25) | 0.14 |
| PWY-6859 | All-trans-farnesol biosynthesis | 0.50 (0.21-1.23) | 0.13 | 0.38 (0.17-0.87) | 0.02 | 0.61 (0.39-0.97) | 0.04 |
| PWY-7013 | L-1,2-propanediol degradation | 0.28 (0.09-0.87) | 0.03 | 0.97 (0.44-2.14) | 0.94 | 0.60 (0.36-0.98) | 0.04 |
| PWY-7039 | Phosphatidate metabolism, as a signaling molecule | 2.39 (1.06-5.38) | 0.04 | 1.08 (0.54-2.14) | 0.83 | 1.24 (0.86-1.78) | 0.25 |

| | | Stool collection time | | | | | |
|---|---|---|---|---|---|---|---|
| | | Collection before breast cancer surgery & neoadjuvant chemotherapy (N=134)[1] | | Collection after breast cancer surgery (N= 167) [2] | | **Overall (N = 301)[3]** | |
| ***Pathways*** | **Function** | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* | *OR (95%CI)[1]* | *P* |
| PWY-7196 | Superpathway of pyrimidine ribonucleosides salvage | 0.85 (0.38-1.86) | 0.68 | 3.09 (1.39-6.84) | 0.01 | 1.49 (0.99-2.24) | 0.06 |
| PWY-7198 | Pyrimidine deoxyribonucleotides de novo biosynthesis IV | 0.33 (0.12-0.91) | 0.03 | 0.75 (0.39-1.44) | 0.38 | 0.71 (0.45-1.12) | 0.14 |
| PWY-7228 | Superpathway of guanosine nucleotides de novo biosynthesis I | 1.05 (0.46-2.39) | 0.91 | 0.31 (0.14-0.69) | $3.88 \times 10^{-3}$ | 0.59 (0.37-0.94) | 0.02 |
| PWY-7288 | Fatty acid β-oxidation (peroxisome, yeast) | 0.66 (0.25-1.72) | 0.39 | 0.36 (0.15-0.89) | 0.03 | 0.75 (0.48-1.17) | 0.21 |
| PWY-7357 | Thiamin formation from pyrithiamine and oxythiamine (yeast) | 2.18 (0.83-5.76) | 0.11 | 2.56 (1.11-5.92) | 0.03 | 1.58 (0.99-2.53) | 0.06 |
| PWY-7371 | 1,4-dihydroxy-6-naphthoate biosynthesis II | 8.60 (2.04-36.28) | $3.40 \times 10^{-3}$ | 0.56 (0.24-1.29) | 0.17 | 1.18 (0.81-1.72) | 0.40 |
| PWY-7383 | Anaerobic energy metabolism (invertebrates, cytosol) | 1.25 (0.55-2.83) | 0.59 | 3.30 (1.38-7.86) | 0.01 | 1.65 (1.06-2.55) | 0.03 |
| PWY-7388 | Octanoyl-[acyl-carrier protein] biosynthesis (mitochondria, yeast) | 0.68 (0.33-1.43) | 0.31 | 0.36 (0.15-0.84) | 0.02 | 0.66 (0.42-1.02) | 0.06 |
| PWY-7400 | L-arginine biosynthesis IV (archaebacteria) | 1.29 (0.58-2.83) | 0.53 | 3.06 (1.28-7.32) | 0.01 | 1.38 (0.89-2.17) | 0.15 |
| PWY-7664 | Oleate biosynthesis IV (anaerobic) | 0.71 (0.36-1.42) | 0.34 | 0.31 (0.13-0.71) | 0.01 | 0.61 (0.40-0.93) | 0.02 |
| PWYG-321 | Mycolate biosynthesis | 0.71 (0.35-1.44) | 0.34 | 0.35 (0.16-0.79) | 0.01 | 0.67 (0.45-1.01) | 0.06 |
| RUMP-PWY | Formaldehyde oxidation I | 0.27 (0.09-0.86) | 0.03 | 1.75 (0.88-3.50) | 0.11 | 0.90 (0.59-1.37) | 0.62 |
| TEICHOICACID-PWY | Teichoic acid (poly-glycerol) biosynthesis | 0.33 (0.11-0.98) | 0.05 | 1.20 (0.62-2.31) | 0.59 | 0.77 (0.50-1.19) | 0.24 |

[1] Models were adjusted for age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] Models were adjusted for the number of days from stool collection time to breast cancer surgery (days), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
[3] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, pre-existing hematological, nephrological and hepatological conditions, diagnosis delay, breast cancer subtypes, TNM cancer stage, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk).
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

## 3. Discussion

In this study, among 301 newly diagnosed Vietnamese breast cancer patients, we evaluated the associations of pre-chemotherapy GI microbial richness and composition, individual microbial taxa, and microbial metabolic pathways with chemotherapy-induced toxicity, including hematological and GI toxicities. We found high pre-chemotherapy GI microbial richness (Chao1 and Shannon indexes) and high abundances of members of class *Clostridia* (in which most of them were from families *Lachnospiraceae, Oscillospiraceae,* and *Ruminococcaceae*) were significantly associated with a reduced risk of severe combined hematological toxicity and neutropenia among breast cancer patients, particularly among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. Conversely, the enrichment of specific microbial taxa from families *Bacteroidaceae*, *Lachnospiraceae*, *Sporanaerobacteraceae*, *Fusobacteriaceae* was significantly associated with an increased risk of severe neutropenia among this sub-population. Moreover, we found a significantly increased association with severe nausea/vomiting for high GI microbial richness (Chao1 indexes) and composition (all beta diversity measurements). We also found that 157 species, 41 genera, and one family belonging to phylum *Firmicutes A,* were associated with an increased risk of severe nausea/vomiting among patients with stool samples collected after breast cancer surgery (all these patients had also received prophylactic antibody treatment). Furthermore, we found that GI microbiota functional capacity in methanogenesis from acetate, pyrimidine deoxyribonucleotide de novo biosynthesis and peptidoglycan biosynthesis II, IV and V (β-lactam resistance) were significantly associated with a reduced risk of severe hematological toxicities. In contrast, GI microbiota functional capacity in the superpathway of polyamine biosynthesis I and phosphatidate metabolism were significantly associated with an increased risk of severe hematological toxicities. In addition, we found that bacteria involved in the reductive TCA cycle I, phosphatidylglycerol biosynthesis I and II, and superpathway of phospholipid biosynthesis I were significantly associated with an increased risk of severe GI toxicities. A high abundance of GI microbiota involving in lactose and galactose and L-proline biosynthesis II (from arginine) was associated with the reduced the risk of severe combined GI toxicities and nausea/vomiting. Our findings suggest that overall GI microbial richness and

multiple microbes may influence the development of hematological and GI toxicities among breast cancer patients.

As mentioned in Chapter 1, the gut microbiota might help reduce the probability of developing chemotherapy-induced toxicities by maintaining barrier homeostasis and protecting against pathogen overgrowth during chemotherapy. Chemotherapeutic agents used in breast cancer treatment interact directly with mucosal and fecal microbiota via biliary excretion into the GI tract, which shortens villi mucosal, and damages the integrity of the mucosal epithelium, and increases intestinal permeability, resulting in commensal microbe translocation.[157,193] Lastly, chemotherapeutic agents may cause dysbiosis of the gut microbiota, disrupting commensal homeostasis in the intestine among cancer patients.[185] This pathological alteration in the microbiota composition might contribute to the development of mucositis as well as the development of neoplastic and autoimmune disorders that may exacerbate systemic toxicity effects and lead to life-threatening systemic infections.[157,193]

Few clinical studies have investigated the pre-chemotherapy GI microbiome's influence on chemotherapy-induced toxicities, including infection, febrile neutropenia, diarrhea, weight gain, and neurological side effects. However, these few studies have been inconclusive. Terrasse and colleagues conducted a study among 75 early-stage breast cancer patients and revealed that alpha diversity and microbial taxa in pre-chemotherapy fecal samples were not significantly associated with side effects/toxicities (i.e., neurological, gastrointestinal, rheumatological toxicities).[199] Conversely, a previous study of 97 patients with AML reported that a higher Shannon index and higher relative abundance of *Porphyromonadaceae* at baseline were associated with an increased probability of remaining infection-free during neutropenia.[162] In our study, we found that high alpha diversity indexes were significantly associated with a reduced risk of severe combined hematological toxicity (i.e., for Chao1 and Shannon index) and severe neutropenia (i.e., only Chao 1). These significant associations of severe combined hematological toxicity or neutropenia with Chao1 index were observed primarily among patients with stool samples collected before breast cancer surgery and neoadjuvant chemotherapy. Shannon and inverse Simpson indexes were significantly associated with a reduced risk of severe neutropenia among the same group of patients. In addition, in the stratified analyses by stool collection

time, high abundances of specific taxa, most of them from families *Lachnospiraceae, Oscillospiraceae,* and *Ruminococcaceae* (belonging to the class *Clostridia,* phylum *Firmicutes A*) such as *Coprococcus eutactus, Dorea scindens, Eubacterium E hallii A, Eubacterium G ventriosum, Intestinimonas butyriciproducens, Faecalibacterium prausnitzii J,* and *Ruminococcus D bicirculans*, were significantly associated with a reduced risk of severe neutropenia among patients whose stool samples were collected before breast cancer surgery. *Faecalibacterium prausnitzii,* a well-known butyrate-producing bacterium belonging to the family *Ruminococcaceae,* represents *Clostridium species*, which could exert anti-inflammation effects on human health via interacting with colonic immune directly or indirectly. The health benefit of *Faecalibacterium prausnitzii* has been demonstrated in recent studies.[246] Moreover, analyses for this subgroup of patients found significantly increased associations with severe neutropenia for some microbial taxa from families *Bacteroidaceae*, *Fusobacteriaceae* belonging to two phyla *Bacteroidota* (i.e., species *Bacteroides A mediterraneensis*), and *Fusobacteriota (*i.e., the family *Fusobacteriaceae* and its species *MGYG-HGUT-03919).* Some microbial taxa from families *Lachnospiraceae*, *Sporanaerobacteraceae* belonging to the phylum *Firmicutes A* (i.e., species *Sporanaerobacter acetigenes*, *Blautia hansenii, Blautia producta, Blautia sp003287895, Dorea scindens,* and *Tyzzerella sp000411335)* also significantly associated with an increased risk of severe neutropenia. It is noteworthy that most microbial taxa that had significant associations with severe neutropenia seem to be susceptible to breast cancer surgery and antibiotic exposure as their associations were primarily confined to one group of patients as defined by stool sample collection time except for the family *Synergistaceae.* The latter appears to be less susceptible to breast cancer surgery and antibiotic exposure in our study. In the overall analyses and subgroup analyses, the abundance of *Synergistaceae* was significantly associated with a reduced risk of severe neutropenia. However, this inverse association of *Synergistaceae* needs to be interpreted with caution. It requires future confirmation because some species within this family, such as *Cloacibacillus* species, have been considered opportunistic human pathogens associated with intestinal infections.[247,248]

Many previous studies explored the association between the gut microbiome and chemotherapy-induced GI toxicity with a predominant focus on diarrhea and proposed possible mechanisms. The changes in GI microbiota might diminish bacterial protective functions and cause intestinal damage, resulting in diarrhea.[189]

For example, decreases in the abundance of *Lactobacillus*, *Bifidobacterium*, *Bacteroides* species were observed while *Escherichia coli* and *Staphylococcus* species tended to be increased in cancer patients with diarrhea. Cancer patients with diarrhea also tended to have more methanogenic archaea, fecal calprotectin, MMP 3 and 9.[189] Another study among patients diagnosed with stage III CRC reported that differentiated microorganisms along with metabolic products, which are relevant to different metabolic pathways in the intestinal micro-ecosystem, were linked to chemotherapy-induced diarrhea.[197] That study found that patients with chemotherapy-induced diarrhea had lower gut microbial community richness and diversity, resulting in a predominance of the species *Klebsiella pneumoniae* and significant differences in specific microbial taxa, including *Proteobacteria, Enterobacteriales, Gammaproteobacteria, Enterobacteriaceae, Klebsiella, Clostridiales, Clostridia, Ruminococcaceae, Bacteroidetes, Bacteroidia, Bacteroidales, Bacteroides* and *Bacteroidaceae* compared with patients without chemotherapy-induced diarrhea.[197] In our study, severe chemotherapy-induced diarrhea was less common than severe nausea/vomiting among breast cancer patients. We focused on evaluating the association between the GI microbiome and the combined GI toxicity and nausea/vomiting. In general, we found no significant associations between the combined GI toxicity and nausea/vomiting with GI microbial richness, except for a significant association between high Chao1 index and a high risk of severe nausea/vomiting in the overall analysis. In addition, we only found that severe nausea/vomiting was significantly associated with variations in GI microbial composition (measured by beta diversity) among patients whose stool samples were collected after surgery suggesting that surgery, possibly due to its accompanied antibiotics treatment may have led to selective growth of some GI toxicity related microbes. Many GI microbial taxa were associated with severe nausea/vomiting among these patients. Approximately 83% of 245 taxa with significant associations with severe nausea/vomiting, including one family, 41 genera, and 157 species (e.g., *Coprococcus eutactus, Ruminiclostridium E siraeum, Ruminiclostridium C sp000435295,* and *Anaeromassilibacillus sp001305115*) were associated with an increased risk of severe nausea/vomiting among patients whose stool samples were collected after breast cancer surgery.

As depicted in **Figure 13**, most participants in our study, particularly patients whose stool samples were collected after breast cancer surgery, had a disrupted gut microbiota. The median time interval from breast cancer surgery to the first date of chemotherapy was approximately 24-25 days. It has been suggested that the

baseline composition of the gut microbiota was mostly restored within 1.5 months after antibiotic treatment, but several common species remained undetectable even after 180 days following the antibiotics treatment.[248,249] Our own data have shown that recovery of gut microbiome diversity takes time (Aim 2 results). We speculate that antibiotic treatment following breast cancer surgery led to gut dysbiosis with overgrown antibiotic-resistant bacteria or pathogens and depletion of beneficial bacteria that protect health. Chemotherapy exacerbates the antibiotics-induced bacterial imbalances, triggering a wide range of uncomfortable GI symptoms such as nausea, vomiting, constipation, diarrhea, or stomatitis. These findings suggest that in Vietnam and other countries where prophylactic antibiotics are commonly given to cancer patients during/after breast cancer surgery, proper measurements may be needed to reduce gut dysbiosis before administration of chemotherapy in order to reduce drug-related GI toxicities. Because delaying chemotherapy after breast cancer surgery may reduce survival chances, starting adjuvant chemotherapy as soon as clinically possible within 31 days of surgery in patients with early and locally advanced breast cancer is highly recommended.[250] Studies are needed to investigate whether diet interventions, prebiotic and/or probiotic supplements after breast cancer survey would improve the gut microbiome and reduce the chemotherapy toxicities.

In general, results from metabolic pathway analyses varied less by the time of stool sample collection. Our study found that the GI microbiota involved in the reductive TCA cycle I, phosphatidylglycerol biosynthesis I and II, and superpathway of phospholipid biosynthesis I were significantly associated with an increased risk of severe GI toxicities, particularly nausea/vomiting. We observed some taxa, such as the genus *Ruminiclostridium E* and four species, including *Ruminiclostridium E siraeum, Ruminiclostridium E sp003512525,* and two *MGYG-HGUT* species (i.e., *03844, 03494*) were associated with a significantly high risk of severe nausea/vomiting, whereas two species within the order *Erysipelotrichales,* (i.e., *Massiliomicrobiota timonensis* and *Absiella innocuum)* were significantly associated with a lower risk of severe nausea/vomiting in both overall analyses and analysis for patients with stool samples collected after breast cancer surgery. In the analysis for metabolic pathways, GI microbiota involved in methanogenesis from acetate, pyrimidine deoxyribonucleotides de novo biosynthesis and peptidoglycan biosynthesis II, IV and V (β-lactam resistance) were significantly associated with a reduced risk of severe hematological toxicities, whereas those involved in the superpathway of polyamine biosynthesis I and phosphatidate metabolism were significantly associated with

an increased risk of severe hematological toxicities. A high abundance of GI microbiota involved in lactose and galactose degradation I and L-proline biosynthesis II (from arginine) was associated with a reduced risk of severe hematological and GI toxicities. *Lactococcus lactis* is known to possess the metabolic pathway of lactose and galactose degradation I. In our analysis, a high abundance of *Lactococcus lactis* was inversely associated with the risk of severe combined GI toxicity (ORs and 95% CI: 0.51 (0.28-0.91); *P* =0.02) and nausea/vomiting (ORs and 95% CI: 0.49 (0.25-0.95); *P* =0.03), but all p values greater than 0.1 after FDR correction. These findings suggest that these bacteria and some other species such as *Coprococcus eutactus, Dorea scindens, Eubacterium E hallii A, Eubacterium G ventriosum, Intestinimonas butyriciproducens, Faecalibacterium prausnitzii J, and Ruminococcus D bicirculans* might have the potential to be served as probiotics for reducing chemotherapy-induced severe GI or hematological toxicity. Preclinical and clinical studies are warranted to explore this potential.

Our study is the largest study to evaluate the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients. As previously mentioned in Aim 2, the availability of blood and urine test results before and during each cycle of chemotherapy/hospital visit is a strength of this study. The chemotherapy-induced hematological toxicities were comprehensively and accurately captured during the first-line chemotherapy treatment and 90 follow-up days after the treatment by a systematic review of the medical charts. GI toxicities were identified through a combination of patient self-reported side effects during the two follow-ups and a review of clinical notes of treating physicians/nurses during each chemotherapy cycle/hospital visit cycle, which minimized the concern related to underestimated non-hematological chemotherapy-induced toxicities by clinicians. Stool sample collection, transportation, and storage followed a standard protocol. In addition, as previously described in Aim 3, we performed the shotgun metagenomic sequencing and used a human bacterial genome from the UHGG collection as the reference, which enhances the taxonomic resolution and accuracy of our metagenome-based study. Last but not least, functional profiling was performed in our study. Our findings on the association between GI microbial metabolic pathways and chemotherapy-induced toxicities provide additional insight into potential biological mechanisms underlying the associations between GI microbiota and chemotherapy-induced toxicity. However, there are several limitations to our study. First, as aforementioned, over half of participants donated stool samples after

breast cancer surgery and prophylactic antibiotics, introducing heterogeneity in baseline microbiome composition. Secondly, the small sample size for stratified analysis by stool collection time is another limitation. These should be taken into consideration when interpreting our findings from the overall analysis or analyses stratified by patients' stool samples collection time.

In conclusion, our metagenomics study found that overall GI microbial richness and several microbial taxa may influence the development of hematological and GI toxicities among breast cancer patients. Our findings suggest that the gut microbiome may be used as a biomarker to predict chemotherapy-induced toxicities or serve as prebiotics/probiotics to prevent and reduce chemotherapy-induced toxicities. However, future larger studies are needed to validate our findings.

# CHAPTER 5

## SPECIFIC AIM 4

**Drug-microbiome interaction and chemotherapy-induced toxicity**

To explore drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients.

### 1. Methods

#### 1.1. Parent Study

This study included 396 newly diagnosed Vietnamese breast cancer patients who have received neoadjuvant or adjuvant chemotherapy during breast cancer treatment as included in Aim 1. In addition, details of breast cancer case recruitment in Aim 1 have been described.

#### 1.2. Population Selection

This analysis included the same study population as included in Aim 3. Briefly, 301 newly diagnosed breast cancer patients in this study provided stool samples at baseline and received neoadjuvant or adjuvant chemotherapy at Vietnam National Cancer Hospital (K Hospitals) and Hanoi Oncology hospital during the follow-up period.

#### 1.3. Outcome Assessment

Four chemotherapy-induced toxicities, including combined hematological toxicity, combined GI toxicity, neutropenia, and nausea/vomiting, as defined in Aim 1, are the primary outcomes for Aim 4. In order to ensure chemotherapy-induced toxicities only occurred after using chemotherapy agents, the highest grade of toxicities was correspondingly captured according to each chemotherapy agent. The combined hematological toxicity refers to having any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia, while the combined GI toxicity was combined to include the four symptoms: nausea, vomiting, diarrhea, constipation, or stomatitis. The two combined types of toxicities and their specific toxicities were grouped as

dichotomous variables (i.e., grade ≥3 vs. grade <3) to evaluate the associations with pre-chemotherapy GI microbiome.

1.4. Exposure Assessment

*Microbiome profiling*

DNA extraction, library preparation, sequencing, and data processing for 301 stool samples were conducted in one batch. Detailed descriptions are given in Aim 2. Sequencing reads were subjected to quality trimming via Trimmomatic v0.39, and Bowtie2 v2.3.0 was used to remove human reads.[229,230] Taxonomic profiling was conducted using Kraken v2.1.1 and Bracken v2.6 to estimate the absolute abundance of microbial taxa with human bacterial genomes from the UHGG collection as the reference.[231-233] Functional profiling was performed using HUMAnN2 v0.11.1 to estimate the relative abundance of GI microbial metabolic pathways.[245]

*Chemotherapeutic agents*

Trained study staff reviewed medical charts and abstracted information on chemotherapeutic regimens and chemotherapeutic medicines during the first-line chemotherapy treatment and directly entered it into the REDCap data management platform. Eleven chemotherapeutic agents, including cyclophosphamide, doxorubicin, epirubicin, paclitaxel, docetaxel, 5-FU, carboplatin, capecitabine, cisplatin, gemcitabine, and vinorelbine, which breast cancer patients received during the first-line treatment, were identified. These 11 chemotherapeutic agents were grouped into a dichotomous variable (used/not used) regardless of chemotherapeutic regimens. Due to the use of carboplatin (4.7%), capecitabine (1.7%), gemcitabine (0.7%), vinorelbine (0.3%), and cisplatin (0.0%) use were rare, and we excluded these five chemotherapeutic agents from our analysis. We also generate dichotomous variables (used/not used) for the use of overall anthracycline (i.e., doxorubicin or epirubicin) and overall taxane (i.e., paclitaxel or docetaxel). Because the sequential anthracycline and taxane regimen was the most common in our study, we included this variable in our analyses as the incorporation of multiple chemotherapeutic agents (i.e., incorporation of taxane with anthracycline-based regimen).

