Clinical characterization of the cancer treatment using the Oncology CDM

2020-02-25

Hokyun Jeon, Seng Chan You MD, MS,
Rae Woong Park MD, PhD

Backgrounds

FOLFOX

FOL - FOLinic acid (leucovorin)
F - Fluorouracil (5-FU)
OX - OXaliplatin

X 12 cycles

Original Study

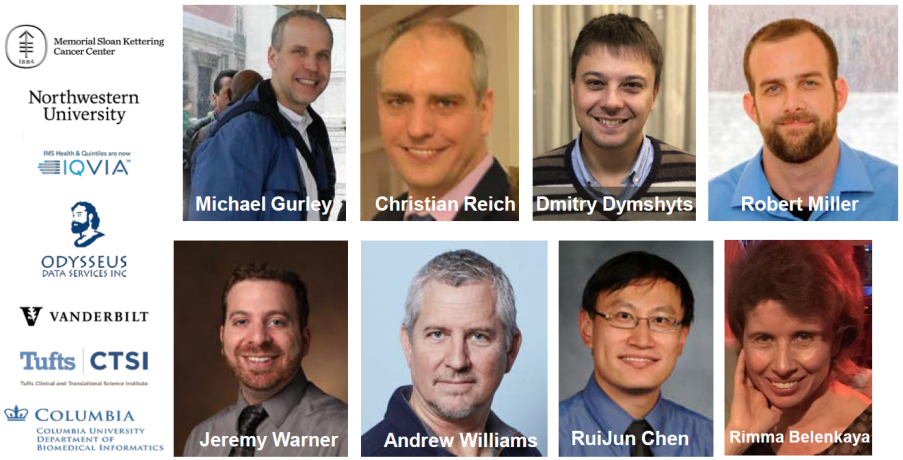
Check for updates

Effects of Proton Pump Inhibitors on FOLFOX and CapeOx Regimens in Colorectal Cancer

Grace G. Wong,[1] Vincent Ha,[1] Michael P. Chu,[2] Deonne Dersch-Mills,[4] Sunita Ghosh,[3] Carole R. Chambers,[5] Michael B. Sawyer[2]

- Cancer patients are treated with **combination of the multiple drugs** in several **cycles**.
- Treatment studies for cancer patient are based on **regimen-level comparison**.

Backgrounds



Episode table

'Treatment regimen' of 'FOLFOX'

Episode_id	1234
Person_id	1
Episode_concept_id	**32531**
Episode_start_datetime	2019-06-09
Episode_parent_id	12345
Episode_object_concept_id	**35806596**
Episode_type_concept_id	**32545**

OMOP concept **'Treatment regimen'**

Foreign key to the disease Episode record

HemOnc concept **'FOLFOX'**

OMOP concept **'Episode algorithmically derived from EHR'**

- Oncology working group proposed **Episode / Episode_event** table to curate regimen-level records of the cancer patients in Common Data Model.

Journal of Biomedical Informatics 96 (2019) 103239

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

HemOnc: A new standard vocabulary for chemotherapy regimen representation in the OMOP common data model

Jeremy L. Warner[a,b,*], Dmitry Dymshyts[c], Christian G. Reich[d], Michael J. Gurley[e], Harry Hochheiser[f], Zachary H. Moldwin[g], Rimma Belenkaya[h], Andrew E. Williams[i], Peter C. Yang[b,j]

As the standard reference and vocabulary of the chemotherapy regimen, the oncology WG used **HemOnc database.**

Backgrounds

FOLFOX4

FOLFOX4: FOLinic acid, Fluorouracil, OXaliplatin

back to top

Regimen

Study	Years of enrollment	Evidence	Comparator	Comparative Efficacy
André et al. 2004 (MOSAIC)	1998-2001	Phase III (E-RT-esc)	FULV	Seems to have superior OS (*)
de Gramont et al. 2012 (AVANT)		Phase III (C)	1. FOLFOX4 & Bevacizumab	Did not meet primary endpoint of DFS
			2. XELOX & Bevacizumab	Did not meet primary endpoint of DFS
Taieb et al. 2014 (PETACC-8)		Phase III (C)	FOLFOX4 & Cetuximab	Did not meet primary endpoint of DFS
Grothey et al. 2018 (IDEA)		Phase III (C)	FOLFOX4 x 6	Inconclusive whether non-inferior DFS36

Note: Reported efficacy for MOSAIC is based on the 2009 update. IDEA is a pooled analysis of six phase III RCTs.

Preceding treatment

- Surgery

Chemotherapy

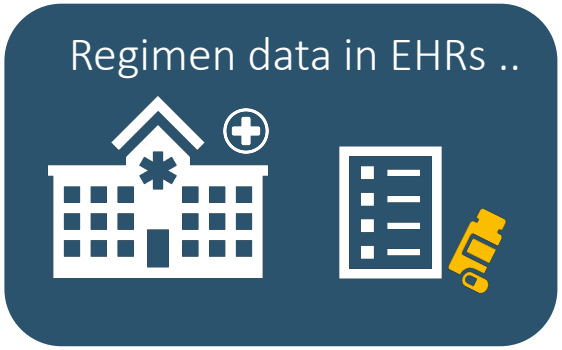
- Fluorouracil (5-FU) 400 mg/m^2 IV bolus once per day on days 1 & 2, then 600 mg/m^2 IV continuous infusion over 22 hours after each bolus, **given second** (total dose per cycle: 2000 mg/m^2)
- Folinic acid (Leucovorin) 200 mg/m^2 IV over 2 hours once per day on days 1 & 2, **given first, with oxaliplatin on day 1**
- Oxaliplatin (Eloxatin) 85 mg/m^2 IV over 2 hours once on day 1, **given first, with leucovorin**

14-day cycle for 12 cycles

HemOnc Database includes 3,362 of chemotherapy regimen descriptions.