1.5. Statistical Analysis

In this study, common taxa were defined if present in >50% of samples; rare taxa were defined if shown in <50% of samples; We limited our analysis of rare taxa to those present in 10-50% of samples. Centered log-ratio (clr) transformation[239] was utilized to normalize the absolute abundance of taxa at each taxonomic level from phylum to species, with zeros replaced by the minimum read count value of the whole dataset.[239] We also limited our analysis of GI microbial metabolic pathways to those present in >10% of samples. Finally, the arcsine square root (asr) transformation was used to normalize the relative abundance of microbial metabolic pathways.

To explore drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients, we evaluated individual drug-microbiome interactions on the associations between pre-chemotherapy GI microbial taxa and metabolic pathways with chemotherapy-induced toxicity. We performed logistic regression analyses stratified by specific chemotherapeutic medication/drug, including cyclophosphamide, 5-FU, overall anthracycline, doxorubicin, epirubicin, overall taxane, paclitaxel, and docetaxel, to evaluate associations between chemotherapy-induced toxicities and the clr-transformed taxa abundance as well as the asr-transformed relative abundance of metabolic pathways. Second, we evaluated multiple drugs-microbiome interactions on the association between pre-chemotherapy GI microbial taxa and metabolic pathways with chemotherapy-induced toxicity. We performed logistic regression analyses stratified by the sequential anthracycline and taxane regimen. The log-likelihood ratio test was used to assess multiplicative interaction by comparing multivariable logistic regression models with and without the cross-product terms of these variables with each taxon. ORs and respective 95% CIs per SD increase in clr-transformed absolute abundance of microbial taxa and per SD increase in asr-transformed relative abundance of metabolic pathways and p-value for interaction, were calculated in models adjusted for potential confounders. The latter included stool collection time, age group at baseline, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, TNM cancer stage, breast cancer subtypes, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake and physical activity (MET/wk) were included in all adjusted models. We also did not mutually adjust for receiving sequential

anthracycline and taxane and specific chemotherapeutic medications, including cyclophosphamide, overall anthracycline, doxorubicin, epirubicin, overall taxane, paclitaxel, and docetaxel into multivariable models because they were highly correlated. In the analyses for GI microbial taxa, FDR of the p-value for interaction was calculated at each taxonomic level separately for overall or stratified analyses to account for multiple testing; The association with FDR-corrected p-value ($P_{FDR}$) of <0.1 was considered statistically significant. In the analyses for GI microbial metabolic pathways, all adjusted p-values >0.1 after FDR correction. Thus, all associations with unadjusted $P<0.05$ were present. All statistical analyses were performed using R version 3.6.3.

## 2. Results

**Table 53**: Highest grade of hematological toxicities by chemotherapeutic agents

| Regimens containing | N | Combined hematological toxicity [a] | | | Neutropenia | | |
|---|---|---|---|---|---|---|---|
| | | Grade <3 | Grade ≥3 | *P*[1] | Grade <3 | Grade ≥3 | *P*[1] |
| | | **Anthracycline-induced toxicities** | | | | | |
| **Doxorubicin** | | | | | | | |
| No | 122 | 86 (70.5) | 36 (29.5) | 0.74 | 97 (79.5) | 25 (20.5) | 0.41 |
| Yes | 179 | 123 (68.7) | 56 (31.3) | | 135 (75.4) | 44 (24.6) | |
| **Epirubicin** | | | | | | | |
| No | 220 | 153 (69.6) | 67 (30.5) | 0.95 | 169 (76.8) | 51 (23.2) | 0.86 |
| Yes | 81 | 56 (69.1) | 25 (30.9) | | 63 (77.8) | 18 (22.2) | |
| | | **Taxane-induced toxicities** | | | | | |
| **Paclitaxel** | | | | | | | |
| No | 109 | 75 (68.8) | 34 (31.2) | 0.14 | 87 (79.8) | 22 (20.2) | 0.85 |
| Yes | 192 | 147 (76.6) | 45 (23.4) | | 155 (80.7) | 37 (19.3) | |
| **Docetaxel** | | | | | | | |
| No | 225 | 165 (73.3) | 60 (26.7) | 0.78 | 179 (79.6) | 46 (20.4) | 0.53 |
| Yes | 76 | 57 (75.0) | 19 (25.0) | | 63 (82.9) | 13 (17.1) | |
| | | **Overall chemotherapy-induced toxicities** | | | | | |
| **Cyclophosphamide** | | | | | | | |
| No | 35 | 23 (65.7) | 12 (34.3) | 0.53 | 28 (80.0) | 7 (20.0) | 0.15 |
| Yes | 266 | 160 (60.2) | 106 (39.8) | | 181 (68.1) | 31 (31.9) | |
| **5-FU** | | | | | | | |
| No | 235 | 143 (60.9) | 92 (39.1) | 0.97 | 163 (69.4) | 72 (30.6) | 0.96 |
| Yes | 66 | 40 (60.6) | 26 (39.4) | | 46 (69.7) | 20 (30.3) | |
| **Sequential anthracycline and taxane** | | | | | | | |
| No | 89 | 59 (66.3) | 30 (33.7) | 0.21 | 68 (76.4) | 21 (23.6) | 0.09 |
| Yes | 212 | 124 (58.5) | 88 (41.5) | | 141 (66.5) | 71 (33.5) | |

[a] Combined hematological toxicity refers to having any of the four toxicities: neutropenia, anemia, lymphopenia, or thrombocytopenia.
[1] p-value for chi-square tests

Among 301 breast cancer patients who donated stool samples and received chemotherapy, no significant differences were observed for all hematological toxicities and GI toxicities by using six chemotherapeutic agents regardless of specific regimens at the first-line chemotherapy (**Tables 53-54**).

**Table 54:** Highest grade of GI toxicities by chemotherapeutic agents

| Regimens containing | N | Combined GI toxicity [a] | | | Nausea/vomiting | | |
|---|---|---|---|---|---|---|---|
| | | Grade <3 | Grade ≥3 | *P*[1] | Grade <3 | Grade ≥3 | *P*[1] |
| | | **Anthracycline-induced toxicities** | | | | | |
| **Doxorubicin** | | | | | | | |
| No | 122 | 113 (92.6) | 9 (7.4) | 0.97 | 116 (95.1) | 6 (4.9) | 0.65 |
| Yes | 179 | 166 (92.7) | 13 (7.3) | | 168 (93.9) | 11 (61) | |
| **Epirubicin** | | | | | | | |
| No | 220 | 207 (94.1) | 13 (5.9) | 0.12 | 209 (95.0) | 11 (5.0) | 0.42 |
| Yes | 81 | 72 (88.9) | 9 (11.1) | | 75 (92.6) | 6 (7.4) | |
| | | **Taxane-induced toxicities** | | | | | |
| **Paclitaxel** | | | | | | | |
| No | 109 | 104 (95.4) | 5 (4.6) | 0.86 | 105 (96.3) | 4 (3.7) | 0.75 |
| Yes | 192 | 184 (95.8) | 8 (4.2) | | 186 (96.7) | 6 (3.1) | |
| **Docetaxel** | | | | | | | |
| No | 225 | 217 (96.4) | 8 (3.6) | 0.26 | 219 (97.3) | 6 (2.7) | 0.28 |
| Yes | 76 | 71 (93.4) | 5 (4.3) | | 72 (94.7) | 4 (5.3) | |
| | | **Overall chemotherapy-induced toxicities** | | | | | |
| **Cyclophosphamide** | | | | | | | |
| No | 35 | 30 (87.5) | 5 (14.3) | 0.85 | 32 (91.4) | 3 (8.6) | 1.00 |
| Yes | 266 | 231 (86.8) | 35 (14.2) | | 237 (89.1) | 29 (10.9) | |
| **5-FU** | | | | | | | |
| No | 235 | 206 (87.7) | 29 (12.3) | 0.36 | 211 (89.8) | 24 (10.2) | 0.66 |
| Yes | 66 | 55 (83.3) | 11 (16.7) | | 58 (87.9) | 8 (12.1) | |
| **Sequential anthracycline and taxane** | | | | | | | |
| No | 89 | 77 (86.5) | 12 (13.5) | 0.95 | 80 (29.7) | 9 (28.1) | 0.85 |
| Yes | 212 | 184 (86.8) | 28 (13.2) | | 189 (70.3) | 23 (71.9) | |

[a] Combined GI toxicity refers to having any of the five symptoms: nausea, vomiting, diarrhea, stomatitis, or constipation.
[1] p-value for chi-square tests

### Interactions between individual chemotherapeutic agents and GI microbial taxa on grade ≥3 chemotherapy-induced toxicities

**Table 55-56** shows significant interactions after FDR correction ($P_{FDR} < 0.1$) between doxorubicin- and docetaxel-containing regimens and GI microbial taxa on grade ≥3 combined hematological toxicity. In phylum *Firmicutes A*, we found the family *Clostridiaceae* was associated with an increased risk of grade ≥3 combined hematological toxicity among patients who received doxorubicin-containing regimens, with OR and 95% CIs of 1.66 (1.14-2.41); $P_{for\ interaction}$=0.001 (**Table 55**). In addition, we found that the family *Aeromonadaceae* within class *Gammaproteobacteria (*phylum *Proteobacteria)* was associated with an increased risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 7.36 (1.09-49.86); $P_{for\ interaction}$=0.005 (**Table 56**).

**Table 55**: Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in doxorubicin-containing regimens stratify analysis[1]

| *Taxonomy* | RA, median (%) | Pre (%) | Doxorubicin-containing regimens: No (n=122) OR (95%CI) | Doxorubicin-containing regimens: Yes (n=179) OR (95%CI) | *P for interaction* | *$P_{FDR}$ for interaction*[2] |
|---|---|---|---|---|---|---|
| Phylum *Firmicutes A* | | | | | | |
| Family *Clostridiaceae* | 0.037 | 97.0 | 0.65 (0.37-1.14) | 1.66 (1.14-2.41) | 0.001 | 0.097 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥Median relative abundance for rare taxa was calculated among carriers
[1] Model was adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa

**Table 56:** Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in docetaxel-containing regimens stratify analysis[1]

| *Taxonomy* | RA, median (%) | Pre (%) | Docetaxel-containing regimens: No (n=225) OR (95%CI) | Docetaxel-containing regimens: Yes (n=76) OR (95%CI) | *P for interaction* | *$P_{FDR}$ for interaction*[2] |
|---|---|---|---|---|---|---|
| Phylum *Proteobacteria* | | | | | | |
| Class *Gammaproteobacteria* | | | | | | |
| Family *Aeromonadaceae* ┼ | 0.006¥ | 14.0 | 0.85 (0.58-1.23) | 7.36 (1.09-49.86) | 0.005 | 0.074 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa

In general, we found no significant interactions between individual chemotherapeutic agents, including cyclophosphamide, 5-FU, overall anthracycline, epirubicin, overall taxane, paclitaxel, and GI microbial taxa on grade ≥3 hematological and GI toxicities.

**Interactions between multiple chemotherapeutic agents and GI microbial taxa on grade ≥3 chemotherapy-induced toxicities**

**Table 57** shows significant interactions between the sequential anthracycline and taxane treatment and seven taxa on grade ≥3 combined hematological toxicity. In the phylum *Actinobacteriota*, two *MGYG-HGUT* species (i.e., *03024* and *009110*), members of the class *Coriobacteriia*, were significantly associated with a high risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.57 (1.12-2.21); $P_{for\ interaction}$=1.08x10$^{-4}$, and 2.15 (1.48-3.11); $P_{for\ interaction}$=1.49 x10$^{-4}$, respectively. Likewise, the family of *Pseudomonadaceae*, a member of the class *Gammaproteobacteria*, had a significantly high risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.60 (1.10-2.33); $P_{for\ interaction}$=0.001. Conversely, in the class *Clostridia*, two genera, including *MGYG-HGUT-02711* (belonging to the family *Acutalibacteraceae*) and *MGYG-HGUT-03297* (belonging to the family *Oscillospiraceae*), showed an inverse association with grade ≥3 combined hematological toxicity, with OR and 95% CIs of 0.55 (0.38-0.79); $P_{for\ interaction}$=0.001, and 0.68 (0.48-0.97); $P_{for\ interaction}$=5.51 x10$^{-5}$, respectively. The significant association for the genus *MGYG-HGUT-03297* was driven by its species *MGYG-HGUT-03297* (OR=0.70, 95% CI: (0.49-0.99); $P_{for\ interaction}$=1.50 x10$^{-4}$). Last but at least, the species *Prevotella intermedia* within the class *Bacteroidia* (belonging to phylum *Bacteroidota)* was associated with a reduced risk of grade ≥3 combined hematological toxicity in the administration of sequential anthracycline and taxane treatment, with OR and 95% CIs of 0.62 (0.41-0.93); $P_{for\ interaction}$=1.70 x10$^{-4}$ (**Table 57**).

Significant interactions between the sequential anthracycline and taxane treatment and the two taxa, the genus *MGYG-HGUT-02711* and the family *Pseudomonadaceae* on grade ≥3 neutropenia, were observed, with OR and 95% CIs of 0.48 (0.32-0.72); $P_{for\ interaction}$=1.54 x10$^{-4}$, and 1.52 (1.07-2.15); $P_{for\ interaction}$=0.004, respectively. In addition, the family *Gemellaceae* within the class *Bacilli (*belonging to phylum *Firmicutes)* had a significantly higher risk of grade ≥3 neutropenia among patients who received

sequential anthracycline and taxane treatment, with OR and 95% CIs of 1.55 (1.10-2.20); $P_{for\ interaction}$=0.004 (**Table 58**).

In the phylum *Firmicutes A,* the genus MGYG-HGUT-03227 within the order *Christensenellales* was associated with an increased risk of grade ≥3 combined GI toxicity in the patients who received sequential anthracycline and taxane treatment, with OR and 95% CIs of 1.88 (1.16-3.05); $P_{for\ interaction}=1.01 \times 10^{-4}$, which driven by the species *MGYG-HGUT-03227* with OR of 1.82 (95%CI: 1.13-2.92; $P_{for\ interaction}=1.31 \times 10^{-4}$). Meanwhile, the genus *MGYG-HGUT-02758*, a member of the order *Monoglobales (*belonging to phylum *Firmicutes A),* was associated with a reduced risk of grade ≥3 combined GI toxicity, with OR and 95% CIs of 0.50 (0.27-0.94); $P_{for\ interaction}=8.49 \times 10^{-5.}$ (**Table 59**). We found no significant interactions between sequential anthracycline and taxane treatment and GI microbial taxa on grade ≥3 nausea/vomiting.

**Table 57:** Association of grade ≥3 combined hematological toxicity with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1]

| Taxonomy | RA, median (%) | Pre (%) | Sequential anthracycline and taxane | | *P for interaction* | $P_{FDR}$ *for interaction*[2] |
|---|---|---|---|---|---|---|
| | | | No (n=89) | Yes (n=212) | | |
| | | | OR (95%CI) | OR (95%CI) | | |
| Phylum *Actinobacteriota* | | | | | | |
| Class *Coriobacteriia* | | | | | | |
| Species *MGYG-HGUT-03024*┼ | 0.005¥ | 15.6 | 0.19 (0.06-0.57) | 1.57 (1.12-2.21) | $1.08x10^{-4}$ | 0.051 |
| Species *MGYG-HGUT-00911*┼ | 0.002¥ | 35.2 | 0.43 (0.21-0.88) | 2.15 (1.48-3.11) | $1.49 x10^{-4}$ | 0.051 |
| Phylum *Bacteroidota* | | | | | | |
| Class *Bacteroidia* | | | | | | |
| Species *Prevotella intermedia*┼ | 0.002¥ | 11.0 | 2.71 (1.28-5.74) | 0.62 (0.41-0.93) | $1.70 x10^{-4}$ | 0.051 |
| Phylum *Firmicutes A* | | | | | | |
| Class *Clostridia* | | | | | | |
| Family *Acutalibacteraceae* | | | | | | |
| Genus *MGYG-HGUT-02711*┼ | 0.008¥ | 40.2 | 2.01 (1.03-3.94) | 0.55 (0.38-0.79) | 0.001 | 0.097 |
| Family *Oscillospiraceae* | | | | | | |
| Genus *MGYG-HGUT-03297*┼ | 0.002¥ | 48.8 | 2.94 (1.47-5.89) | 0.68 (0.48-0.97) | $5.51 x10^{-5}$ | 0.019 |
| Species *MGYG-HGUT-03297*┼ | 0.002¥ | 48.8 | 2.62 (1.35-5.08) | 0.70 (0.49-0.99) | $1.50 x10^{-4}$ | 0.051 |
| Phylum *Proteobacteria* | | | | | | |
| Class *Gammaproteobacteria* | | | | | | |
| Family *Pseudomonadaceae*┼ | 0.004¥ | 16.6 | 0.32 (0.13-0.80) | 1.60 (1.10-2.33) | 0.001 | 0.040 |

Common taxa (prevalence≥50% in the population); ┼ Rare taxa: 10% ≤prevalence< 50% in the population. ¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 58**: Association of grade ≥3 neutropenia with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1]

| Taxonomy | RA, median (%) | Pre (%) | Sequential anthracycline and taxane | | *P for interaction* | $P_{FDR}$ *for interaction*[2] |
|---|---|---|---|---|---|---|
| | | | No (n=89) | Yes (n=212) | | |
| | | | OR (95%CI) | OR (95%CI) | | |
| Phylum *Firmicutes_A* | | | | | | |
| Class *Clostridia* | | | | | | |
| Family *Acutalibacteraceae* | | | | | | |
| Genus *MGYG-HGUT-02711* ┼ | 0.008¥ | 40.2 | 2.50 (1.19-5.25) | 0.48 (0.32-0.72) | 1.54 x10$^{-4}$ | 0.054 |
| Phylum *Firmicutes* | | | | | | |
| Class *Bacilli* | | | | | | |
| Family *Gemellaceae* ┼ | 0.003¥ | 49.2 | 0.34 (0.14-0.80) | 1.55 (1.10-2.20) | 0.004 | 0.065 |
| Phylum *Proteobacteria* | | | | | | |
| Class *Gammaproteobacteria* | | | | | | |
| Family *Pseudomonadaceae* ┼ | 0.004¥ | 16.6 | 0.19 (0.05-0.71) | 1.52 (1.07-2.15) | 0.004 | 0.065 |

Common taxa (prevalence≥50% in the population); ┼ Rare taxa: 10% ≤prevalence< 50% in the population. ¥ Median relative abundance for rare taxa was calculated among carriers
[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.
ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa
Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Table 59**: Association of grade ≥3 combined GI toxicity with GI microbiome taxa in the sequential anthracycline and taxane- stratify analysis[1]

| | | | Sequential anthracycline and taxane | | | |
|---|---|---|---|---|---|---|
| | | | No (n=89) | Yes (n=212) | *P for* | $P_{FDR}$ *for* |
| Taxonomy | RA, median (%) | Pre (%) | OR (95%CI) | OR (95%CI) | *interaction* | *interaction*[2] |
| Phylum *Firmicutes A* | | | | | | |
| Class *Clostridia* | | | | | | |
| Order *Christensenellales* | | | | | | |
| Genus *MGYG-HGUT-03227*┼ | 0.002¥ | 34.9 | 0.24 (0.04-1.34) | 1.88 (1.16-3.05) | $1.01 \times 10^{-4}$ | 0.018 |
| Species *MGYG-HGUT-03227*┼ | 0.002¥ | 34.9 | 0.29 (0.07-1.27) | 1.82 (1.13-2.92) | $1.31 \times 10^{-4}$ | 0.092 |
| Order *Monoglobales* | | | | | | |
| Genus *MGYG-HGUT-02758*┼ | 0.002¥ | 36.5 | 5.53 (1.55-19.81) | 0.50 (0.27-0.94) | $8.49 \times 10^{-5}$ | 0.018 |

Common taxa (prevalence≥50% in the population); ┼Rare taxa: 10% ≤prevalence< 50% in the population. ¥Median relative abundance for rare taxa was calculated among carriers

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] FDR were calculated at each taxonomic level by common and rare taxa. $P_{FDR}$ <0.1 is considered statistically significant.

ORs and 95% CIs per SD increase in clr-transformed absolute abundance of taxa

Abbreviation: RA: Relative abundance; FDR: false discovery rates; Pre: Prevalence.

**Interactions between chemotherapeutic agents and GI microbial metabolic pathways on grade ≥3 chemotherapy-induced toxicities**

**Table 60-61** shows significant interactions between doxorubicin- and docetaxel-containing regimens and GI microbial metabolic pathways on grade ≥3 combined hematological toxicity. Among those patients, three of eight metabolic pathways were significantly associated with a high risk of grade ≥3 combined hematological toxicity among patients who received doxorubicin-containing regimens. Three pathways that are related to degradation function, including the superpathway of allantoin degradation in yeast and sucrose degradation III (sucrose invertase), were associated with a significantly high risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.91 (1.30-2.82) for ALLANTOINDEG-PWY ($P_{for\ interaction}$=0.18); 1.67 (1.12-2.47) for PWY-621 ($P_{for\ interaction}$=0.001), respectively. Furthermore, the pathway (PWY-5751) that is related to phenylethanol biosynthesis was associated with a significantly high risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.56 (1.07-2.29), $P_{for\ interaction}$=0.018 (**Table 60**). In addition, we found that the pathway (PWY3O-1109) that is related to the superpathway of 4-hydroxybenzoate biosynthesis (yeast) was associated with an increased risk of grade ≥3 combined hematological toxicity (OR=2.69, 95% Cis: 1.03-7.02; $P_{for\ interaction}$=0.049), whereas the pathway (PWY-7200) that is related the Superpathway of pyrimidine deoxyribonucleoside salvage was associated with significantly reduced risk of grade ≥3 combined hematological toxicity (OR=9.41, 95% Cis: 0.18-0.94; $P_{for\ interaction}$=0.042) among patients who received docetaxel-containing regimens (**Table 61**).

**Tables 62-64** show GI microbial metabolic pathways significantly interacted with the sequential anthracycline and taxane treatment on grade ≥3 combined hematological toxicity, neutropenia, and combined GI toxicity. Eleven of 21 metabolic pathways were significantly associated with a high risk of grade ≥3 combined hematological toxicity (**Table 62**). In addition, the pathways that are related to degradation function were significantly associated with an increased risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.41 (1.02-1.95); $P_{for\ interaction}$=0.009 for the superpathway of hexuronide and hexuronate degradation (GALACT-GLUCUROCAT-PWY), 1.59 (1.13-2.25); $P_{for\ interaction}$=0.003 for including the superpathway of β-D-glucuronide and D-glucuronate degradation

(GLUCUROCAT-PWY), 1.41 (1.01-1.97); $P_{for\ interaction}$=0.005 for the superpathway of glucose and xylose degradation (PWY-6901), 1.72 (1.21-2.45), $P_{for\ interaction}$=0.001 for D-fructuronate degradation (PWY-7242); 1.54 (1.09-2.16), $P_{for\ interaction}$=0.003 for 4-deoxy-L-threo-hex-4-enopyranuronate degradation (PWY-6507), and 1.70 (1.21-2.40), $P_{for\ interaction}$=0.04 for the superpathway of allantoin degradation in yeast (ALLANTOINDEG-PWY), respectively. In addition, four pathways that are related to biosynthesis function, including the superpathway of L-tyrosine biosynthesis (PWY-6630), the superpathway of thiamin diphosphate biosynthesis II (PWY-6895), Phenylethanol biosynthesis (PWY-5751) and Taxadiene biosynthesis (engineered; PWY-7392), were associated with a significantly high risk of grade ≥3 combined hematological toxicity, with OR and 95% CIs of 1.46 (1.04-2.05), $P_{for\ interaction}$=0.02; 1.57 (1.11-2.21), $P_{for\ interaction}$=0.04; 1.44 (1.03-2.02), $P_{for\ interaction}$=0.005; and *1.46 (1.05-2.02),* $P_{for\ interaction}$=0.03, respectively.

Significant interactions between the sequential anthracycline and taxane treatment and six metabolic pathways, including GALACT-GLUCUROCAT-PWY, GLUCUROCAT-PWY, PWY-6507, PWY-6630, and PWY-6895, were observed for grade ≥3 neutropenia. The superpathway of β-D-glucuronide and D-glucuronate degradation (GLUCUROCAT-PWY) and D-fructuronate degradation (PWY-7242) were significantly associated with an increased risk of grade ≥3 neutropenia, with OR and 95% CIs of 1.63 (1.14-2.33), $P_{for\ interaction}=4.29 \times 10^{-4}$ and 1.77 (1.22-2.55), $P_{for\ interaction}=3.72 \times 10^{-4}$ (**Table 63**).

Furthermore, the pathway (P162-PWY) that is related to L-glutamate degradation V (via hydroxyglutarate) was associated with a significantly high risk of grade ≥3 combined GI toxicity, with OR and 95% CIs of 1.77 (1.11-2.81), $P_{for\ interaction}$=0.03 (**Table 64**).

**Table 60:** Association of grade ≥3 combined hematological toxicity with GI microbial metabolic pathways in doxorubicin-containing regimens stratify analysis

| Pathways | Function | Doxorubicin-containing regimens[1] No (n=122) OR (95%CI) | Doxorubicin-containing regimens[1] Yes (n=179) OR (95%CI) | P for interaction[2] |
|---|---|---|---|---|
| ALLANTOINDEG-PWY | Superpathway of allantoin degradation in yeast | 0.85 (0.51-1.40) | 1.91 (1.30-2.82) | 0.018 |
| COBALSYN-PWY | Adenosylcobalamin salvage from cobinamide I | 0.52 (0.31-0.85) | 0.82 (0.57-1.16) | 0.020 |
| PWY-4722 | Creatinine degradation II | 0.37 (0.17-0.80) | 1.03 (0.73-1.46) | 0.004 |
| PWY-5751 | Phenylethanol biosynthesis | 0.83 (0.50-1.38) | 1.56 (1.07-2.29) | 0.018 |
| PWY-621 | Sucrose degradation III (sucrose invertase) | 0.49 (0.28-0.87) | 1.67 (1.12-2.47) | 0.001 |
| PWY-6470 | Peptidoglycan biosynthesis V (β-lactam resistance) | 0.55 (0.32-0.93) | 0.72 (0.50-1.03) | 0.046 |
| PWY-821 | Superpathway of sulfur amino acid biosynthesis (Saccharomyces cerevisiae) | 0.59 (0.36-0.96) | 1.34 (0.94-1.90) | 0.027 |
| SER-GLYSYN-PWY | Superpathway of L-serine and glycine biosynthesis I | 0.53 (0.32-0.91) | 1.27 (0.89-1.80) | 0.014 |

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] All $P_{FDR\ for\ interaction}$ >0.1
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 61:** Association of grade ≥3 combined hematological toxicity with GI microbial metabolic pathways in docetaxel-containing regimens stratify analysis

| Pathways | Function | Docetaxel-containing regimens[1] No (n=225) OR (95%CI) | Docetaxel-containing regimens[1] Yes (n=76) OR (95%CI) | P for interaction[2] |
|---|---|---|---|---|
| DENITRIFICATION-PWY | Nitrate reduction I (denitrification) | 1.42 (1.03-1.95) | 0.94 (0.46-1.96) | 0.030 |
| PWY3O-1109 | Superpathway of 4-hydroxybenzoate biosynthesis (yeast) | 0.94 (0.70-1.27) | 2.69 (1.03-7.02) | 0.049 |
| PWY-7200 | Superpathway of pyrimidine deoxyribonucleoside salvage | 1.29 (0.94-1.77) | 0.41 (0.18-0.94) | 0.042 |
| URSIN-PWY | Ureide biosynthesis | 1.43 (1.02-1.99) | 0.66 (0.24-1.79) | 0.028 |

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] All $P_{FDR\ for\ interaction}$ >0.1
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 62:** Association of grade ≥3 combined hematological toxicity with GI metabolic pathways in the sequential anthracycline and taxane- stratify analysis

| | | Sequential anthracycline and taxane[1] | | |
|---|---|---|---|---|
| | | No (n=89) | Yes (n=212) | |
| Pathways | **Function** | OR (95%CI) | OR (95%CI) | P for interaction[2] |
| ALLANTOINDEG-PWY | Superpathway of allantoin degradation in yeast | 1.06 (0.57-1.96) | 1.70 (1.21-2.40) | 0.039 |
| ASPASN-PWY | Superpathway of L-aspartate and L-asparagine biosynthesis | 2.08 (1.10-3.92) | 0.99 (0.71-1.37) | 0.011 |
| COBALSYN-PWY | Adenosylcobalamin salvage from cobinamide I | 0.35 (0.18-0.69) | 0.82 (0.58-1.15) | 0.018 |
| GALACT-GLUCUROCAT-PWY | Superpathway of hexuronide and hexuronate degradation | 0.53 (0.26-1.09) | 1.41 (1.02-1.95) | 0.009 |
| GLUCUROCAT-PWY | Superpathway of β-D-glucuronide and D-glucuronate degradation | 0.50 (0.23-1.10) | 1.59 (1.13-2.25) | 0.003 |
| GLYCOGENSYNTH-PWY | Glycogen biosynthesis I (from ADP-D-Glucose) | 0.45 (0.22-0.93) | 0.97 (0.71-1.34) | 0.017 |
| PWY-1269 | CMP-3-deoxy-D-manno-octulosonate biosynthesis I | 7.86 (2.35-26.32) | 1.02 (0.73-1.44) | 0.001 |
| PWY-241 | C4 photosynthetic carbon assimilation cycle, NADP-ME type | 0.78 (0.43-1.41) | 1.67 (1.18-2.37) | 0.021 |
| PWY-5154 | L-arginine biosynthesis III (via N-acetyl-L-citrulline) | 1.99 (1.08-3.69) | 0.73 (0.52-1.02) | 0.005 |
| PWY-5751 | Phenylethanol biosynthesis | 0.80 (0.35-1.80) | 1.44 (1.03-2.02) | 0.005 |
| PWY-6470 | Peptidoglycan biosynthesis V (β-lactam resistance) | 0.39 (0.18-0.86) | 0.74 (0.54-1.02) | 0.011 |
| PWY-6507 | 4-deoxy-L-threo-hex-4-enopyranuronate degradation | 0.42 (0.21-0.87) | 1.54 (1.09-2.16) | 0.003 |
| PWY-6630 | Superpathway of L-tyrosine biosynthesis | 0.71 (0.39-1.28) | 1.46 (1.04-2.05) | 0.019 |
| PWY-6892 | Thiazole biosynthesis I (E. coli) | 0.32 (0.15-0.70) | 0.94 (0.68-1.29) | 0.021 |
| PWY-6895 | Superpathway of thiamin diphosphate biosynthesis II | 0.70 (0.41-1.21) | 1.57 (1.11-2.21) | 0.038 |
| PWY-6901 | Superpathway of glucose and xylose degradation | 0.47 (0.25-0.91) | 1.41 (1.01-1.97) | 0.005 |
| PWY-7242 | D-fructuronate degradation | 0.55 (0.28-1.08) | 1.72 (1.21-2.45) | 0.001 |
| PWY-7392 | Taxadiene biosynthesis (engineered) | 0.84 (0.46-1.54) | 1.46 (1.05-2.02) | 0.033 |
| PWY-7539 | 6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia) | 1.91 (1.01-3.61) | 0.93 (0.66-1.31) | 0.049 |
| PWY0-1586 | Peptidoglycan maturation (meso-diaminopimelate containing) | 2.48 (1.17-5.24) | 1.04 (0.76-1.43) | 0.010 |
| PWY66-422 | D-galactose degradation V (Leloir pathway) | 0.51 (0.27-0.99) | 0.92 (0.67-1.27) | 0.033 |

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] All $P_{FDR\ for\ interaction}$ >0.1

ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 63:** Association of grade ≥3 neutropenia with GI microbial metabolic pathways in the sequential anthracycline and taxane- stratify analysis

| Pathways | Function | Sequential anthracycline and taxane[1] | | P for interaction[2] |
|---|---|---|---|---|
| | | No (n=89) | Yes (n=212) | |
| | | OR (95%CI) | OR (95%CI) | |
| COBALSYN-PWY | Adenosylcobalamin salvage from cobinamide I | 0.39 (0.18-0.84) | 0.87 (0.62-1.23) | 0.006 |
| DAPLYSINESYN-PWY | L-lysine biosynthesis I | 0.74 (0.39-1.41) | 1.62 (1.13-2.32) | 0.022 |
| ENTBACSYN-PWY | Enterobactin biosynthesis | 0.79 (0.40-1.57) | 1.48 (1.05-2.09) | 0.025 |
| GALACT-GLUCUROCAT-PWY | Superpathway of hexuronide and hexuronate degradation | 0.58 (0.29-1.18) | 1.44 (1.03-2.03) | 0.006 |
| GLCMANNANAUT-PWY | Superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation | 0.48 (0.24-0.96) | 1.21 (0.87-1.67) | 0.015 |
| GLUCUROCAT-PWY | Superpathway of β-D-glucuronide and D-glucuronate degradation | 0.40 (0.16-1.00) | 1.63 (1.14-2.33) | $4.29 \times 10^{-4}$ |
| P4-PWY | Superpathway of L-lysine, L-threonine and L-methionine biosynthesis I | 0.74 (0.39-1.43) | 1.62 (1.14-2.31) | 0.021 |
| PWY-1269 | CMP-3-deoxy-D-manno-octulosonate biosynthesis I | 6.60 (1.95-22.34) | 1.29 (0.89-1.86) | 0.031 |
| PWY-241 | C4 photosynthetic carbon assimilation cycle, NADP-ME type | 0.63 (0.32-1.24) | 1.77 (1.22-2.57) | 0.005 |
| PWY-5345 | Superpathway of L-methionine biosynthesis (by sulfhydrylation) | 0.53 (0.28-1.03) | 1.49 (1.03-2.14) | 0.005 |
| PWY-6507 | 4-deoxy-L-threo-hex-4-enopyranuronate degradation | 0.53 (0.26-1.07) | 1.58 (1.11-2.26) | 0.004 |
| PWY-6628 | Superpathway of L-phenylalanine biosynthesis | 0.63 (0.32-1.24) | 1.50 (1.04-2.14) | 0.035 |
| PWY-6629 | Superpathway of L-tryptophan biosynthesis | 0.51 (0.25-1.04) | 1.58 (1.10-2.27) | 0.014 |
| PWY-6630 | Superpathway of L-tyrosine biosynthesis | 0.55 (0.27-1.11) | 1.48 (1.04-2.13) | 0.016 |
| PWY-6895 | Superpathway of thiamin diphosphate biosynthesis II | 0.66 (0.36-1.20) | 1.60 (1.10-2.33) | 0.036 |
| PWY-7094 | Fatty acid salvage | 0.77 (0.39-1.54) | 1.49 (1.06-2.10) | 0.043 |
| PWY-7242 | D-fructuronate degradation | 0.46 (0.20-1.02) | 1.77 (1.22-2.55) | $3.72 \times 10^{-4}$ |
| PWY-7290 | Escherichia coli serotype O86 O-antigen biosynthesis | 0.55 (0.20-1.54) | 1.44 (1.03-2.03) | 0.042 |
| PWY0-1479 | tRNA processing | 0.61 (0.31-1.23) | 1.46 (1.03-2.06) | 0.029 |
| PWY0-1586 | Peptidoglycan maturation (meso-diaminopimelate containing) | 2.81 (1.25-6.33) | 1.02 (0.73-1.42) | 0.013 |
| PWY0-781 | Aspartate superpathway | 0.76 (0.39-1.46) | 1.56 (1.10-2.22) | 0.031 |

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).

[2] All $P_{FDR\ for\ interaction}$ >0.1

ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

**Table 64:** Association of grade ≥3 combined GI toxicity with GI microbial metabolic pathways in the sequential anthracycline and taxane- stratify analysis

| Pathways | **Function** | **Sequential anthracycline and taxane**[1] | | P for interaction[2] |
|---|---|---|---|---|
| | | No (n=89) | Yes (n=212) | |
| | | OR (95%CI) | OR (95%CI) | |
| PWY-7198 | Pyrimidine deoxyribonucleotides de novo biosynthesis IV | 0.29 (0.09-0.93) | 1.41 (0.86-2.32) | 0.019 |
| PWY-7196 | Superpathway of pyrimidine ribonucleosides salvage | 3.76 (1.28-11.03) | 0.93 (0.59-1.48) | 0.025 |
| P162-PWY | L-glutamate degradation V (via hydroxyglutarate) | 0.19 (0.02-1.49) | 1.77 (1.11-2.81) | 0.026 |

[1] Models were adjusted for stool collection time (after breast cancer surgery vs. before breast cancer surgery & neoadjuvant chemotherapy), age group, income levels, residence, menopausal status, BMI levels, comorbidity, diagnosis delay, breast cancer subtypes, TNM cancer stage, dose-dense chemotherapy, fiber intake, fat intake, carbohydrate intake, and physical activity (MET/wk).
[2] All $P_{FDR\ for\ interaction}$ >0.1
ORs and 95% CIs per SD increase in asr-transformed relative abundance of pathways

## 3. Discussion

This study explored drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among 301 newly diagnosed Vietnamese breast cancer patients. In association with severe combined hematological toxicity, our study found the family *Clostridiaceae* had a significant interaction with doxorubicin-containing regimens, while the family *Aeromonadaceae* had a significant interaction with docetaxel-containing regimens, both reached the FDR correction threshold of significance ($P_{FDR} < 0.1$). In addition, significant interactions between administration of sequential anthracycline and taxane and bacteria on severe hematological and GI toxicities, particularly neutropenia, were found for some taxa, including families *Pseudomonadaceae*, *Acutalibacteraceae* and *Oscillospiraceae,* which might be due to GI microbiota involvement in the taxadiene biosynthesis (engineered; PWY-7392) and three degradation functions, including the superpathway of β-D-glucuronide and D-glucuronate degradation (GLUCUROCAT-PWY), 4-deoxy-L-threo-hex-4-enopyranuronate degradation (PWY-6507), and D-fructuronate degradation (PWY-7242). These microbial taxa and metabolic pathways were associated with a significantly high risk of severe hematological toxicity among patients who received a sequential anthracycline and taxane treatment. Our findings suggest interactions between the gut microbiome and chemotherapeutic agents in developing chemotherapy-induced toxicity among breast cancer patients.

The interaction between gut microbes and non-antibiotic drugs such as anticancer drugs is complex, and bidirectional.[251] As described previously, many preclinical and clinical studies have attempted to elucidate the gut microbiota changes caused by chemotherapy. The richness and composition of the GI microbiota during and after chemotherapy treatment exhibited marked alternations in both diversity and composition.[157] On the other hand, the gut microbiome can also affect an individual's response to a drug by enzymatically transforming its structure and altering its bioavailability, bioactivity, or toxicity.[252] To our knowledge, this is the first study to date evaluating drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among breast cancer patients. A recent preclinical study supported the hypothesis that paclitaxel may impair barrier integrity by reducing the number and function of beneficial gut bacteria (e.g., *Akkermansia muciniphila*), resulting in systemic exposure to bacterial metabolites and products – which could result in altered brain function via the gut-immune-brain axis.[204]

Our study identified effect modifications by individual/multiple drugs-microbiome interactions on the association between GI microbial taxa/metabolic pathways and chemotherapy-induced toxicity via stratified analysis. The sample size of 301 participants might not be adequately powered to detect modest drug-microbiome interaction. Some true associations between GI microbiome and chemotherapy-induced toxicity may have been missed and associations might be underestimated or overestimated due to the small sample size, particularly when stratified by stool collection time. This may explain the largely null results that we found regarding the interactions between individual chemotherapeutic agents, including cyclophosphamide, 5-FU, anthracycline, epirubicin, taxane, paclitaxel, and GI microbial taxa on severe chemotherapy-induced toxicities in our study. Among all the individual drug-microbiome interactions evaluated, we only found that two families, including *Clostridiaceae* and *Aeromonadaceae,* had significant interactions with doxorubicin- or docetaxel-containing regimens on severe combined hematological toxicity after FDR correction. However, it is worth noting that breast cancer patients often receive multiple chemotherapeutic agents during chemotherapy. In our study, approximately 70% of breast cancer patients received sequential anthracycline and taxane regimens, with paclitaxel being the predominant taxane used, such as AC-P/FAC-P/ EC-P/FEC-P. In addition, using 5-FU in a combination with anthracycline was prescribed for approximately 21.0% of patients. Thus, it is difficult to evaluate specific individual drug-microbiome interaction in our study population.

We found that sequential anthracycline and taxane treatment modified the association between the family *Pseudomonadaceae* (belonging to the class *Gammaproteobacteria* within the phylum *Proteobacteria)* with severe combined hematological toxicity and neutropenia. This bacteria family was associated with a significantly increased risk of severe hematological toxicities in the presence of sequential anthracycline and taxane treatment but was inversely associated with the risk of severe combined hematological toxicity in the absence of sequential anthracycline and taxane treatment. These effect modifications could be linked to the significant interaction that we found between severe hematological toxicities and some GI microbial metabolic pathways. For example, *Pseudomonas aeruginosa*, a member of the family *Pseudomonadaceae*, has been known to be involved in the pathways[253,254] including the superpathway of β-D-glucuronide and D-glucuronate degradation (GLUCUROCAT-PWY), 4-deoxy-L-threo-hex-4-enopyranuronate degradation (PWY-6507), and

D-fructuronate degradation (PWY-7242), was significantly associated with an increased risk of severe hematological toxicities among patients who received sequential anthracycline and taxane treatment.

Moreover, we found a significant interaction between the GI microbial taxadiene biosynthesis and severe combined hematological toxicity in the presence of sequential anthracycline and taxane treatment. Taxadiene (Taxa-4,11-diene) is the first dedicated intermediate in taxol (paclitaxel) biosynthesis. The enzymes catalyzing the steps of the Taxadiene biosynthesis pathway have been assembled from some bacteria belonging to the class *Gammaproteobacteria,* such as *Escherichia coli K-12 substr. MG1655* and *Pseudomonas herbicola.*[254] In our study, docetaxel-containing regimens had effect modification on the association of the family *Aeromonadacea,* another member of the class *Gammaproteobacteria,* with severe combined hematological toxicity. We speculate that members of both families, *Pseudomonadaceae and Aeromonadacea,* might be involved in the taxadiene biosynthesis. This hypothesis needs to be further investigated. Our findings suggested that drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity may partially explain our observation in Aim 1 that breast cancer patients who received chemotherapy with sequential anthracycline and taxane were more likely to experience higher incidences of severe hematological toxicities. In addition, the drug-microbiome interaction may explain that patients receiving docetaxel-containing regimens (e.g., the regimens of TAC or TC) or dose-dense chemotherapy (dose-dense AC followed by docetaxel) had a high incidence of neutropenia[87,88] as well as might partially address racial disparity on chemotherapy-induced hematological toxicity. Asians experienced a higher incidence and severity rate of hematological toxicity than Caucasians when given G-CSF use was consistent. [221] Further studies are needed to understand the impact of drug-microbiome interaction.

Furthermore, we observed *Prevotella intermedia* within the class *Bacteroidia* (belonging to the phylum *Bacteroidota*), and some taxa, members of families *Acutalibacteraceae* and *Oscillospiraceae* (belonging to the class *Clostridia*) were inversely associated with severe combined hematological toxicity in patients who received sequential anthracycline and taxane treatment. Unfortunately, we found no linked pathways from the analysis for GI microbial metabolic pathways. However, if confirmed, these findings may offer a potential

microbial intervention to prevent developing hematological toxicity among patients who received sequential anthracycline and taxane treatment.

The main limitation of the present study is the limited sample size for subgroup/stratified analyses. However, our study is thus far the largest on this topic. On the other hand, stool samples were collected only once before chemotherapy, making it difficult to evaluate the dynamics of gut microbiota caused by chemotherapeutic agents and their relationship to the development of chemotherapy-induced toxicities.

In conclusion, our results provide early evidence supporting the hypothesis that interaction between the GI microbiome and multiple chemotherapeutic agents may contribute to developing or preventing chemotherapy-induced toxicity among breast cancer patients. However, further larger studies are needed to validate our findings and understand how the microbiome metabolizes agents or ameliorates the efficacy of cancer treatment.

## OVERALL CONCLUSION

We conducted a case-only study with four specific aims based on a prospective follow-up of 501 newly diagnosed Vietnamese breast cancer patients recruited into the Vietnamese Breast Cancer Study (VBCS) to investigate the influence of the pre-chemotherapy GI microbiota on chemotherapy-induced toxicity among breast cancer patients.

**Specific aim 1**: We described the incidence of chemotherapy-induced toxicity and evaluated the associations between the chemotherapy-induced toxicity and clinical and demographic factors. Among 396 Vietnamese breast cancer patients who received chemotherapy, we found a substantial proportion of severe hematological (38.6%) and GI (12.9%) toxicities associated with the administration of chemotherapeutic agents at the first-line treatment. Neutropenia and nausea/vomiting were the most common hematological toxicities and GI toxicities among breast cancer patients, with 29.5% and 10.1% of patients experiencing grade ≥ 3. Participants, particularly those who received chemotherapy with a sequential anthracycline and taxane treatment, were more likely to experience severe chemotherapy-induced hematological and GI toxicities. In addition, a pre-existing nephrological condition and dose-dense chemotherapy were significantly associated with an increased risk of severe hematological toxicities, particularly neutropenia. Moreover, the triple-negative/basal-like subtype was significantly associated with high risks of severe hematological and GI toxicities compared with other breast cancer subtypes.

**Specific aim 2**: We evaluated the associations of GI microbial richness and composition and individual microbial taxa with non-clinical and clinical factors among 356 Vietnamese breast cancer patients who donated fecal samples at baseline before systemic treatment regardless of their status receiving breast cancer surgery. The GI microbiome profile of breast cancer patients differ significantly before and after breast cancer surgery which was always followed by prophylaxis antibiotic treatment. We found a significantly lower GI microbial richness (measured by alpha diversity indexes), different composition (measured by beta diversity) and significant alteration in taxa abundance of approximately 40% of investigated 2,864 GI microbial taxa among breast cancer patients who underwent breast cancer surgery followed by prophylaxis antibiotic treatment. In addition, breast cancer patients who experienced a serious delay in diagnosis and treatment, particularly

patients with stool samples collected before surgery, had lower GI microbial richness and composition and reduced proportions of carriers of species *Roseburia hominis* and *Firmicutes* populations, including 48 genera and six families within the class *Clostridia*. Furthermore, our study showed that age, income, and geographic residence influence the GI microbiome of breast cancer patients, with considerably declined abundances or proportions of carriers of many microbial taxa among elderly patients, patients having low income, or those living in rural areas. Finally, we found that some GI microbial taxa were significantly associated with cancer stages, breast cancer subtypes, ER and HER2 status, and BMI levels.

**Specific aim 3**: In this study among 301 Vietnamese breast cancer patients, who donated stool samples at baseline and received neoadjuvant or adjuvant chemotherapy, we evaluated the associations of pre-chemotherapy GI microbial richness and composition, individual microbial taxa, and metabolic pathways with chemotherapy-induced toxicities, including hematological and GI toxicities at the first-line chemotherapy.

We found high pre-chemotherapy GI microbial richness (Chao1 and Shannon indexes) and high abundances of specific taxa (most of them from the families *Lachnospiraceae, Oscillospiraceae,* and *Ruminococcaceae* such as *Coprococcus eutactus, Dorea scindens, Eubacterium E hallii A, Eubacterium G ventriosum, Intestinimonas butyriciproducens, Faecalibacterium prausnitzii J,* and *Ruminococcus D bicirculans*) were significantly associated with a reduced risk of severe hematological toxicities among breast cancer among patients with stool samples collected pre-surgery. Conversely, the enrichment of specific microbial taxa from the families *Bacteroidaceae*, *Lachnospiraceae*, *Sporanaerobacteraceae*, and *Fusobacteriaceae* was significantly associated with an increased risk of severe neutropenia. Moreover, we found that 157 species, 41 genera, and one family belonging to the phylum *Firmicutes A,* most of which may be less susceptible to breast cancer surgery and antibiotic exposure, were associated with an increased risk of severe nausea/vomiting among breast cancer patients with stool sample collected after surgery. In addition, we found gut microbial functional capacity in relation to the reductive TCA cycle I, phosphatidylglycerol biosynthesis I and II, and superpathway of phospholipid biosynthesis I was significantly associated with an increased risk of severe GI toxicities, particularly nausea/vomiting, associations that did not differ by time of stool sample collection. Furthermore, GI microbiota involved in the superpathway of polyamine biosynthesis I and phosphatidate metabolism were significantly associated with an increased risk of severe hematological

toxicities, whereas those involved in the pathways methanogenesis from acetate, pyrimidine deoxyribonucleotides de novo biosynthesis and peptidoglycan biosynthesis II, IV and V (β-lactam resistance) was significantly associated with a reduced risk of severe hematological toxicities. In addition, a high abundance of GI microbiota involved in lactose and galactose and L-proline biosynthesis II (from arginine) were associated with a reduced risk of severe combined GI toxicities and nausea/vomiting. Finally, our findings suggest that overall GI microbial richness and multiple microbes may influence the development of hematological and GI toxicities among breast cancer patients. Diet interventions, prebiotics, or probiotics for restoring normal gut microbiota after breast cancer surgery may be considered in clinical care and practice for Vietnamese breast cancer patients who routinely receive prophylactic antibiotics in surgery before chemotherapy.

**Specific aim 4**: We explored drug-microbiome interaction on the association between pre-chemotherapy GI microbiome and chemotherapy-induced toxicity among 301 newly diagnosed Vietnamese breast cancer patients. Significant interactions with the administration of sequential anthracycline and taxane treatment on severe hematological and GI toxicities, particularly neutropenia, were found for some taxa, including *Pseudomonadaceae Acutalibacteraceae, Oscillospiraceae,* which may be linked with their functions in the taxadiene biosynthesis and involvement in the superpathway of β-D-glucuronide and D-glucuronate degradation, 4-deoxy-L-threo-hex-4-enopyranuronate degradation, and D-fructuronate degradation. These findings provide novel information on the possible biological mechanisms underlying the interactions between the gut microbiome and chemotherapeutic agents in developing chemotherapy-induced toxicity among breast cancer patients.

Our study provides evidence of the role of the pre-chemotherapy GI microbiota on the development of chemotherapy-induced toxicities and suggests restoring normal gut microbiota as a potential preventive measure to reduce chemotherapy toxicity among breast cancer patients. Further larger studies are needed to validate our findings and understand how the pre-chemotherapy GI microbiome and the changes of GI microbiota during, and post-chemotherapy contribute to developing short-term and long-term chemotherapy toxicity as well as treatment efficacy among breast cancer patients. Research on the latter is particularly important as some chemotherapy toxicity associated with microbes might also be related to treatment efficacy.

In addition, studies are needed to investigate whether diet interventions, prebiotics, and/or probiotics supplementation prior to chemotherapy or during/post-chemotherapy would reduce short-term chemotherapy toxicities and/or long-term chemotherapy toxicities. Furthermore, besides chemotherapy, targeted therapy and immunotherapy have made noteworthy progress in the clinical treatment of breast cancer. Investigating the role of the GI microbiome in the toxicities of these cancer treatments is an uncharted research territory. Knowledge gain would open a new avenue to increase treatment adherence and enhance efficacy. Future studies are needed to determine the best strategy to manipulate the GI microbiota to improve long-term cancer outcomes and quality of life for breast cancer survivors.

## VIII. APPENDIX

**Supplementary Table S1:** Human studies investigating changes to the diversity of GI microbiota among breast cancer patients in comparison with healthy controls and by clinical features

| Author [ref] | Type of study | Sample size and characteristics | Method | Main finding | Changes to the GI microbiome |
|---|---|---|---|---|---|
| **Case-control study** | | | | | |
| Bertazzoni et al., 2006 [154] | Case-control study | Feces from 18 breast cancer patients and 30 healthy women.<br>Breast cancer patients were four premenopausal women and 14 postmenopausal women. | Gram-stain, morphological, and biochemical analysis | The bacteria flora composition in feces breast cancer patients was different from that of healthy women in both number and species. | Premenopausal breast cancer patients had a significant increase in the number of *Enterobacteriaceae*, aerobic *Streptococci*, *Lactobacilli*, and anaerobic bacteria, including *Clostridia, bacteroides*, and anaerobic *lactobacilli* in feces when compared with healthy controls. In menopausal period, breast cancer patients showed a remarkable increase in *Bacteroides, Clostridia*, and anaerobic *Lactobacilli*.<br>Compared with healthy controls, several different bacterial species were identified in breast cancer patients: *S. constellatus*, S. intermedius; *Ps. saccharolyticus* and *Ps. asaccharolyticus*; *E. rectale*; *L. fermentans*, *L. plantarum*; *C. limosum*, *C. symbiosum*, *C. sordellii*, *C. glycolium*; *B. gengivalis*, and *B. asaccharolyticus*. |
| Goedert et al., 2015 [146] | Population-based case-control study | Urine and feces from 48 postmenopausal breast cancer patients and 48 paired control patients.<br>Patients were 42 ER+, 37 PR+, and 5 HER2-positive.<br>Clinical stages were 11 in situ (stage 0), 25 at stage I, ten at stage II, and two at cancer stage III. | Illumina sequencing and taxonomy16S rRNA genes | Postmenopausal breast cancer patients had an altered fecal microbiota composition (beta-diversity) and an estrogen-independent low alpha-diversity, compared with control patients. | The relative abundance of several taxa was different between breast cancer patients and control patients. Particularly at the family level, cases had higher *Clostridiaceae*, *Faecalibacterium*, and *Ruminococcaceae* and lower *Dorea* and *Lachnospiraceae*. |
| Goedert et al., 2018 [147] | Population-based case-control study | Urine and feces from 48 postmenopausal breast cancer patients (75% stage 0-1, 88% estrogen-receptor positive) and 48 paired controls.<br>The original sample study was published.[146] | 16S V4 rRNA gene amplicon sequencing, the DADA2 package, and SILVA (ver123). | Breast cancer patients had a significant estrogen-independent association with the IgA-positive and IgA-negative gut microbiota. | Case and controls differed significantly in the composition of the IgA-positive microbiota and the IgA-negative microbiota fractions. Cases were more likely than controls to carry IgA-coated Betaproteobacteria *Parasutterella*, particularly carry IgA-coated Betaproteobacteria *Parasutterella excrementihominis*. Conversely, cases were less likely than control to carry eight taxa, including IgA-coated *Firmicutes Clostridiales Ruminococcaceae Oscillibacter*, IgA-noncoated Bacteroidetes *Alistipes indistinctus*, and six IgA-noncoated *Firmicutes Clostridiales* taxa.<br>Breast cancer patients showed a lower richness (number of observed species) and α-diversity (Chao 1 index), which was significantly more marked in the IgA-positive than the IgA-negative microbiota. |

| | | | | | |
|---|---|---|---|---|---|
| Zhu et al., 2018 [155] | Case-control study | Feces from premenopausal breast cancer patients (n=18), premenopausal healthy women (n=25), postmenopausal breast cancer patients (n=44), and postmenopausal healthy women (n=46) | Illumina sequencing (HiSeq x10 platform) | The gut microbiota composition differed significantly between postmenopausal breast cancer patients and healthy controls, while it was similar between premenopausal breast cancer patients and healthy controls. | The number of species, mean Chao1 index, and beta diversity were higher in postmenopausal breast cancer patients than in postmenopausal controls.<br>Bacterial species were enriched in postmenopausal breast cancer patients, including *Escherichia coli, Citrobacter koseri, Acinetobacter radioresistens, Enterococcus gallinarum, Shewanella putrefaciens, Erwinia amylovora, Actinomyces sp. HPA0247, Salmonella enterica, and Fusobacterium nucleatum.*<br>*Eubacterium eligens* and *Lactobacillus vaginalis* were less abundant in postmenopausal breast cancer patients compared with controls. |
| Miko et al., 2018 [148] | Population-based case-control study | Serum samples and feces were collected from 48 healthy women and 48 postmenopausal breast cancer patients.<br>The original sample study was published.[146] | qPCR assays were used to measure the abundance of the baiH DNA in fecal DNA samples | The abundance of baiH ORF in several bacterial species was significantly lower in breast cancer patients. | Compared with healthy controls, the abundance of baiH of *Clostridium sordelli*, *Staphylococcus haemolyticus*, *Escherichia coli*, and *Pseudomonas putida* was significantly lower in breast cancer patients, and in line with the lower LCA level and LCA/CDCA ratio.<br>Early-stage breast cancer patients (stage 0-I) showed a more pronounced decrease in the abundance of baiH of *Bacteroides thetaiotaomicron*, *Clostrium sordellie*, *Staphylococcus haemolyticus*, *Escherichia coli,* and *Pseudomonas putida.* |
| Kovacs et al., 2019 [149] | Population-based case-control study | Serum samples and feces were collected from 48 healthy women and 48 postmenopausal breast cancer patients.<br>The original sample study was published.[146] | qPCR assays were used to measure the abundance of DNA coding for LdcC and CadA in fecal DNA samples. | The abundance of DNA coding for LdcC and CadA in several bacterial species was significantly decreased in breast cancer patients. | Compared with healthy individuals, the abundance of *Escherichia coli* CadA and E.*coli*, *Enterobacter cloacae,* and *Hafnia alvei* LdcC DNA were slightly declined in breast cancer patients. The early-stage breast cancer patients (Stage 0) showed a more pronounced decrease in CadA and LdcC abundance when compared with the pool of all patients.<br>In the feces of stage I patients, *Escherichia coli* LdcC protein levels were markedly lower than in the feces of healthy controls, in line with the lower fecal DNA abundance. |
| He et al., 2021[255] | Case-control study | Feces were collected from 54 premenopausal breast cancer patients and 28 premenopausal healthy controls. | 16S V3-V4 rRNA sequences and targeted metabolomics | The composition and symbiosis of gut microbiota in patients with premenopausal breast cancer changed significantly compared with that in premenopausal healthy controls. | Compared with premenopausal healthy controls, the intestinal bacteria, and their interrelationships in premenopausal women with breast cancer changed significantly, with a reduced abundance of short-chain fatty acids (SCFA)-producing bacteria and significantly lower levels of intestinal SCFA-producing enzymes.<br>*Pediococcus* and *Desulfovibrio* could distinguish premenopausal breast cancer patients from premenopausal healthy women. |
| Hou et al., 2021 [156] | Population-based Case-control study | Feces were collected from 67 age-matched female controls (50 premenopausal women and 17 postmenopausal women) and 200 breast cancer patients (100 | 16S V3-V4 rRNA sequences | The gut microbiota in premenopausal patients differed from that in postmenopausal breast cancer patients. In addition, | The alpha diversity was significantly reduced in premenopausal breast cancer patients, and the beta diversity differed significantly between breast cancer patients and controls.<br>The 14 microbial makers were identified in the different menopausal status of breast cancer. Premenopausal breast cancer patients had significantly higher *Anaerostipes* and *Bacteroides fragilis*, whereas postmenopausal breast cancer patients had significantly higher |

| | | | | | |
|---|---|---|---|---|---|
| | | premenopausal patients and 100 post-menopausal patients) with stage I-II | | the functional pathways differed between breast cancer patients and controls. | *Proteobacteria* and *Klebsiella pneumoniae.* The above four bacterial taxa were not affected by age.<br>Compared with premenopausal controls, premenopausal breast cancer was enriched with the pathways contributing to the abundance of the microbiome against the steroid-related and oncogenic-related pathways. Moreover, postmenopausal breast cancer patients exhibited greater enrichment of steroid-related and chemical carcinogenesis pathways than premenopausal breast cancer patients. |
| Yang & Wang et al., 2021 [151] | Case-control study | Feces were collected from 83 patients with invasive breast cancer and 19 patients with benign breast tumors.<br>Malignant patients were 30 postmenopausal, 51 ER+, 47 PR+, 37 HER2-positive, and 62 Ki67 ≥30%<br>Clinical stages were three at stage I, 29 at stage II, and 48 at cancer stage III. | 16S V4 rRNA sequences | Patients with malignant breast tumors had a distinct enrichment of gut microbiome by different clinicopathological factors, including ER, PR, Ki-67 levels, HER2 status, and tumor grade | Microbiome community richness was higher in the benign group than in the malignant group. The metabolic pathways in patients with malignant breast tumors, especially the lipopolysaccharide biosynthesis pathways, was significantly different from those in patients with a benign tumor.<br>Patients with malignant tumors possessed elevated levels of *Citrobacter*, whereas a great majority of the microbiota elevated in those with benign tumors included *Clostridium*, *Faecalibacterium*, *Lachnospira*, *Erysipelotrichaceae*, *Romboutsia*, *Fusicatenibacter*, *Xylophilus*, and *Arcanobacterium.* |
| **<u>Case-only study</u>** | | | | | |
| Luu et al., 2017 [150] | Case-only study | Feces from 31 patients with early-stage breast cancer.<br>Over 90% patients with ER+/PR+ and 15% of patients overexpressed HER2. Clinical stages were stage 0 (n=15), stage I (n=7), stage II (n=7) and Stage III (n=2). Histological grading was 16.1% grade 1, 61.3% grade 2 and 22.6% grade 3.<br>Twenty-three patients had a normal BMI, and eight were overweight or obese (BMI > 25 kg/m$^2$). | Real-time qPCR targeting 16S rRNA sequences | Intestinal microbiota composition in breast cancer patients differs according to clinical characteristics and BMI. | The total number of Bacteroidetes, *C. leptum cluster*, *C. coccoides cluster*, *F. prausnitzii, and Blautia sp.* were significantly higher in Clinical stage II/III breast cancer than in clinical-stage 0/I.<br>*Blautia sp* was significantly associated with more severe histological grades.<br>In overweight and obese patients, the number of total Firmicutes, *F. prausnitzii*, *Blautia sp.,* and *E. lenta* bacteria was significantly lower than that found in patients with normal BMI. |
| Fruge et al., 2018 [153] | Clinical trial | Feces from 32 overweight or obese women diagnosed with early-stage breast | 16S V4 rRNA sequences | Body composition was inversely associated with AM, | Women with a high relative abundance of AM (HAM patients) had lower fat mass when compared with low AM relative abundance (LAM patients). |

| | | | | | |
|---|---|---|---|---|---|
| | | cancer (stage 0 to II) in a presurgical weight-loss trial.<br>Participants were dichotomized per the median relative abundance of *Akkermansia muciniphila* (AM) at baseline. | | microbiome diversity, and positively interleukin-6 level in early-stage breast cancer patients. | Alpha-diversity measures (Chao1 and Shannon index) were higher in women with HAM at baseline and attenuated after weight loss.<br>Higher *Prevotella* and *Lactobacillus* and lower *Clostridium*, *Campylobacter*, and *Helicobacter* genera in HAM patients vs. LAM patients.<br>At baseline, the interleukin-6 level was associated with species richness and fat mass, but not AM. |
| Wu et al., 2020 [152] | Case-only study | Fecal samples were collected prior to chemotherapy of 37 incident breast cancer patients with mostly Hispanic women, 25 HER2-negative and 12 HER2-positive | 16S V4 rRNA sequences | HER2 status and age at menarche had significant associations with gut microbiome alpha diversity measures and specific microbial composition. | Compared with HER2-negative breast cancer patients, HER2-positive patients showed 12-23% lower alpha diversity, lower abundance of *Firmicutes* and higher abundance of *Bacteroidetes*.<br>Women with early menarche (age ≤11 age) was associated with lower OUT, Chao 1 index, and lower abundance of *Firmicutes* when compared with women with later menarche (age ≥12 age). |

* bai: bile acid-inducible operon (wherein the baiH ORF codes for 7-HSDH, a key enzyme in lithocholic acid biosynthesis); LCA: lithocholic acid; CDCA: chenodeoxycholic acid; LdcC: constitutive lysine decarboxylase; CadA: acid-inducible lysine decarboxylase.

**Supplementary Table S2:** Chemotherapeutic treatment and GI microbiota changes in clinical studies

| Author [ref] | Subjects | Treatment | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| Nyhlén et al., 2002 [[158]] | Feces from 9 patients with acute leukemia | Different combination of 9 intravenously administered antineoplastic drugs | Standard microbiological culture techniques | An increase in the count of *Bacteroides* spp. in 3 of 9 patients during treatment and an increased count of yeasts in 2 of 5 patients during chemotherapy-induced neutropenia. | During chemotherapy, no significant changes in the numbers of bacteria or *Candida* spp.: *Enterococci*, *Streptococci*, *Staphylococci*, *bacilli*, and *Escherichia coli*). The number of *Bacteroides* spp. significantly increased in 3/9 patients, whereas *Lactobacilli*, *Bifidobacterial*, *Peptostreptococci,* and *Clostridia* spp. were stable during treatment. The Gram-positive aerobic microflora was unchanged in all five patients during neutropenia, but the number of yeasts increased in 2 of 5 patients. |
| Van Vliet et al., 2009 [[159]] | Feces from 9 pediatric patients with acute myeloid leukemia (AML) | Four consecutive chemotherapy cycles (ADE I and ADE II: high-dose cytarabine, daunorubicin, and etoposide; MACE: amsacrine, high-dose cytarabine, and etoposide; MidAC: mitoxantrone, and high-dose cytarabine) with antibiotic prophylaxis | PCR-Denaturing gradient gel electrophoresis (DGGE) fingerprinting and fluorescent in situ hybridization (FISH) | A tremendous decrease in intestinal microbial diversity and a disturbed balance between aerobic and anaerobic bacteria were observed in patients with AML during treatment. | The total number of bacteria significantly decreased during treatment (comparing healthy control samples) and was restored at six weeks after the last chemotherapy cycles. There were 70-20,000-fold decrease in the number of anaerobic bacteria: *Bacteroides* species, *Clostridium* cluster XIVa, *Faecalibacterium prausnitzii*, and *Bifidobacterium* species. At the end-of-treatment, the total *Clostridium* XIVa and *F. prausnitzii* recovered, while both *Bacteroides* and *Bifidobacterium* species were 10-300-fold lower (compared with healthy control samples). In patients with ALM, *Enterococci* significantly increased, but *streptococci* diminished, and no gram-negative Enterobacteriaceae were detected during treatment. In addition, the number of streptococci was restored at six weeks after the last chemotherapy cycle. |
| Zwielehner et al., 2011 [[169]] | Fecal samples from 17 ambulant patients and 17 gender-, age-, and lifestyle-matched healthy controls | Different chemotherapeutic treatments with or without the present concomitant antibiotics. | TaqMan qPCR, denaturing gradient gel electrophoresis (DGGE) fingerprint, and 454 high-throughput sequencings. | Chemotherapeutic treatment-induced changes in fecal microbiota with a coincidental development of *C. difficile* and a decrease in *Bifidobacterium, Lactobacillus*, *Veillonella,* and *Faecalibacterium prausnitzii*. | The absolute number of fecal bacteria among cancer patients was significantly lower than the healthy control. Decreases in bacterial abundances following chemotherapy were mainly related to decreases in *Bacteroides*, *Bifidobacteria*, *Clostridium* cluster IV, and *Clostridium* cluster XIVa. Patients who received antibiotics had higher bacterial abundances than those without concomitant antibiotics. The incidence of *C. difficile* in two patients immediately after chemotherapy was accompanied by a reduction of *Bifidobacterium*, *Lactobacillus*, *Veillonella, Escherichia coli /Shigella,* and particularly *Faecalibacterium prausnitzii* (from 9% to undetected). |
| Montassier et al., 2014 [[163]] | Pre- and post-chemotherapy fecal samples of 8 adult patients with non-Hodgkin lymphoma | Bone marrow transplantation conditioning chemotherapy for five consecutive days including high-dose | 16S V5-V6 rRNA gene amplicon sequencing and 454 high-throughput pyrosequencing | The fecal microbiota of patients exhibited a steep reduction in alpha diversity and significant differences in the composition of | Overall diversity decreased significantly in evenness measured by the Shannon diversity index and richness measured by phylogenetic diversity following a cycle of chemotherapy. At the phylum level, the fecal microbiota of patients after chemotherapy showed a decrease in *Firmicutes* and an increase in *Bacteroidetes*. |

| Author [ref] | Subjects | Treatment | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| | | Carmustine, Etoposide, Aracytine, and Melphalan | | the intestinal microbiota in response to chemotherapy | At the genus level, fecal microbiota of patients showed a profound decrease in *Blautia*, *Faecalibacterium* and *Roseburia*, and *Bifidobacterium*, and increases in *Bacteroides* and *Escherichia* after chemotherapy compared with before chemotherapy. |
| Montassier et al., 2015 [164] | Feces from 28 Patients with Non-Hodgkin's lymphoma | Myeloablative condition regimens for five consecutive days including high-dose Carmustine, Etoposide, Aracytine, and Melphalan without concomitant antibiotics | 16S V5-V6 rRNA gene amplicon sequencing | The fecal microbiota of patients exhibited a rapid and marked decreased overall diversity and a distinct disruption in bacteria composition. | After chemotherapy, a reduction in overall diversity with a decreased evenness (alpha diversity) and richness (Faith's phylogenetic diversity).<br>At the phylum level, the fecal microbiota of patients after chemotherapy showed significant declines in the abundance of *Firmicutes* and *Actinobacteria* and an increase in *Proteobacteria*.<br>At the genus level, fecal samples after chemotherapy exhibited decreased abundances in *Ruminococcus, Oscillospira*, *Blautia*, *Lachnospira*, *Roseburia*, *Dorea*, *Coprococus*, *Anaerostipes*, *Clostridium*, *Collinsella*, *Adlercreutzia,* and *Bifidobacterium*, whereas showed increased abundances in *Citrobacter*, *Klebsiella*, *Enterococcus*, *Megasphaera,* and *Parabacteroides*.<br>*Actinomyces*, *Mobiluncus*, *Scardovia*, *Slackia*, *Prevotella*, *Mitsuokella*, *Oxalobacter,* and *Erysipelotrichaceae* were unchanged before and after chemotherapy, which was considered to be resistant to chemotherapy. |
| Galloway-Peña et al., 2016 [160] | Pre- and post-chemotherapy (every 96h until neutrophil recovery) buccal and fecal specimens were collected twice weekly from 34 patients with newly diagnosed acute myelogenous leukemia (AML) | Different induction chemotherapy with antimicrobial prophylaxis | 16S V4 rRNA sequencing using an Illumina MiSeq system | A statistically significant proportion of the patients who had alpha diversity decreases in both the oral and stool samples over the course of the induction chemotherapy. | There was an overall statistically significant decrease in microbial diversity over the course of induction chemotherapy in both the oral and stool samples. However, the loss in diversity was not common, and some patients gained diversity during chemotherapy.<br>They observed statistically significant increases for *Lactobacillus* in both oral and stool samples, and significant decreases were primarily observed for anaerobic genera, such as *Blautia, Prevotella*, and *Leptotrichia.*<br>Patients who had received a carbapenem antibiotic for over 72 hours were significantly more likely to have a decrease in both oral and stool alpha diversity over the course of induction chemotherapy compared with those who did not receive a carbapenem antibiotics. |
| Kong, Gao and Yan et al., 2017 [165] | Fecal samples were collected preoperatively, postoperatively, and after the first to fifth cycles of postoperative chemotherapy from 43 CRC patients who received | Multiple cycles of capecitabine and plus oxaliplatin (CapeOx) therapy | High throughput 16S rRNA amplicon sequencing | The CapeOx regimen was found to alter intestinal microbiota dramatically. | CapeOx therapy reduces the abundance of cancer-promoting bacteria (*Enterococcus*) and certain pathogenic bacteria (*Escherichia-Shigella, Morganella, Pyramidobacter, and Proteus*). It also noted that CapeOx therapy led to an increased abundance of many conditionally pathogenic bacteria such as *Collinsella, Anaerostipes, Bilophila, Comamonas, Weissella, Bacteroides,* and *Eggerthella.* The ratio of *Bacteroidetes* to *Firmicutes* was also found an increase after chemotherapy<br>The lactate-utilizing and butyrate-producing bacteria *Butyricimonas* and *Butyricicoccus* increased while the conditional pathogens of |

| Author [ref] | Subjects | Treatment | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| | radical surgery and adjuvant chemotherapy | | | | *Veillonella* decreased after CapeOx therapy. After chemotherapy, the lactate-utilizing microbiota shifted from *Veillonella* to *Butyricimonas* and *Butyricicoccus*.<br>The "rebound effect" of chemotherapy-adapted bacteria was observed. The abundance of *Dorea, Ruminococcaceae_UCG-010, Streptococcus, Prevotella_9, Mogibacterium,* and *Roseburia* fluctuated after one or two cycles of chemotherapy. However, after five chemotherapeutic cycles, their abundance recovered. |
| Dang et al., 2018 [166] | Fecal samples were collected from 69 individuals divided into four groups: healthy individuals (n=33), CRC patients before treatment (n=17), chemotherapy-treated CRC patients (n=14), and surgically treated CRC patients (n=5) | 6-8 cycles of the chemotherapeutic cocktail of oxaliplatin and tegafur (a precursor of 5′-FU) | 16S V4-V5 rRNA sequencing | Some microbial groups were tightly associated with CRC patients undergoing chemotherapy | *Fusobacterium nucleatum* was shown to confer chemoresistance during CRC therapy, and certain bacterial strains or genera, such as the genus *Sutterella* and species *Veillonella dispar*, were associated with CRC patients who were treated with chemotherapeutic cocktails.<br>Two species, *Prevotella copri* and *Bacteroides plebeius*, were only enriched in patients treated with chemotherapy. |
| Galloway-Peña et al., 2020 [162] | Oral swabs and stool samples were obtained biweekly from baseline until neutrophil recovery following induction chemotherapy from 97 patients with acute myelogenous leukemia (AML) | Different induction chemotherapy with antimicrobial prophylaxis | 16S V4 rRNA sequencing using an Illumina MiSeq system | Oral and stool microbiome changed over the course of chemotherapy treatment | There was a significant decrease in Shannon diversity over the course of IC for both oral and stool samples.<br>*Clostridiales* and *Blautia* were significantly higher in baseline vs the end-of-study stool samples. Similarly, several taxa were significantly higher among oral baseline samples, including *Viellonellaceae*, *Prevotellaceae*, and *Gemella*, whereas *Staphylococcus* was enriched at the end of the study. |
| Tong et al., 2020 [168] | Fecal samples were collected preoperatively, postoperatively, and after the first to fifth cycles of chemotherapy from 18 ovarian cancer patients | 6 cycles of regimen of TC (Carboplatin and paclitaxel) and TP chemotherapy (Cisplatin and paclitaxel) | Genomic DNA extraction, PCR amplification and 16S rRNA sequencing | Chemotherapy may have a differential influence on gut microbiota. | Compared with before chemotherapy, the abundance of *Bacteroidetes* and *Firmicute*s increased, and the abundance of *Proteobacteria* decreased after chemotherapy.<br>The abundance of *anaerobic* bacteria, such as *Bacteroides, Collinsella,* and *Blautia*, exhibited an increasing tendency after multiple cycles of chemotherapy, whereas the abundance of *Veillonella, Lachnospiraceae_unclassified, Roseburia, Akkermansia* and *Bifidobacterium* increased at the first to third cycles of chemotherapy and decreased at the subsequence cycles of chemotherapy. |

| Author [ref] | Subjects | Treatment | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| Shuwen et al., 2020 [167] | Stool samples of 15 stage II-III postoperative and 11 advanced CRC patients who completed 8 cycles of the XELOX regimen and 11 CRC patients who completed 8 cycles of the FOLFIRI regimen | 8 cycles of the XELOX (Oxaliplatin + Capecitabine) regimen or FOLFIRI (Irinotecan + leucovorin + 5-fluorouracil) regimen | 16S ribosomal RNA gene and ITS ribosomal RNA gene sequencing | The community structure of gut bacteria and fungi changes in chemotherapy on CRCs | The abundances of *Veillonella, Humicola, Tremellomycetes,* and *Malassezia* were increased in post-operative CRC patients treated with the XELOX regimen.<br>The abundances of *Faecalibacterium, Clostridiales, phascolarctobacterium, Humicola* and *Rhodotorula* were decreased, and the abundances of *Candida, Magnusiomyces, Tremellomycetes, Dipodascaceae, Saccharomycetales, Malassezia* and *Lentinula* were increased in advanced CRC patients treated with the FOLFIRI regimen.<br>The abundance of *Humicola, Rhodotorula*, and *Magnusiomyces* was decreased, and the abundances of *Candida, Tremellomycetes, Dipodascaceae, Saccharomycetales, Malassezia,* and *Lentinula* were increased in advanced CRC patients treated with the FOLFIRI regimen combined with cetuximab compared with those treated with the FOLFIRI regimen alone. |

**Supplementary Table S3:** Chemotherapeutic treatment and GI microbiota changes in pre-clinical studies.

| Author [ref] | Subjects | Administration | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| **5-Fluorouracil (5-FU)** | | | | | |
| Bültzingslöwen et al., 2003 [170] | Samples from the small intestine and large intestine of female Lewis rats | 5-FU (injected) | Bacteria culture techniques | 5-FU treatment caused an increase in the number of Gram-negative anaerobes in the large intestine and an increased translocation to mesenteric lymph nodes | The total number of anaerobic bacteria and facultative anaerobes in the small intestine was unchanged. There was a predominated shift in the types of facultative (i.e., from gram-positive cocci to gram-negative rods). No increase in the number of anaerobes was observed in the large intestine, whereas the gram-negative facultative increased and the gram-positive facultative decreased, leading to an increase in the total number of facultative anaerobes in the colon. |
| Stringer et al.,2009 [171] | Fecal samples and colon samples from female Dark Agouti rats | 5-FU (injected) | Standard microbiological culture techniques and Real-time PCR | 5-FU treatment resulted in significant changes to intestinal flora in rats. | Microbiological culture methods showed increases in *Clostridium spp.* (after 24h) and *Escherichia spp.* (from 2-6h and again from 48-72h) and decreases in *Lactobacillus spp.* (from 1-2h and 12-24h) in colon. In feces, there were decreases in Clostridium spp. (2h following treatment), *Escherichia* spp. (from 2-24h after treatment), Proteus spp. (from 2-6h after treatment) and *Streptococcus* spp. (at 72h). Real-time PCR showed fluctuation in *Bifidobacterium spp.* between time points, decreases in *Bacteroides spp.* (at 48h), *Lactobacillus spp.* (from 12-24h) and *Enterococcus spp.* (from 2-48h), whereas exhibited increases in *Clostridium spp.* (after treatment), *Escherichia coli* (at 48h) and *Staphylococcus spp.* (at 24h). *E. faecalis*, *S. pneumoniae*, *L. acidophilus*, *B. lactis*, *C. botulinum,* and *S. epidermidis* also showed susceptibility to 5-FU. |
| Hamouda et al., 2017 [172] | Fecal samples from male C57BL/6J mice | 5-FU (50mg/kg injected)/ 5-FU with the administration of ampicillin | 16S V3-V4 rRNA gene amplicon sequencing and qPCR | 5-FU treatment induced a decrease in the abundance of intestinal *Firmicutes*, but an increase in *Bacteroidetes* and *Verrucomicrobia*. | Repeated administration of 5-FU decreased the abundance of *Firmicutes* while increased the abundance of *Bacteroidetes* and *Verrucomicrobia*. However, these responses were completely blocked by co-administered ampicillin, which increased the abundance of *Firmicutes* and decreased the abundance of *Bacteroidetes* and *Verrucomicrobia.* |
| Li et al., 2017 [173] | Feces and cecum contents from male BALB/c mice | 5-FU (50mg/kg intraperitoneal injections)/ 5-FU with fecal transplantation | 16S V3-V4 rRNA gene amplicon sequencing and qPCR | 5-FU treatment greatly diminished the community richness and diversity and altered the abundance of microbiota in both feces and cecum contents. | In cecum contents and feces, administration of 5-FU significantly decreased *Firmicutes*/*Bacteroides* ratio. 5-FU treatment reduced the phylum *Firmicutes*, *Proteobacteria* and *Cyanobacteria, Ordoribacter, Candidatus Saccharimonas,* and *Marvinbryantia* but increased *Helicobacter* and *Thalassopira* in feces. 5-FU treatment increased the abundance of *Verrucomicrobia* and significantly changed the abundance of *Blautia*, *Alistipes*, *Coprococcus*, *Roseburia*, *Akkermansia*, *Bilophila*, *Candidatus Saccharimonas*, and *Mucispirilum* in cecum contents. |
| Vanlancker et al. 2017 [175] | Mucosal and luminal samples from an *in vitro* | 5-FU (a dose of 10 μM) and irinotecan (SN-38; a dose of 10 μM) | Clustering analysis of denaturing gradient gel electrophoresis | 5-FU and SN-38 only displayed a minor impact on colon microbial functionality and | There were no clear shifts in the microbial profile of all donors, and only a minor difference between control and 5-FU treated samples or SN-38 treated samples could be observed. |

| Author [ref] | Subjects | Administration | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| | mucosal simulator of the human intestinal microbial system | | (DGGE) and 16S V3-V4 rRNA gene amplicon sequencing | composition in the luminal and mucosal gut microbiota. | At the genus level, 5-FU treatment increased the relative abundance of *Bacteroides* and decreased the abundance of *Escherichia*/*Shigella* in the lumen of few donors. 5-FU increased the microbial diversity and the abundance of some bacteria (e.g., *Anaeroglobus*, *Roseburia*, and *Parabacteroides*) in the mucus but did not influence the diversity in the lumen at the end of treatment.<br>SN-38 did not greatly impact the microbial diversity indices and specific genera for all donors. However, some donor-specific changes could be observed: increases in *Cloacibacillus* and *Alistipes* in the lumen and an increase in *Roseburia* in the mucus. |
| Yuan et al., 2018 [[174]] | Feces from colorectal cancer female BALB/c mice | 5-FU/ 5-FU combined with an antibiotic cocktail/ 5-FU with probiotic | 16S V3-V4 rRNA gene amplicon sequencing | 5-FU treatment reduced the overall alpha diversity and altered microbial composition. | 5-FU treatment significantly decreased *Actinobacteria, Alistipes*, *Lactobacillus* and increased the relative abundance of *Enterobacter*, *Lachnospiraceae_Nk4 A136_group*, *Escherichia-Shigella*, *Alloprevotella*, *Bacteroides*, *Rikenella*, *Blautia, Mucispirillum,* and *Mycoplasma*. 5-FU treatment induced profound losses of the species: *lactobacillus_animalis* and *Helicobacter_hepaticus* and dominance of *Lachnospiraceae_bacterium_COE1*, *Bacteroides_vulgatus*, *Mycoplasma_sualvi*, and *Escherichia coli*. |
| **Cyclophosphamide** | | | | | |
| Viaud et al., 2013 [[176]] | Samples of the small intestine from mice in bearing subcutaneous cancers | CTX | 454 pyrosequencing technology and 16S V1-3 rRNA gene amplicon sequencing | CTX induced changes in microbiota composition of microbiota in small intestine and led to the translocation of Gram+ bacterial species into secondary lymphoid organs. | One week after treatment, CTX led to a reduction of species of the Firmicutes phylum distributed within four genera and groups (*Clostridium* cluster XIVa, *Roseburia*, *unclassified Lachnospiraceae*, *Corprococcus* in the small intestine mucosa of treated mice.<br>The total bacterial load of the small intestinal microbiota and the bacterial counts of the Clostridium leptum group (cluster IV) were unchanged, whereas cyclophosphamide reduced the abundance of Lactobacilli and Enterococci at seven days post-treatment.<br>Several Gram+ bacterial species, including *Lactobacillus johnsonii*, *Lactobacillus murinus,* and *Enterococcus hirae* were cultured from mesenteric lymph nodes and spleens. |
| Yang et al., 2013 [[177]] | Fecal samples and intestine tissues from male Balb/c mice | CTX (Intraperitoneal injection) at 25 mg/kg, 50 mg/kg and 100 mg/kg for 5 days | Standard microbiological culture techniques | Treatment with CTX, especially at high doses, altered mucosal barrier, and colonization resistance, and increased intestinal permeability significantly. | Detecting changes of colonization resistance with predominantly anaerobic resident microflora (*Bifidobacterium* and *Lactobacillus*) and increases the bacterial counts of potentially pathogenic bacteria (*Escherichia coli*, *Enterobacteriaceae*, *Pseudomonas,* and *enterococci*) after treatment.<br>Treatment with CTX at 100 mg/kg induced significantly higher *Enterobacteriaceae*, *Pseudomonas,* and *enterococci* group counts compared with the other treatment.<br>Treatment with CTX at 50 mg/kg and 100 mg/kg led to significant increases in intestinal permeability. |
| Xu et al., 2015 [[178]] | Fecal samples from pathogen-free male C57BL/6 mice | CTX<br>(Intraperitoneal injections) one a week for 28 consecutive days | 454 pyrosequencing technology | CTX administered intraperitoneally reduced bacterial diversity and shifted the fecal microbiota composition. | The bacterial community α-diversity (Shannon & Simpson indices) was lower in the CTX-treated mice group.<br>CTX treatment also decreased the relative abundance of *Bacteroides*, which led to an increase *Firmicutes*/*Bacteroides* ratio. |

| Author [ref] | Subjects | Administration | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| | | | | | Administration of CTX significantly increased the relative abundance of class *Bacilli*, *Clostridia*, C. *Coriobacteriia*, and *Mollicutes,* and the family *Lacnospiraceae, Coriobacteriaceae*, *Lactobacillaceae*, and *Staphylococcaceae*; decreased the class *Bacteroidia* and *AlphaProteobacteria,* and the family *Prevotellaceae*, S24-7, *Alcaligenaceae*, and *Rhodospirillaceae;* disappeared *Verrucomicrobia* and *Streptococcaceae*. |
| **Gemcitabine** | | | | | |
| Panebianco et al., 2018 [179] | Feces from pancreatic cancer xenografted mice | Gemcitabine (100 μ/10g injected) | 16S V3-V4 rRNA gene amplicon sequencing | Gemcitabine induced significant changes in intestinal microbiota with a shift towards an inflammation-related bacterial profile. | Gemcitabine reduced the proportion of *Firmicutes* and *Bacteroidetes* but increased *Protobacteria* and *Verrucomicrobia.* Gemcitabine also decreased the relative abundance of *Bacteroidales* order, *Lachnospiraceae* family, and *Ruminococcaceae*, while almost disappeared the genus of *Erysipelatoclostrium*, *Alistipes,* and *Anaerotruncus.* At the species level, gemcitabine significantly increased *Akkermansia muciniphila*, *Escherichia. coli* and *Peptoclostridium difficile but* decreased *Bacteroides acidifaciens* and *Lactobacillus animalis*. |
| **Irinotecan** | | | | | |
| Stringer et al.,2007 [181] | Fecal samples and colon samples from female Dark Agouti rats | Irinotecan (intraperitoneal injection) | Standard microbiological culture techniques | Microflora changes were observed 6, 12, and 24 hours after treatments, with a relative modification in the presence of bacteria compared with control rats. | In colon, there was an increase in the levels of *Escherichia* spp. (between 6-24 hours), *Clostridium* spp. (at 2 hours), *Enterococcus* spp. (at 6 hours), Serratia spp. (at 60 min – 24 hours), *Staphylococcus* spp. (at 60 min and 48 hours), Bacillus spp. (at 6 hours), *Peptostreptococcus* spp. (at 30-60 min) and *Lactobacillus* spp. (over time). Proteus spp. and Streptococcus spp. were undetected at 2 hours, while Veillonella spp. were not detected at 30 mins and 2-6 hours. In feces, several bacteria were not detected during the follow-up period: *Bacillus* spp (at 12 hours and 72 hours), *Bifidobacterium* spp. (at 60 mins and 48-72 hours), Clostridium spp. (at 60 min), Veillonella spp. (until 12 hours), and Actinobacillus spp. (at all time points). *Prosteus* spp. reached their highest levels at 24-72 hours, while the levels of *Peptostreptococcus* spp, *Clostridium* and *Enterobacter* spp. were highest at 30min to 2 hours. |
| Stringer et al.,2008 [182] | Feces from female Dark Agouti rats | Irinotecan (intraperitoneal injection) | Real-time PCR | Fecal flora changed quantitatively after treatment with increases in the β-glucuronidase-producing bacteria and decreases in the 'beneficial' bacteria | There was an increase in the level of the β-glucuronidase-producing bacteria: *Staphylococcus spp.* (from 2-12h), *Clostridium spp.* (at 48h) and *Escherichia coli* (from 24-48h), whereas the levels of two beneficial bacteria were decreased for *Lactobacillus spp.* (from 12-48h), and *Bifidobacterium spp.* (at all time points). *Bacteroides spp.* (a β-glucuronidase-producing, major component of intestinal flora) decreased from 6-24h and at 72h. |
| Lin et al., 2012 [183] | Cecal samples and feces from tumor-bearing rats | *Dose-intensive regimen*: Irinotecan (injected) | PCR-denaturing Gradient Gel electrophoresis and qPCR | Chemotherapy changed the intestinal microbiota composition, with increases in the abundance of clostridial | For dose-intensive treatment, changes in cecal microbiota were decreases in the total bacteria number and all bacterial groups (except the *Closterium* cluster XI), particularly the *Bacteroides* group and *Clostridium* clusters IV and XIVa on the third-day treatment. By day 7, the number of total bacteria and the Bacteroides group were recovered. |

| Author [ref] | Subjects | Administration | Techniques | Main finding | GI microbiota changes after treatment |
|---|---|---|---|---|---|
| | | *Low-dose regimen*: Irinotecan and 5-FU (injected) | | clusters XI and *Enterobacteriaceae* after treatment. | The abundance of *Clostridium* cluster XI and *Enterobacteriaceae* remained higher than Day 0, whereas *Clostridium* cluster XIVa, *Lactobacillus* group, and *Bifidobacterium* spp. remained significantly lower. Clostridium cluster I was undetected at all time points. For irinotecan and 5-FU therapy, *Clostridium* cluster XI, *Clostridium* cluster XIVa, and *Enterobacteriaceae* increase, while Clostridium cluster IV declined in cecal samples at Day 11. Chemotherapy-induced changes in fecal microbiota were less pronounced than those in cecal microbiota. |
| Forsgård et al., 2017 [184] | Feces from 48 male Sprague-Dawley rats | irinotecan (200 mg/kg)/5-FU (150mg/kg)/ Oxaliplatin (15 mg/kg) | 16S rRNA gene amplicon sequencing | Irinotecan led to an increase in the relative abundance of Fusobacteria and *Proteobacteria*, whereas 5-FU and Oxaliplatin induced a minor change in the fecal microbiota composition. | Administration of irinotecan caused a significant reduction in fecal microbiota diversity. At the end of the experiment, irinotecan treated group showed a significantly decreased relative abundance of *Actinobacteria*, *Bacteroides*, and *Synergistetes* and significant increases in *Fusobacteria* and *Proteobacteria* compared with the control group. 5-FU treatment induced an increase in *Verrucomicrobia*, while Oxaliplatin caused a rise of *Proteobacteria* compared with the control group at the end of the experiment. |
| **Methotrexate** | | | | | |
| Fijlstra et al., 2015 [180] | Feces from male Wistar Unilever outbred rats | MTX (90mg/kg intravenously injections) | Fluorescence in situ hybridization (FISH) | Substantial decreases in the absolute number and diversity of intestinal microbiota and increases in the relative number of enteropathogenic bacteria | MTX induced an overall decrease in most bacteria on day four after treatment. A decrease in the number of anaerobic and aerobic bacteria after MTX treatment: *Clostridium* cluster XIVa, *Ruminococci*, *Eubacterium cylindroides* group, bifidobacterial, mouse intestinal bacteria, C. *ramosum* (on day 4), Bacteroides (on day 6), Lactobacilli and enterococci (on day 2) and Streptococci (on day 4 and 6). There was a reduction in the relative number of almost anaerobes as well as aerobic bacteria after MTX treatment, but the relative number of *Bacteroides*, *Lactobacilli*, *Enterococc*i, and *Enterobacteriaceae* was increased. |

**Supplementary Table S4:** Patient self-report chemotherapy-induced side effects

Have you experienced the following symptoms during cancer chemotherapy in the last six months? [For each symptom, please report the most serious one]

| Symptom | None | Mild | Moderate | Severe<br>[*Notify your doctor or go to hospital immediately if you have]* | How many days has the symptom lasted? |
|---|---|---|---|---|---|
| 1. Nausea | 0 | 1-Can eat | 2-Eating/ drinking less than normal | 3- Cannot eat or drink | __ days |
| 2. Vomiting | 0 | 1-Vomied once during a day | 2-Vomited 2 to 5 times dung the day | 3-Vomited 6 or more times during the day | __ days |
| 3. Diarrhea | 0 | 1-Loose stools | 2-Watery stools, many more than normal | 3- Constant or bloody, or causing you to feel dizzy | __ days |
| 4. Constipation | 0 | 1-No bowel movement for two days | 2-No bowel movement for 3 to 4 days | 3-No bowel movement for more than four days or swollen abdomen | __ days |
| 5. Sore mouth | 0 | 1-Soreness or painless ulcer | 2-Soreness of a painful ulcer but can eat | 3-Painful ulcer and cannot eat | __ days |
| 6. Pain or difficulty with swallowing | 0 | 1-Pain but can eat | 2-pain requiring soft or liquid diet | 3-unable to eat at all | __ days |
| 7. Fatigue (Felling weak) | 0 | 1-Able to do normal activities with some effort | 2-In bed less than half of the day | 3-In bed more than half the day | __ days |
| 8. High fever | 0 | 1-37.0°C to 38°C | 2-38°C to 40°C | 3- Greater than 40°C | __ days |
| 9. Allergic Reaction | 0 | 1-Rash, no fever | 2-Rash, fever < 38°C | 3-Hives, Fever> 38°C, difficulty breathing, seek immediate treatment | __ days |
| 10. Itching or Rash | 0 | 1-Scattered skin rash with redness/mild itching | 2-Generalized rash with sores | 3-Rash with open sores | __ days |
| 11. Cough | 0 | 1-Mild | 2-Dry persistent, controlled with medications | 3-Not controlled with medications | __ days |
| 12. Muscle or Joint pain | 0 | 1-Sore but does not require medicine | 2-Requires medicines for pain | 3- Pain medicine does not help | __ days |
| 13. Tingling or numbness in hands | 0 | 1-Tingling sensation | 2-Tingling, some numbness | 3-Numbness, interfering with function (e.g., can't hold a cup) | __ days |

**Supplementary Table S5:** National Cancer Institute Common Toxicity Criteria

| Category | Toxicity | Grade0 | Grade1 | Grade2 | Grade3 | Grade4 | Normal range |
|---|---|---|---|---|---|---|---|
| Fever/no infection | Fever/no infection | None | 37.1 - 38.0° | 38.1 - 40.0° | >40° for < 24 hours | >40° for > 24 hours or fever accompanied by hypotension | |
| Cardiac | Ischemia | None | Nonspecific T-wave flattening | Asymptomatic, ST/T wave changes suggesting ischemia | Angina without evidence for infarction | Acute myocardial infarction | |
| Blood/ bone marrow | WBC (x1000/mm3) | >4.0 | 3.0-3.9 | 2.0-2.9 | 1.0-1.9 | < 1.0 | |
| Blood/ bone marrow | PLT (x1000/mm3) | WNL | 75.0-WNL | 50.0-74.9 | 25.0-49.9 | < 25.0 | 150-450 G/L |
| Blood/ bone marrow | Hgb (gm/L) | WNL | 100.0-WNL | 80.0-100.0 | 65-79 | < 65 | 125-145g/L |
| Blood/ bone marrow | NEUT/ANC (x1000/mm3) | >2.0 | 1.5-1.9 | 1.0-1.4 | 0.5-0.9 | < 0.5 | Calculate by % NEUT * WBC |
| Blood/ bone marrow | Lymph (x1000/mm3) | >2.0 | 1.5-1.9 | 1.0-1.4 | 0.5-0.9 | < 0.5 | Calculate by % Lymph * WBC |
| Infection | Infection | None | Mid | Moderate | Sever | Life-Threating | |
| Hepatic | Bilirubin | WNL | -- | < 1.5 xN | 1.5-3.0 x N | > 3.0 xN | 0.2 -1.0 mg/dL or 3.4- 17.1 µmol/L |
| Hepatic | SGOT/SPT (AST/ALT) | WNL | <2.5 x N | 2.6-5.0 x N | 5.1 – 20.0 x N | > 20.0 x N | <37 U/L |
| Renal/bladder | Creatinine | WNL | <1.5 x N | 1.5 – 3.0 x N | 3.1-6.0 x N | > 6.0 x N | <37 U/L |
| Renal/bladder | Proteinuria | No change | 1+ or <0.3 gm% (<3 gm/L) | 2-3+ or 0.3-1.0 gm% (3-10 gm/L) | 4+ or > 1.0 gm% (>10 gm/L) | Nephrotic syndrome | |
| Renal/bladder | Hematuria | Neg. | Microscopic | Gross, no clots | Gross + clots | Requires Tx. | |

* 1000/mm$^3$ = 1 G/L; * WNL: with nnormal limits; * N: Normal
WBC: White blood cells count; PLT: platelet count; Hgb: Hemoglobin; NEUT/ANC: Neutrophil.

## REFERENCE


1 Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, doi:10.3322/caac.21660 (2021).
2 Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394-424, doi:10.3322/caac.21492 (2018).
3 Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer Statistics, 2021. *CA Cancer J Clin* **71**, 7-33, doi:10.3322/caac.21654 (2021).
4 Porter, P. "Westernizing" women's risks? Breast cancer in lower-income countries. *N Engl J Med* **358**, 213-216, doi:10.1056/NEJMp0708307 (2008).
5 Vineis, P. & Wild, C. P. Global cancer patterns: causes and prevention. *Lancet* **383**, 549-557, doi:10.1016/S0140-6736(13)62224-2 (2014).
6 Barnes, B. B., Steindorf, K., Hein, R., Flesch-Janys, D. & Chang-Claude, J. Population attributable risk of invasive postmenopausal breast cancer and breast cancer subtypes for modifiable and non-modifiable risk factors. *Cancer Epidemiol* **35**, 345-352, doi:10.1016/j.canep.2010.11.003 (2011).
7 Tamimi, R. M. *et al.* Population Attributable Risk of Modifiable and Nonmodifiable Breast Cancer Risk Factors in Postmenopausal Breast Cancer. *Am J Epidemiol* **184**, 884-893, doi:10.1093/aje/kww145 (2016).
8 Valencia, O. M. *et al.* The Role of Genetic Testing in Patients With Breast Cancer: A Review. *JAMA Surg* **152**, 589-594, doi:10.1001/jamasurg.2017.0552 (2017).
9 Darbre, P. D. & Fernandez, M. F. Environmental oestrogens and breast cancer: long-term low-dose effects of mixtures of various chemical combinations. *J Epidemiol Community Health* **67**, 203-205, doi:10.1136/jech-2012-201362 (2013).
10 Lacey, J. V., Jr. *et al.* Breast cancer epidemiology according to recognized breast cancer risk factors in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Cohort. *BMC Cancer* **9**, 84, doi:10.1186/1471-2407-9-84 (2009).
11 Chlebowski, R. T. Nutrition and physical activity influence on breast cancer incidence and outcome. *Breast* **22 Suppl 2**, S30-37, doi:10.1016/j.breast.2013.07.006 (2013).
12 Nguyen, T. T. & Hoang, M. V. Non-communicable diseases, food and nutrition in Vietnam from 1975 to 2015: the burden and national response. *Asia Pac J Clin Nutr* **27**, 19-28, doi:10.6133/apjcn.032017.13 (2018).
13 Thuan TV, Anh PT, DV, T. & Huong TTT. Cancer control in Vietnam. Where are we? *Cancer control* (2016).
14 Nguyen, S. M. *et al.* Projecting Cancer Incidence for 2025 in the 2 Largest Populated Cities in Vietnam. *Cancer Control* **26**, 1073274819865274, doi:10.1177/1073274819865274 (2019).
15 Parkin, D. M., Whelan, S. L., Ferlay, J., Raymond, L. & Young, J. *Cancer Incidence in Five Continents Vol. VII*. (IARC Press, 1997).
16 Pham, D. X., Ho, T.-Q. H., Bui, T. D., Ho-Pham, L. T. & Nguyen, T. V. Trends in breast cancer incidence in Ho Chi Minh City 1996 - 2015: a registry-based study. doi:10.21203/rs.3.rs-21931/v1 (2020).
17 Parkin, D. M., Whelan, S. L., Ferlay, J., Teppo, L. & Thomas, D. B. *Cancer Incidence in Five Continents Vol. VIII*. (IARC Press, 2002).
18 General Statistics Office & United Nations Population Fund. *Vietnam population projection, 2014-2049*. (Vietnam News Agency Publishing House, 2016).
19 International Agency for Research on Cancer (IARC). *Global Cancer Observatory—Vietnam Population fact sheets*, <http://gco.iarc.fr/today/data/factsheets/populations/704-viet-nam-fact-sheets.pdf> (2020).
20 Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* **127**, 2893-2917, doi:10.1002/ijc.25516 (2010).
21 Jenkins, C. *et al.* Breast cancer services in Vietnam: a scoping review. *Glob Health Action* **11**, 1435344, doi:10.1080/16549716.2018.1435344 (2018).

22 Vu Hong, T., Nguyen Ba, D., Skoog, L., Ta Thanh, V. & Tani, E. Breast Cancer Survival Defined by Biological Receptor and Menopausal Status in Vietnamese Women. *Cancer Control* **26**, 1073274819865279, doi:10.1177/1073274819865279 (2019).
23 Lan, N. H., Laohasiriwong, W. & Stewart, J. F. Survival probability and prognostic factors for breast cancer patients in Vietnam. *Glob Health Action* **6**, 1-9, doi:10.3402/gha.v6i0.18860 (2013).
24 Howlader, N. *et al.* SEER Cancer Statistics Review, 1975-2016, National Cancer Institute., (Bethesda, MD, https://seer.cancer.gov/csr/1975_2016/, based on November 2018 SEER data submission, posted to the SEER web site, April 2019., 2019).
25 Bleyer, A. & Welch, H. G. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med* **367**, 1998-2005, doi:10.1056/NEJMoa1206809 (2012).
26 Nguyen, S. M. *et al.* Delay in the diagnosis and treatment of breast cancer in Vietnam. *Cancer Med* **10**, 7683-7691, doi:10.1002/cam4.4244 (2021).
27 Senkus, E. *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **26 Suppl 5**, v8-30, doi:10.1093/annonc/mdv298 (2015).
28 Berry, D. A. *et al.* Effect of screening and adjuvant therapy on mortality from breast cancer. *N Engl J Med* **353**, 1784-1792, doi:10.1056/NEJMoa050518 (2005).
29 Anampa, J., Makower, D. & Sparano, J. A. Progress in adjuvant chemotherapy for breast cancer: an overview. *BMC Med* **13**, 195, doi:10.1186/s12916-015-0439-8 (2015).
30 Carels, N., Spinasse, L. B., Tilli, T. M. & Tuszynski, J. A. Toward precision medicine of breast cancer. *Theor Biol Med Model* **13**, 7, doi:10.1186/s12976-016-0035-4 (2016).
31 Teven, C. M., Schmid, D. B., Sisco, M., Ward, J. & Howard, M. A. Systemic Therapy for Early-Stage Breast Cancer: What the Plastic Surgeon Should Know. *Eplasty* **17**, e7 (2017).
32 Goldhirsch, A. *et al.* Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* **24**, 2206-2223, doi:10.1093/annonc/mdt303 (2013).
33 Rampurwala, M. M., Rocque, G. B. & Burkard, M. E. Update on adjuvant chemotherapy for early breast cancer. *Breast Cancer (Auckl)* **8**, 125-133, doi:10.4137/BCBCR.S9454 (2014).
34 Masood, S. Neoadjuvant chemotherapy in breast cancers. *Womens Health (Lond)* **12**, 480-491, doi:10.1177/1745505716677139 (2016).
35 Gonzalez-Neira, A. Pharmacogenetics of chemotherapy efficacy in breast cancer. *Pharmacogenomics* **13**, 677-690, doi:10.2217/pgs.12.44 (2012).
36 Alexander, M., Coenders, F., Murray, D., Kirsa, S. & Seymour, J. Challenging historical perspectives of the 24-h chemotherapy day: Flexible chemotherapy dose-timing guidelines. *Asia Pac J Clin Oncol* **12**, e57-64, doi:10.1111/ajco.12132 (2016).
37 Thirumaran, R., Prendergast, G. C. & Gilman, P. B. in *Cancer Immunotherapy* (eds George C. Prendergast & Elizabeth M. Jaffee) 101-116 (Academic Press, 2007).
38 Bonadonna, G. *et al.* 30 years' follow up of randomised studies of adjuvant CMF in operable breast cancer: cohort study. *BMJ* **330**, 217, doi:10.1136/bmj.38314.622095.8F (2005).
39 Grant, C. H. & Gourley, C. in *Cancer Treatment and the Ovary* (eds Richard A. Anderson & Norah Spears) 21-33 (Academic Press, 2015).
40 Early Breast Cancer Trialists' Collaborative, G. *et al.* Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet* **379**, 432-444, doi:10.1016/S0140-6736(11)61625-5 (2012).
41 Weber-Schöndorfer, C. & Schaefer, C. in *Drugs During Pregnancy and Lactation (Second Edition)* (eds Christof Schaefer, Paul Peters, & Richard K. Miller) 335-367 (Academic Press, 2007).
42 Mamounas, E. P. *et al.* Paclitaxel after doxorubicin plus cyclophosphamide as adjuvant chemotherapy for node-positive breast cancer: results from NSABP B-28. *J Clin Oncol* **23**, 3686-3696, doi:10.1200/JCO.2005.10.517 (2005).
43 Smith, I. C. *et al.* Neoadjuvant chemotherapy in breast cancer: significantly enhanced response with docetaxel. *J Clin Oncol* **20**, 1456-1466, doi:10.1200/JCO.2002.20.6.1456 (2002).
44 Jones, S. E. *et al.* Phase III trial comparing doxorubicin plus cyclophosphamide with docetaxel plus cyclophosphamide as adjuvant therapy for operable breast cancer. *J Clin Oncol* **24**, 5381-5387, doi:10.1200/JCO.2006.06.5391 (2006).

45 Giordano, S. H., Lin, Y. L., Kuo, Y. F., Hortobagyi, G. N. & Goodwin, J. S. Decline in the use of anthracyclines for breast cancer. *J Clin Oncol* **30**, 2232-2239, doi:10.1200/JCO.2011.40.1273 (2012).
46 De Laurentiis, M. *et al.* Taxane-based combinations as adjuvant chemotherapy of early breast cancer: a meta-analysis of randomized trials. *J Clin Oncol* **26**, 44-53, doi:10.1200/JCO.2007.11.3787 (2008).
47 Early Breast Cancer Trialists' Collaborative, G. Increasing the dose intensity of chemotherapy by more frequent administration or sequential scheduling: a patient-level meta-analysis of 37 298 women with early breast cancer in 26 randomised trials. *Lancet* **393**, 1440-1452, doi:10.1016/S0140-6736(18)33137-4 (2019).
48 Bonilla, L. *et al.* Dose-dense chemotherapy in nonmetastatic breast cancer: a systematic review and meta-analysis of randomized controlled trials. *J Natl Cancer Inst* **102**, 1845-1854, doi:10.1093/jnci/djq409 (2010).
49 Citron, M. L. Dose-Dense Chemotherapy: Principles, Clinical Results and Future Perspectives. *Breast Care (Basel)* **3**, 251-255, doi:10.1159/000148914 (2008).
50 Al-Mahayri, Z. N., Patrinos, G. P. & Ali, B. R. Toxicity and Pharmacogenomic Biomarkers in Breast Cancer Chemotherapy. *Front Pharmacol* **11**, 445, doi:10.3389/fphar.2020.00445 (2020).
51 Wolters, R. *et al.* A comparison of international breast cancer guidelines - do the national guidelines differ in treatment recommendations? *Eur J Cancer* **48**, 1-11, doi:10.1016/j.ejca.2011.06.020 (2012).
52 Health, M. o. (2018).
53 Azim, H. A., Jr., de Azambuja, E., Colozza, M., Bines, J. & Piccart, M. J. Long-term toxic effects of adjuvant chemotherapy in breast cancer. *Ann Oncol* **22**, 1939-1947, doi:10.1093/annonc/mdq683 (2011).
54 Montemurro, F. *et al.* Self-evaluation of Adjuvant Chemotherapy-Related Adverse Effects by Patients With Breast Cancer. *JAMA Oncol* **2**, 445-452, doi:10.1001/jamaoncol.2015.4720 (2016).
55 Pearce, A. *et al.* Incidence and severity of self-reported chemotherapy side effects in routine care: A prospective cohort study. *PLoS One* **12**, e0184360, doi:10.1371/journal.pone.0184360 (2017).
56 Benson, A. B., 3rd *et al.* Recommended guidelines for the treatment of cancer treatment-induced diarrhea. *J Clin Oncol* **22**, 2918-2926, doi:10.1200/JCO.2004.04.132 (2004).
57 Abraham, J. E. *et al.* A nested cohort study of 6,248 early breast cancer patients treated in neoadjuvant and adjuvant chemotherapy trials investigating the prognostic value of chemotherapy-related toxicities. *BMC Med* **13**, 306, doi:10.1186/s12916-015-0547-5 (2015).
58 Suter, T. M. & Meier, B. Detection of anthracycline-induced cardiotoxicity: is there light at the end of the tunnel? *Ann Oncol* **13**, 647-649 (2002).
59 Schneider, B. P. *et al.* Genome-Wide Association Study for Anthracycline-Induced Congestive Heart Failure. *Clin Cancer Res* **23**, 43-51, doi:10.1158/1078-0432.CCR-16-0908 (2017).
60 Kus, T. *et al.* Polymorphism of CYP3A4 and ABCB1 genes increase the risk of neuropathy in breast cancer patients treated with paclitaxel and docetaxel. *Onco Targets Ther* **9**, 5073-5080, doi:10.2147/OTT.S106574 (2016).
61 Floyd, J., Mirza, I., Sachs, B. & Perry, M. C. Hepatotoxicity of chemotherapy. *Semin Oncol* **33**, 50-67, doi:10.1053/j.seminoncol.2005.11.002 (2006).
62 Bielopolski, D. *et al.* Paclitaxel-induced pneumonitis in patients with breast cancer: case series and review of the literature. *J Chemother* **29**, 113-117, doi:10.1179/1973947815Y.0000000029 (2017).
63 Shajahan, J., Pillai, P. S. & Jayakumar, K. N. A Prospective Comparative Study of the Toxicity Profile of 5-Flurouracil, Adriamycin, Cyclophosphamide Regimen VS Adriamycin, Paclitaxel Regimen in Patients with Locally Advanced Breast Carcinoma. *J Clin Diagn Res* **9**, FC01-06, doi:10.7860/JCDR/2015/15939.6864 (2015).
64 Howard-Anderson, J., Ganz, P. A., Bower, J. E. & Stanton, A. L. Quality of life, fertility concerns, and behavioral health outcomes in younger breast cancer survivors: a systematic review. *J Natl Cancer Inst* **104**, 386-405, doi:10.1093/jnci/djr541 (2012).
65 Carelle, N. *et al.* Changing patient perceptions of the side effects of cancer chemotherapy. *Cancer* **95**, 155-163, doi:10.1002/cncr.10630 (2002).
66 Lyman, G. H., Dale, D. C. & Crawford, J. Incidence and predictors of low dose-intensity in adjuvant breast cancer chemotherapy: a nationwide study of community practices. *J Clin Oncol* **21**, 4524-4531, doi:10.1200/JCO.2003.05.002 (2003).

67 Vandyk, A. D., Harrison, M. B., Macartney, G., Ross-White, A. & Stacey, D. Emergency department visits for symptoms experienced by oncology patients: a systematic review. *Support Care Cancer* **20**, 1589-1599, doi:10.1007/s00520-012-1459-y (2012).

68 Bonadonna, G., Valagussa, P., Moliterni, A., Zambetti, M. & Brambilla, C. Adjuvant cyclophosphamide, methotrexate, and fluorouracil in node-positive breast cancer: the results of 20 years of follow-up. *N Engl J Med* **332**, 901-906, doi:10.1056/NEJM199504063321401 (1995).

69 Lash, R. S. *et al.* A Systematic Review of Emergency Department Use Among Cancer Patients. *Cancer Nurs* **40**, 135-144, doi:10.1097/NCC.0000000000000360 (2017).

70 Klepin, H. D. *et al.* Comorbidity, chemotherapy toxicity, and outcomes among older women receiving adjuvant chemotherapy for breast cancer on a clinical trial: CALGB 49907 and CALGB 361004 (alliance). *J Oncol Pract* **10**, e285-292, doi:10.1200/JOP.2014.001388 (2014).

71 Qi, W. *et al.* The effect of reduced RDI of chemotherapy on the outcome of breast cancer patients. *Sci Rep* **10**, 13241, doi:10.1038/s41598-020-70187-8 (2020).

72 Friese, C. R. *et al.* Treatment-associated toxicities reported by patients with early-stage invasive breast cancer. *Cancer* **123**, 1925-1934, doi:10.1002/cncr.30547 (2017).

73 Bayo, J., Prieto, B. & Rivera, F. Comparison of Doctors' and Breast Cancer Patients' Perceptions of Docetaxel, Epirubicin, and Cyclophosphamide (TEC) Toxicity. *Breast J* **22**, 293-302, doi:10.1111/tbj.12571 (2016).

74 Rothwell, P. M. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* **365**, 82-93, doi:10.1016/S0140-6736(04)17670-8 (2005).

75 Scharf, O. & Colevas, A. D. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. *J Clin Oncol* **24**, 3933-3938, doi:10.1200/JCO.2005.05.3959 (2006).

76 Pitrou, I., Boutron, I., Ahmad, N. & Ravaud, P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* **169**, 1756-1761, doi:10.1001/archinternmed.2009.306 (2009).

77 Fromme, E. K., Eilers, K. M., Mori, M., Hsieh, Y. C. & Beer, T. M. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J Clin Oncol* **22**, 3485-3490, doi:10.1200/JCO.2004.03.025 (2004).

78 Di Maio, M. *et al.* Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. *J Clin Oncol* **33**, 910-915, doi:10.1200/JCO.2014.57.9334 (2015).

79 Basch, E. *et al.* Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. *Lancet Oncol* **7**, 903-909, doi:10.1016/S1470-2045(06)70910-X (2006).

80 Shapiro, C. L. & Recht, A. Side effects of adjuvant treatment of breast cancer. *N Engl J Med* **344**, 1997-2008, doi:10.1056/NEJM200106283442607 (2001).

81 Hassett, M. J., O'Malley, A. J., Pakes, J. R., Newhouse, J. P. & Earle, C. C. Frequency and cost of chemotherapy-related serious adverse effects in a population sample of women with breast cancer. *J Natl Cancer Inst* **98**, 1108-1117, doi:10.1093/jnci/djj305 (2006).

82 Du, X. L., Osborne, C. & Goodwin, J. S. Population-based assessment of hospitalizations for toxicity from chemotherapy in older women with breast cancer. *J Clin Oncol* **20**, 4636-4642, doi:10.1200/JCO.2002.05.088 (2002).

83 Ozer, H. *et al.* 2000 update of recommendations for the use of hematopoietic colony-stimulating factors: evidence-based, clinical practice guidelines. American Society of Clinical Oncology Growth Factors Expert Panel. *J Clin Oncol* **18**, 3558-3585, doi:10.1200/JCO.2000.18.20.3558 (2000).

84 Smith, T. J. *et al.* 2006 update of recommendations for the use of white blood cell growth factors: an evidence-based clinical practice guideline. *J Clin Oncol* **24**, 3187-3205, doi:10.1200/JCO.2006.06.4451 (2006).

85 Levine, M. N. *et al.* Randomized trial of intensive cyclophosphamide, epirubicin, and fluorouracil chemotherapy compared with cyclophosphamide, methotrexate, and fluorouracil in premenopausal women with node-positive breast cancer. National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* **16**, 2651-2658, doi:10.1200/JCO.1998.16.8.2651 (1998).

86 Eskander, R. N. & Tewari, K. S. Impact of chemotherapy-induced neutropenia on survival in patients with breast, ovarian and cervical cancer: a systematic review. *Journal of Hematological Malignancies* **2**, 63, doi:10.5430/jhm.v2n3p63 (2012).
87 Kim, S. B. *et al.* Docetaxel-based adjuvant therapy for breast cancer patients in Asia-Pacific region: Results from 5 years follow-up on Asia-Pacific Breast Initiative-I. *Asia Pac J Clin Oncol* **12**, 125-132, doi:10.1111/ajco.12454 (2016).
88 Kim, S. B., Kok, Y. T., Thuan, T. V., Chao, T. Y. & Shen, Z. Z. Safety Results of Docetaxel-(Taxotere(R))-Based Chemotherapy in Early Breast Cancer Patients of Asia-Pacific Region: Asia-Pacific Breast Initiative II. *J Breast Cancer* **18**, 356-364, doi:10.4048/jbc.2015.18.4.356 (2015).
89 Smith, T. J. *et al.* Recommendations for the Use of WBC Growth Factors: American Society of Clinical Oncology Clinical Practice Guideline Update. *J Clin Oncol* **33**, 3199-3212, doi:10.1200/JCO.2015.62.3488 (2015).
90 Rodgers, G. M., 3rd *et al.* Cancer- and chemotherapy-induced anemia. *J Natl Compr Canc Netw* **10**, 628-653, doi:10.6004/jnccn.2012.0064 (2012).
91 Groopman, J. E. & Itri, L. M. Chemotherapy-induced anemia in adults: incidence and treatment. *J Natl Cancer Inst* **91**, 1616-1634, doi:10.1093/jnci/91.19.1616 (1999).
92 Schwartz, R. N. Anemia in patients with cancer: incidence, causes, impact, management, and use of treatment guidelines and protocols. *Am J Health Syst Pharm* **64**, S5-13; quiz S28-30, doi:10.2146/ajhp060601 (2007).
93 Steensma, D. P. Is anemia of cancer different from chemotherapy-induced anemia? *J Clin Oncol* **26**, 1022-1024, doi:10.1200/JCO.2007.15.3874 (2008).
94 Kirshner, J., Hatch, M., Hennessy, D. D., Fridman, M. & Tannous, R. E. Anemia in stage II and III breast cancer patients treated with adjuvant doxorubicin and cyclophosphamide chemotherapy. *Oncologist* **9**, 25-32, doi:10.1634/theoncologist.9-1-25 (2004).
95 Goldrick, A. *et al.* Anemia is a common but neglected complication of adjuvant chemotherapy for early breast cancer. *Curr Oncol* **14**, 227-233, doi:10.3747/co.2007.156 (2007).
96 Chaumard, N. *et al.* Incidence and risk factors of anemia in patients with early breast cancer treated by adjuvant chemotherapy. *Breast* **21**, 464-467, doi:10.1016/j.breast.2011.10.009 (2012).
97 Sun, C. C. *et al.* Rankings and symptom assessments of side effects from chemotherapy: insights from experienced patients with ovarian cancer. *Support Care Cancer* **13**, 219-227, doi:10.1007/s00520-004-0710-6 (2005).
98 Boussios, S., Pentheroudakis, G., Katsanos, K. & Pavlidis, N. Systemic treatment-induced gastrointestinal toxicity: incidence, clinical presentation and management. *Ann Gastroenterol* **25**, 106-118 (2012).
99 Roscoe, J. A. *et al.* Patient expectation is a strong predictor of severe nausea after chemotherapy: a University of Rochester Community Clinical Oncology Program study of patients with breast carcinoma. *Cancer* **101**, 2701-2708, doi:10.1002/cncr.20718 (2004).
100 Booth, C. M. *et al.* Chemotherapy-induced nausea and vomiting in breast cancer patients: a prospective observational study. *J Support Oncol* **5**, 374-380 (2007).
101 Shih, V., Wan, H. S. & Chan, A. Clinical predictors of chemotherapy-induced nausea and vomiting in breast cancer patients receiving adjuvant doxorubicin and cyclophosphamide. *Ann Pharmacother* **43**, 444-452, doi:10.1345/aph.1L437 (2009).
102 Warr, D. G., Street, J. C. & Carides, A. D. Evaluation of risk factors predictive of nausea and vomiting with current standard-of-care antiemetic treatment: analysis of phase 3 trial of aprepitant in patients receiving adriamycin-cyclophosphamide-based chemotherapy. *Support Care Cancer* **19**, 807-813, doi:10.1007/s00520-010-0899-5 (2011).
103 Bourdeanu, L. *et al.* Chemotherapy-induced nausea and vomiting in Asian women with breast cancer receiving anthracycline-based adjuvant chemotherapy. *J Support Oncol* **10**, 149-154, doi:10.1016/j.suponc.2011.10.007 (2012).
104 Lee, K. M. *et al.* Late chronotypes are associated with neoadjuvant chemotherapy-induced nausea and vomiting in women with breast cancer. *Chronobiol Int* **34**, 480-491, doi:10.1080/07420528.2017.1295978 (2017).

105 Nawa-Nishigaki, M. *et al.* Control of Nausea and Vomiting in Patients Receiving Anthracycline/Cyclophosphamide Chemotherapy for Breast Cancer. *Anticancer Res* **38**, 877-884, doi:10.21873/anticanres.12297 (2018).
106 Kawazoe, H. *et al.* Patient-related Risk Factors for Nausea and Vomiting with Standard Antiemetics in Patients with Breast Cancer Receiving Anthracycline-based Chemotherapy: A Retrospective Observational Study. *Clin Ther* **40**, 2170-2179, doi:10.1016/j.clinthera.2018.10.004 (2018).
107 Velcheti, V. & Punekar, S. R. *Handbook of cancer treatment-related toxicities*. (Elsevier, 2021).
108 Hesketh, P. J. Chemotherapy-induced nausea and vomiting. *N Engl J Med* **358**, 2482-2494, doi:10.1056/NEJMra0706547 (2008).
109 Basch, E. *et al.* Antiemetics: American Society of Clinical Oncology clinical practice guideline update. *J Clin Oncol* **29**, 4189-4198, doi:10.1200/JCO.2010.34.4614 (2011).
110 Mosa, A. S. M., Hossain, A. M., Lavoie, B. J. & Yoo, I. Patient-Related Risk Factors for Chemotherapy-Induced Nausea and Vomiting: A Systematic Review. *Front Pharmacol* **11**, 329, doi:10.3389/fphar.2020.00329 (2020).
111 Stein, A., Voigt, W. & Jordan, K. Chemotherapy-induced diarrhea: pathophysiology, frequency and guideline-based management. *Ther Adv Med Oncol* **2**, 51-63, doi:10.1177/1758834009355164 (2010).
112 Gibson, R. J. & Stringer, A. M. Chemotherapy-induced diarrhoea. *Curr Opin Support Palliat Care* **3**, 31-35, doi:10.1097/SPC.0b013e32832531bb (2009).
113 Keefe, D. M. *et al.* Updated clinical practice guidelines for the prevention and treatment of mucositis. *Cancer* **109**, 820-831, doi:10.1002/cncr.22484 (2007).
114 Lichtman, S. M. *et al.* International Society of Geriatric Oncology Chemotherapy Taskforce: evaluation of chemotherapy in older patients--an analysis of the medical literature. *J Clin Oncol* **25**, 1832-1843, doi:10.1200/JCO.2007.10.6583 (2007).
115 Elad, S. *et al.* MASCC/ISOO clinical practice guidelines for the management of mucositis secondary to cancer therapy. *Cancer* **126**, 4423-4431, doi:10.1002/cncr.33100 (2020).
116 Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. & Relman, D. A. The application of ecological theory toward an understanding of the human microbiome. *Science* **336**, 1255-1262, doi:10.1126/science.1224203 (2012).
117 Backhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 852, doi:10.1016/j.chom.2015.05.012 (2015).
118 Dominguez-Bello, M. G. *et al.* Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107**, 11971-11975, doi:10.1073/pnas.1002601107 (2010).
119 Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789-799, doi:10.1016/j.cell.2014.09.053 (2014).
120 David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome Biol* **15**, R89, doi:10.1186/gb-2014-15-7-r89 (2014).
121 Tulstrup, M. V. *et al.* Antibiotic Treatment Affects Intestinal Permeability and Gut Microbial Composition in Wistar Rats Dependent on Antibiotic Class. *PLoS One* **10**, e0144854, doi:10.1371/journal.pone.0144854 (2015).
122 Backhed, F. *et al.* Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe* **17**, 690-703, doi:10.1016/j.chom.2015.04.004 (2015).
123 Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science* **341**, 1237439, doi:10.1126/science.1237439 (2013).
124 David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559-563, doi:10.1038/nature12820 (2014).
125 Dzutsev, A., Goldszmid, R. S., Viaud, S., Zitvogel, L. & Trinchieri, G. The role of the microbiota in inflammation, carcinogenesis, and cancer therapy. *Eur J Immunol* **45**, 17-31, doi:10.1002/eji.201444972 (2015).
126 Bosch, T. C. & McFall-Ngai, M. J. Metaorganisms as the new frontier. *Zoology (Jena)* **114**, 185-190, doi:10.1016/j.zool.2011.04.001 (2011).
127 Kamada, N., Chen, G. Y., Inohara, N. & Nunez, G. Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* **14**, 685-690, doi:10.1038/ni.2608 (2013).

128 Fulde, M. & Hornef, M. W. Maturation of the enteric mucosal innate immune system during the postnatal period. *Immunol Rev* **260**, 21-34, doi:10.1111/imr.12190 (2014).
129 Ijssennagger, N. *et al.* Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Proc Natl Acad Sci U S A* **112**, 10038-10043, doi:10.1073/pnas.1507645112 (2015).
130 Reinhardt, C. *et al.* Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodelling. *Nature* **483**, 627-631, doi:10.1038/nature10893 (2012).
131 Neuman, H., Debelius, J. W., Knight, R. & Koren, O. Microbial endocrinology: the interplay between the microbiota and the endocrine system. *FEMS Microbiol Rev* **39**, 509-521, doi:10.1093/femsre/fuu010 (2015).
132 Canfora, E. E., Jocken, J. W. & Blaak, E. E. Short-chain fatty acids in control of body weight and insulin sensitivity. *Nat Rev Endocrinol* **11**, 577-591, doi:10.1038/nrendo.2015.128 (2015).
133 Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222-227, doi:10.1038/nature11053 (2012).
134 Yano, J. M. *et al.* Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264-276, doi:10.1016/j.cell.2015.02.047 (2015).
135 Devlin, A. S. & Fischbach, M. A. A biosynthetic pathway for a prominent class of microbiota-derived bile acids. *Nat Chem Biol* **11**, 685-690, doi:10.1038/nchembio.1864 (2015).
136 Haiser, H. J. *et al.* Predicting and manipulating cardiac drug inactivation by the human gut bacterium Eggerthella lenta. *Science* **341**, 295-298, doi:10.1126/science.1235872 (2013).
137 Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635-1638, doi:10.1126/science.1110591 (2005).
138 Rea, D. *et al.* Microbiota effects on cancer: from risks to therapies. *Oncotarget* **9**, 17915-17927, doi:10.18632/oncotarget.24681 (2018).
139 Claesson, M. J. *et al.* Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**, 178-184, doi:10.1038/nature11319 (2012).
140 Cammarota, G. *et al.* The involvement of gut microbiota in inflammatory bowel disease pathogenesis: potential for therapy. *Pharmacol Ther* **149**, 191-212, doi:10.1016/j.pharmthera.2014.12.006 (2015).
141 de Theije, C. G. *et al.* Altered gut microbiota and activity in a murine model of autism spectrum disorders. *Brain Behav Immun* **37**, 197-206, doi:10.1016/j.bbi.2013.12.005 (2014).
142 Sekirov, I., Russell, S. L., Antunes, L. C. & Finlay, B. B. Gut microbiota in health and disease. *Physiol Rev* **90**, 859-904, doi:10.1152/physrev.00045.2009 (2010).
143 Zhang, Y. J. *et al.* Impacts of gut bacteria on human health and diseases. *Int J Mol Sci* **16**, 7493-7519, doi:10.3390/ijms16047493 (2015).
144 Spor, A., Koren, O. & Ley, R. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol* **9**, 279-290, doi:10.1038/nrmicro2540 (2011).
145 Fernandez, M. F. *et al.* Breast Cancer and Its Relationship with the Microbiota. *Int J Environ Res Public Health* **15**, doi:10.3390/ijerph15081747 (2018).
146 Goedert, J. J. *et al.* Investigation of the association between the fecal microbiota and breast cancer in postmenopausal women: a population-based case-control pilot study. *J Natl Cancer Inst* **107**, doi:10.1093/jnci/djv147 (2015).
147 Goedert, J. J. *et al.* Postmenopausal breast cancer and oestrogen associations with the IgA-coated and IgA-noncoated faecal microbiota. *Br J Cancer* **118**, 471-479, doi:10.1038/bjc.2017.435 (2018).
148 Miko, E. *et al.* Lithocholic acid, a bacterial metabolite reduces breast cancer cell proliferation and aggressiveness. *Biochim Biophys Acta Bioenerg* **1859**, 958-974, doi:10.1016/j.bbabio.2018.04.002 (2018).
149 Kovacs, T. *et al.* Cadaverine, a metabolite of the microbiome, reduces breast cancer aggressiveness through trace amino acid receptors. *Sci Rep* **9**, 1300, doi:10.1038/s41598-018-37664-7 (2019).
150 Luu, T. H. *et al.* Intestinal Proportion of Blautia sp. is Associated with Clinical Stage and Histoprognostic Grade in Patients with Early-Stage Breast Cancer. *Nutr Cancer* **69**, 267-275, doi:10.1080/01635581.2017.1263750 (2017).
151 Yang, P., Wang, Z., Peng, Q., Lian, W. & Chen, D. Comparison of the Gut Microbiota in Patients with Benign and Malignant Breast Tumors: A Pilot Study. *Evol Bioinform Online* **17**, 11769343211057573, doi:10.1177/11769343211057573 (2021).

152 Wu, A. H. *et al.* Gut microbiome associations with breast cancer risk factors and tumor characteristics: a pilot study. *Breast Cancer Res Treat* **182**, 451-463, doi:10.1007/s10549-020-05702-6 (2020).
153 Fruge, A. D. *et al.* Fecal Akkermansia muciniphila Is Associated with Body Composition and Microbiota Diversity in Overweight and Obese Women with Breast Cancer Participating in a Presurgical Weight Loss Trial. *J Acad Nutr Diet*, doi:10.1016/j.jand.2018.08.164 (2018).
154 Bertazzoni, M. E. *et al.* Intestinal Microflora as an Alternative Metabolic Source of Estrogens in Women with Uterine Leiomyoma and Breast Cancer. *Annals of the New York Academy of Sciences* **595**, 473-479, doi:10.1111/j.1749-6632.1990.tb34337.x (2006).
155 Zhu, J. *et al.* Breast cancer in postmenopausal women is associated with an altered gut metagenome. *Microbiome* **6**, 136, doi:10.1186/s40168-018-0515-3 (2018).
156 Hou, M. F. *et al.* Comprehensive profiles and diagnostic value of menopausal-specific gut microbiota in premenopausal breast cancer. *Exp Mol Med* **53**, 1636-1646, doi:10.1038/s12276-021-00686-9 (2021).
157 Touchefeu, Y. *et al.* Systematic review: the role of the gut microbiota in chemotherapy- or radiation-induced gastrointestinal mucositis - current evidence and potential clinical applications. *Aliment Pharmacol Ther* **40**, 409-421, doi:10.1111/apt.12878 (2014).
158 Nyhlen, A., Ljungberg, B., Nilsson-Ehle, I. & Nord, C. E. Impact of combinations of antineoplastic drugs on intestinal microflora in 9 patients with leukaemia. *Scand J Infect Dis* **34**, 17-21 (2002).
159 van Vliet, M. J. *et al.* Chemotherapy treatment in pediatric patients with acute myeloid leukemia receiving antimicrobial prophylaxis leads to a relative increase of colonization with potentially pathogenic bacteria in the gut. *Clin Infect Dis* **49**, 262-270, doi:10.1086/599346 (2009).
160 Galloway-Pena, J. R. *et al.* The role of the gastrointestinal microbiome in infectious complications during induction chemotherapy for acute myeloid leukemia. *Cancer* **122**, 2186-2196, doi:10.1002/cncr.30039 (2016).
161 Galloway-Pena, J. R. *et al.* Characterization of oral and gut microbiome temporal variability in hospitalized cancer patients. *Genome Med* **9**, 21, doi:10.1186/s13073-017-0409-1 (2017).
162 Galloway-Pena, J. R. *et al.* Gut Microbiome Signatures Are Predictive of Infectious Risk Following Induction Therapy for Acute Myeloid Leukemia. *Clin Infect Dis* **71**, 63-71, doi:10.1093/cid/ciz777 (2020).
163 Montassier, E. *et al.* 16S rRNA gene pyrosequencing reveals shift in patient faecal microbiota during high-dose chemotherapy as conditioning regimen for bone marrow transplantation. *Microb Ecol* **67**, 690-699, doi:10.1007/s00248-013-0355-4 (2014).
164 Montassier, E. *et al.* Chemotherapy-driven dysbiosis in the intestinal microbiome. *Aliment Pharmacol Ther* **42**, 515-528, doi:10.1111/apt.13302 (2015).
165 Kong, C. *et al.* Alterations in intestinal microbiota of colorectal cancer patients receiving radical surgery combined with adjuvant CapeOx therapy. *Sci China Life Sci* **62**, 1178-1193, doi:10.1007/s11427-018-9456-x (2019).
166 Deng, X. *et al.* Comparison of Microbiota in Patients Treated by Surgery or Chemotherapy by 16S rRNA Sequencing Reveals Potential Biomarkers for Colorectal Cancer Therapy. *Front Microbiol* **9**, 1607, doi:10.3389/fmicb.2018.01607 (2018).
167 Shuwen, H. *et al.* Effects of postoperative adjuvant chemotherapy and palliative chemotherapy on the gut microbiome in colorectal cancer. *Microb Pathog* **149**, 104343, doi:10.1016/j.micpath.2020.104343 (2020).
168 Tong, J. *et al.* Changes of Intestinal Microbiota in Ovarian Cancer Patients Treated with Surgery and Chemotherapy. *Cancer Manag Res* **12**, 8125-8135, doi:10.2147/CMAR.S265205 (2020).
169 Zwielehner, J. *et al.* Changes in human fecal microbiota due to chemotherapy analyzed by TaqMan-PCR, 454 sequencing and PCR-DGGE fingerprinting. *PLoS One* **6**, e28654, doi:10.1371/journal.pone.0028654 (2011).
170 Von Bultzingslowen, I., Adlerberth, I., Wold, A. E., Dahlen, G. & Jontell, M. Oral and intestinal microflora in 5-fluorouracil treated rats, translocation to cervical and mesenteric lymph nodes and effects of probiotic bacteria. *Oral Microbiol Immunol* **18**, 278-284 (2003).
171 Stringer, A. M. *et al.* Gastrointestinal microflora and mucins may play a critical role in the development of 5-Fluorouracil-induced gastrointestinal mucositis. *Exp. Biol. Med. (Maywood)* **234**, 430-441, doi:10.3181/0810-RM-301 (2009).

172 Hamouda, N. *et al.* Apoptosis, Dysbiosis and Expression of Inflammatory Cytokines are Sequential Events in the Development of 5-Fluorouracil-Induced Intestinal Mucositis in Mice. *Basic Clin Pharmacol Toxicol* **121**, 159-168, doi:10.1111/bcpt.12793 (2017).

173 Li, H. L. *et al.* Alteration of Gut Microbiota and Inflammatory Cytokine/Chemokine Profiles in 5-Fluorouracil Induced Intestinal Mucositis. *Front Cell Infect Microbiol* **7**, 455, doi:10.3389/fcimb.2017.00455 (2017).

174 Yuan, L. *et al.* The influence of gut microbiota dysbiosis to the efficacy of 5-Fluorouracil treatment on colorectal cancer. *Biomed Pharmacother* **108**, 184-193, doi:10.1016/j.biopha.2018.08.165 (2018).

175 Vanlancker, E., Vanhoecke, B., Stringer, A. & Van de Wiele, T. 5-Fluorouracil and irinotecan (SN-38) have limited impact on colon microbial functionality and composition in vitro. *PeerJ* **5**, e4017, doi:10.7717/peerj.4017 (2017).

176 Viaud, S. *et al.* The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science* **342**, 971-976, doi:10.1126/science.1240537 (2013).

177 Yang, J., Liu, K. X., Qu, J. M. & Wang, X. D. The changes induced by cyclophosphamide in intestinal barrier and microflora in mice. *Eur J Pharmacol* **714**, 120-124, doi:10.1016/j.ejphar.2013.06.006 (2013).

178 Xu, X. & Zhang, X. Effects of cyclophosphamide on immune system and gut microbiota in mice. *Microbiol Res* **171**, 97-106, doi:10.1016/j.micres.2014.11.002 (2015).

179 Panebianco, C. *et al.* Influence of gemcitabine chemotherapy on the microbiota of pancreatic cancer xenografted mice. *Cancer Chemother Pharmacol* **81**, 773-782, doi:10.1007/s00280-018-3549-0 (2018).

180 Fijlstra, M. *et al.* Substantial decreases in the number and diversity of microbiota during chemotherapy-induced gastrointestinal mucositis in a rat model. *Support Care Cancer* **23**, 1513-1522, doi:10.1007/s00520-014-2487-6 (2015).

181 Stringer, A. M. *et al.* Chemotherapy-induced diarrhea is associated with changes in the luminal environment in the DA rat. *Exp. Biol. Med. (Maywood)* **232**, 96-106 (2007).

182 Stringer, A. M. *et al.* Faecal microflora and β-glucuronidase expression are altered in an irinotecan-induced diarrhea model in rats. *Cancer Biology & Therapy* **7**, 1919-1925, doi:10.4161/cbt.7.12.6940 (2008).

183 Lin, X. B. *et al.* Irinotecan (CPT-11) chemotherapy alters intestinal microbiota in tumour bearing rats. *PLoS One* **7**, e39764, doi:10.1371/journal.pone.0039764 (2012).

184 Forsgard, R. A. *et al.* Chemotherapy-induced gastrointestinal toxicity is associated with changes in serum and urine metabolome and fecal microbiota in male Sprague-Dawley rats. *Cancer Chemother Pharmacol* **80**, 317-332, doi:10.1007/s00280-017-3364-z (2017).

185 Buchta Rosean, C. M. & Rutkowski, M. R. The influence of the commensal microbiota on distal tumor-promoting inflammation. *Semin Immunol* **32**, 62-73, doi:10.1016/j.smim.2017.06.002 (2017).

186 Lassen, K. *et al.* Consensus review of optimal perioperative care in colorectal surgery: Enhanced Recovery After Surgery (ERAS) Group recommendations. *Arch Surg* **144**, 961-969, doi:10.1001/archsurg.2009.170 (2009).

187 Bucaneve, G. *et al.* Levofloxacin to prevent bacterial infection in patients with cancer and neutropenia. *N Engl J Med* **353**, 977-987, doi:10.1056/NEJMoa044097 (2005).

188 Cullen, M. *et al.* Antibacterial prophylaxis after chemotherapy for solid tumors and lymphomas. *N Engl J Med* **353**, 988-998, doi:10.1056/NEJMoa050078 (2005).

189 Stringer, A. M. *et al.* Biomarkers of chemotherapy-induced diarrhoea: a clinical study of intestinal microbiome alterations, inflammation and circulating matrix metalloproteinases. *Support Care Cancer* **21**, 1843-1852, doi:10.1007/s00520-013-1741-7 (2013).

190 Alexander, J. L. *et al.* Gut microbiota modulation of chemotherapy efficacy and toxicity. *Nat Rev Gastroenterol Hepatol* **14**, 356-365, doi:10.1038/nrgastro.2017.20 (2017).

191 Yip, L. Y. & Chan, E. C. Investigation of Host-Gut Microbiota Modulation of Therapeutic Outcome. *Drug Metab Dispos* **43**, 1619-1631, doi:10.1124/dmd.115.063750 (2015).

192 Paulos, C. M. *et al.* Microbial translocation augments the function of adoptively transferred self/tumor-specific CD8+ T cells via TLR4 signaling. *J Clin Invest* **117**, 2197-2204, doi:10.1172/JCI32205 (2007).

193 Ciftciler, R. & Ciftciler, A. E. The importance of microbiota in hematology. *Transfus Apher Sci*, 103320, doi:10.1016/j.transci.2021.103320 (2021).

194 Rattanathammethee, T. *et al.* Gut microbiota profiles of treatment-naive adult acute myeloid leukemia patients with neutropenic fever during intensive chemotherapy. *PLoS One* **15**, e0236460, doi:10.1371/journal.pone.0236460 (2020).
195 Daillere, R. *et al.* Enterococcus hirae and Barnesiella intestinihominis Facilitate Cyclophosphamide-Induced Therapeutic Immunomodulatory Effects. *Immunity* **45**, 931-943, doi:10.1016/j.immuni.2016.09.009 (2016).
196 Guthrie, L., Gupta, S., Daily, J. & Kelly, L. Human microbiome signatures of differential colorectal cancer drug metabolism. *NPJ Biofilms Microbiomes* **3**, 27, doi:10.1038/s41522-017-0034-1 (2017).
197 Fei, Z. *et al.* Gut microbiome associated with chemotherapy-induced diarrhea from the CapeOX regimen as adjuvant chemotherapy in resected stage III colorectal cancer. *Gut Pathog* **11**, 18, doi:10.1186/s13099-019-0299-4 (2019).
198 Uzan-Yulzari, A. *et al.* The intestinal microbiome, weight, and metabolic changes in women treated by adjuvant chemotherapy for breast and gynecological malignancies. *BMC Med* **18**, 281, doi:10.1186/s12916-020-01751-2 (2020).
199 Terrisse, S. *et al.* Intestinal microbiota influences clinical outcome and side effects of early breast cancer treatment. *Cell Death Differ* **28**, 2778-2796, doi:10.1038/s41418-021-00784-1 (2021).
200 Wallace, B. D. *et al.* Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**, 831-835, doi:10.1126/science.1191175 (2010).
201 Kurita, A. *et al.* Streptomycin alleviates irinotecan-induced delayed-onset diarrhea in rats by a mechanism other than inhibition of beta-glucuronidase activity in intestinal lumen. *Cancer Chemother Pharmacol* **67**, 201-213, doi:10.1007/s00280-010-1310-4 (2011).
202 Frank, M. *et al.* TLR signaling modulates side effects of anticancer therapy in the small intestine. *J Immunol* **194**, 1983-1995, doi:10.4049/jimmunol.1402481 (2015).
203 Shen, S. *et al.* Gut microbiota is critical for the induction of chemotherapy-induced pain. *Nat Neurosci* **20**, 1213-1216, doi:10.1038/nn.4606 (2017).
204 Ramakrishna, C. *et al.* Dominant Role of the Gut Microbiota in Chemotherapy Induced Neuropathic Pain. *Sci Rep* **9**, 20324, doi:10.1038/s41598-019-56832-x (2019).
205 Hanson, B. M. & Weinstock, G. M. The importance of the microbiome in epidemiologic research. *Ann Epidemiol* **26**, 301-305, doi:10.1016/j.annepidem.2016.03.008 (2016).
206 Roy, S. & Trinchieri, G. Microbiota: a key orchestrator of cancer therapy. *Nat Rev Cancer* **17**, 271-285, doi:10.1038/nrc.2017.13 (2017).
207 Nguyen, S. M. *et al.* Association of fruit, vegetable and animal food intakes with breast cancer risk overall and by molecular subtype among Vietnamese women. *Cancer Epidemiol Biomarkers Prev*, doi:10.1158/1055-9965.EPI-21-1085 (2022).
208 Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* **95**, 103208, doi:https://doi.org/10.1016/j.jbi.2019.103208 (2019).
209 Caplan, L. Delay in breast cancer: implications for stage at diagnosis and survival. *Front Public Health* **2**, 87, doi:https://doi.org/10.3389/fpubh.2014.00087 (2014).
210 Vavra, K. L., Saadeh, C. E., Rosen, A. L., Uptigrove, C. E. & Srkalovic, G. Improving the relative dose intensity of systemic chemotherapy in a community-based outpatient cancer center. *J Oncol Pract* **9**, e203-211, doi:10.1200/JOP.2012.000810 (2013).
211 Ding, W. *et al.* Anthracycline versus nonanthracycline adjuvant therapy for early breast cancer: A systematic review and meta-analysis. *Medicine (Baltimore)* **97**, e12908, doi:10.1097/MD.0000000000012908 (2018).
212 Barcenas, C. H. *et al.* Risk of hospitalization according to chemotherapy regimen in early-stage breast cancer. *J Clin Oncol* **32**, 2010-2017, doi:10.1200/JCO.2013.49.3676 (2014).
213 Rajan, S. S., Stearns, S. C., Lyman, G. H. & Carpenter, W. R. Effect of primary prophylactic G-CSF use on systemic therapy administration for elderly breast cancer patients. *Breast Cancer Res Treat* **130**, 255-266, doi:10.1007/s10549-011-1553-8 (2011).
214 Rajan, S. S., Lyman, G. H., Carpenter, W. R. & Stearns, S. C. Chemotherapy characteristics are important predictors of primary prophylactic CSF administration in older patients with breast cancer. *Breast Cancer Res Treat* **127**, 511-520, doi:10.1007/s10549-010-1216-1 (2011).
215 Rajan, S. S., Lyman, G. H., Stearns, S. C. & Carpenter, W. R. Effect of primary prophylactic granulocyte-colony stimulating factor use on incidence of neutropenia hospitalizations for elderly early-

stage breast cancer patients receiving chemotherapy. *Med Care* **49**, 649-657, doi:10.1097/MLR.0b013e318215c42e (2011).
216 Gadisa, D. A., Assefa, M., Tefera, G. M. & Yimer, G. Patterns of Anthracycline-Based Chemotherapy-Induced Adverse Drug Reactions and Their Impact on Relative Dose Intensity among Women with Breast Cancer in Ethiopia: A Prospective Observational Study. *J Oncol* **2020**, 2636514, doi:10.1155/2020/2636514 (2020).
217 Rasic, A., Sofic, A., Beslija, S., Rasic, I. & Hasanbegovic, B. Effects of adding taxane to anthracycline-based neoadjuvant chemotherapy in locally advanced breast cancer. *Med Glas (Zenica)* **16**, 77-82, doi:10.17392/964-19 (2019).
218 Crawford, J., Dale, D. C. & Lyman, G. H. Chemotherapy-induced neutropenia: risks, consequences, and new directions for its management. *Cancer* **100**, 228-237, doi:10.1002/cncr.11882 (2004).
219 Huang, R. S. & Ratain, M. J. Pharmacogenetics and pharmacogenomics of anticancer agents. *CA Cancer J Clin* **59**, 42-55, doi:10.3322/caac.20002 (2009).
220 Lee, L., Cheung, W. Y., Atkinson, E. & Krzyzanowska, M. K. Impact of comorbidity on chemotherapy use and outcomes in solid tumors: a systematic review. *J Clin Oncol* **29**, 106-117, doi:10.1200/JCO.2010.31.3049 (2011).
221 Han, H. S. *et al.* Racial differences in acute toxicities of neoadjuvant or adjuvant chemotherapy in patients with early-stage breast cancer. *Eur J Cancer* **47**, 2537-2545, doi:10.1016/j.ejca.2011.06.027 (2011).
222 Chow, L. W. C. *et al.* Toxicity profile differences of adjuvant docetaxel/cyclophosphamide (TC) between Asian and Caucasian breast cancer patients. *Asia Pac J Clin Oncol* **13**, 372-378, doi:10.1111/ajco.12682 (2017).
223 Carroll, J., Protani, M., Walpole, E. & Martin, J. H. Effect of obesity on toxicity in women treated with adjuvant chemotherapy for early-stage breast cancer: a systematic review. *Breast Cancer Res Treat* **136**, 323-330, doi:10.1007/s10549-012-2213-3 (2012).
224 Barpe, D. R., Rosa, D. D. & Froehlich, P. E. Pharmacokinetic evaluation of doxorubicin plasma levels in normal and overweight patients with breast cancer and simulation of dose adjustment by different indexes of body mass. *Eur J Pharm Sci* **41**, 458-463, doi:10.1016/j.ejps.2010.07.015 (2010).
225 Lichtman, S. M. *et al.* Effect of Pretreatment Renal Function on Treatment and Clinical Outcomes in the Adjuvant Treatment of Older Women With Breast Cancer: Alliance A171201, an Ancillary Study of CALGB/CTSU 49907. *J Clin Oncol* **34**, 699-705, doi:10.1200/JCO.2015.62.6341 (2016).
226 Peterson, L. L. *et al.* Association between renal function and chemotherapy-related toxicity in older adults with cancer. *J Geriatr Oncol* **8**, 96-101, doi:10.1016/j.jgo.2016.10.004 (2017).
227 Boxer, L. & Dale, D. C. Neutropenia: causes and consequences. *Semin Hematol* **39**, 75-81, doi:10.1053/shem.2002.31911 (2002).
228 Polyak, K. Breast cancer: origins and evolution. *J Clin Invest* **117**, 3155-3163, doi:10.1172/JCI33295 (2007).
229 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
230 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357 (2012).
231 Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome biology* **20**, 257 (2019).
232 Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science* **3**, e104 (2017).
233 Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol* **39**, 105-114, doi:10.1038/s41587-020-0603-3 (2021).
234 Shao, Y. *et al.* Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117-121 (2019).
235 Simon, H. Y., Siddle, K. J., Park, D. J. & Sabeti, P. C. Benchmarking metagenomics tools for taxonomic classification. *Cell* **178**, 779-794 (2019).
236 Weiss, S. *et al.* Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**, 27 (2017).
237 Oksanen, J. *et al.* Package 'vegan'. *Community ecology package, version* **2**, 1-295 (2013).

238 Tang, Z.-Z., Chen, G. & Alekseyenko, A. V. PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. *Bioinformatics* **32**, 2618-2625 (2016).
239 Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V. & Egozcue, J. J. It's all relative: analyzing microbiome data as compositions. *Annals of epidemiology* **26**, 322-329 (2016).
240 Duncan, S. H. *et al.* Proposal of Roseburia faecis sp. nov., Roseburia hominis sp. nov. and Roseburia inulinivorans sp. nov., based on isolates from human faeces. *Int J Syst Evol Microbiol* **56**, 2437-2441, doi:10.1099/ijs.0.64098-0 (2006).
241 Liu, H. *et al.* Butyrate: A Double-Edged Sword for Health? *Adv Nutr* **9**, 21-29, doi:10.1093/advances/nmx009 (2018).
242 Riviere, A., Selak, M., Lantin, D., Leroy, F. & De Vuyst, L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front Microbiol* **7**, 979, doi:10.3389/fmicb.2016.00979 (2016).
243 Amin, A. L., Purdy, A. C., Mattingly, J. D., Kong, A. L. & Termuhlen, P. M. Benign breast disease. *Surg Clin North Am* **93**, 299-308, doi:10.1016/j.suc.2013.01.001 (2013).
244 Nga do, T. T. *et al.* Antibiotic sales in rural and urban pharmacies in northern Vietnam: an observational study. *BMC Pharmacol Toxicol* **15**, 6, doi:10.1186/2050-6511-15-6 (2014).
245 Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nature methods* **15**, 962 (2018).
246 Guo, P. T., Zhang, K., Ma, X. & He, P. L. Clostridium species as probiotics: potentials and challenges. *J Anim Sci Biotechno* **11**, doi:ARTN 24
10.1186/s40104-019-0402-1 (2020).
247 Domingo, M. C. *et al.* Cloacibacillus sp., a Potential Human Pathogen Associated with Bacteremia in Quebec and New Brunswick. *J Clin Microbiol* **53**, 3380-3383, doi:10.1128/JCM.01137-15 (2015).
248 Palleja, A. *et al.* Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat Microbiol* **3**, 1255-1265, doi:10.1038/s41564-018-0257-9 (2018).
249 Ramirez, J. *et al.* Antibiotics as Major Disruptors of Gut Microbiota. *Front Cell Infect Microbiol* **10**, 572912, doi:10.3389/fcimb.2020.572912 (2020).
250 Raphael, M. J. *et al.* The relationship between time to initiation of adjuvant chemotherapy and survival in breast cancer: a systematic review and meta-analysis. *Breast Cancer Res Treat* **160**, 17-28, doi:10.1007/s10549-016-3960-3 (2016).
251 Weersma, R. K., Zhernakova, A. & Fu, J. Interaction between drugs and the gut microbiome. *Gut* **69**, 1510-1519, doi:10.1136/gutjnl-2019-320204 (2020).
252 Koppel, N., Maini Rekdal, V. & Balskus, E. P. Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356**, doi:10.1126/science.aag2770 (2017).
253 Schoch, C. L. *et al.* NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)* **2020**, doi:10.1093/database/baaa062 (2020).
254 Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res* **48**, D445-D453, doi:10.1093/nar/gkz862 (2020).
255 He, C., Liu, Y., Ye, S., Yin, S. & Gu, J. Changes of intestinal microflora of breast cancer in premenopausal women. *Eur J Clin Microbiol Infect Dis* **40**, 503-513, doi:10.1007/s10096-020-04036-x (2021).