- The composition of the drugs in 1 cycle treatment
- Gap dates between each cycle
- The recommended number of cycles

Backgrounds



Although the structure for chemotherapy records are constructed by Oncology WG, it is not easy to extract chemotherapy regimen from **Electronic Health Records (EHRs)** compared to registry data.

- Most of the cases, the chemotherapy regimen descripted in **narrative medical notes**.
- So, **manual process** is inevitable to extract chemotherapy regimen information from EHRs.

Objective

1. Development of R packages for

- A package to **extract regimen-level treatment records** from EHRs / Claim data.
- An R package to generate **episode / episode_event table** based on extracted treatment.
- An R package to visualize of **cohort pathways.**
- **Validation** of suggested algorithm

2. Proof-of-concept study:

- The onset time identification of the **chemotherapy-induced neutropenia** after various chemotherapy regimens

Study Schema

Drug Exposure

Drug_exposure_start_date	Drug
2014-08-14	Megestrol Acetate
2014-08-14	calcium polycarbophil
2014-08-14	Lactulose
2014-08-18	Metoclopramide
2014-08-18	Dexamethasone
2014-08-18	Fluorouracil
2014-08-18	Dexamethasone
2014-08-18	calcium polycarbophil
2014-08-18	Oxaliplatin
2014-08-18	Magnesium Oxide
2014-08-18	Leucovorin
2014-08-20	Atropine Sulfate
2014-08-20	calcium polycarbophil

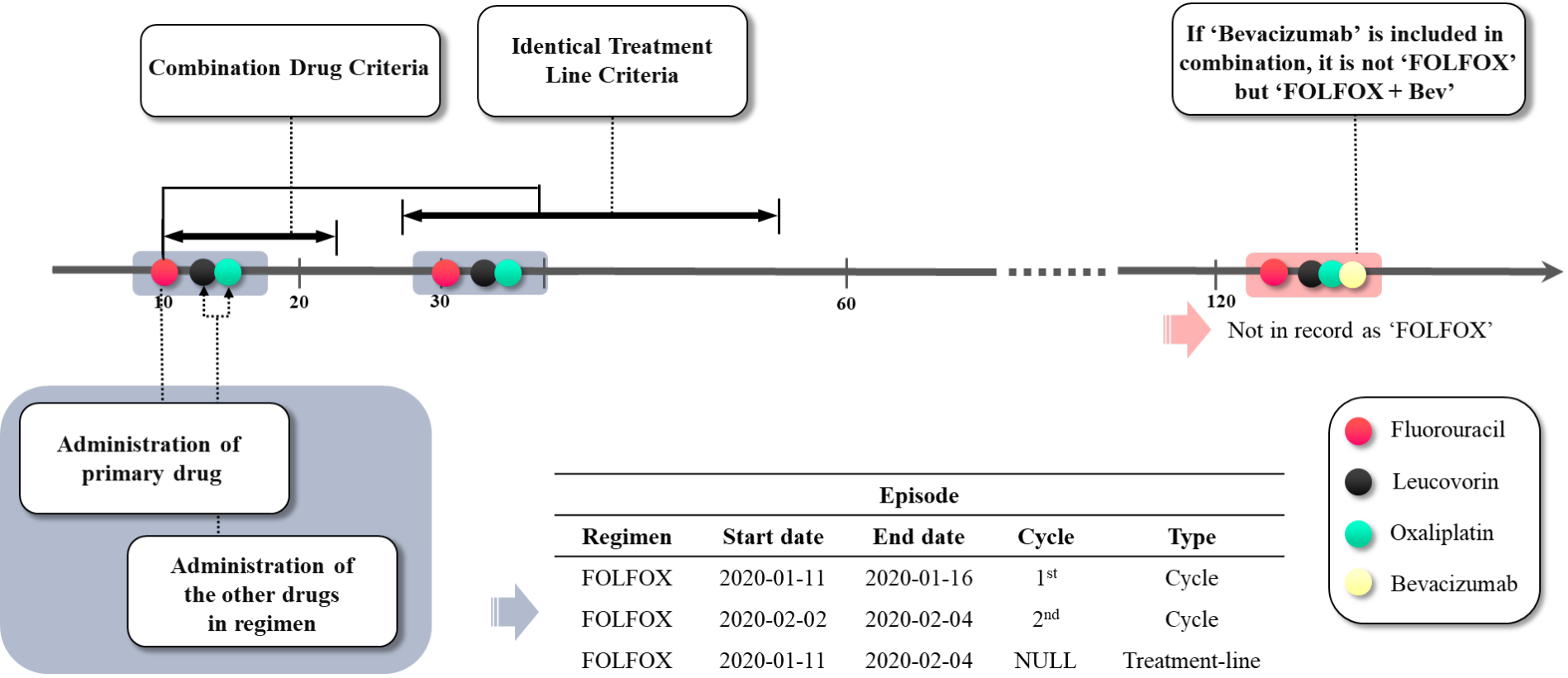
Treatment Episode table

	Episode_source_concept_id	Episode_Number	Episode_start_Datetime	Episode_end_datetime
FOLFOX	35804199	1	2005-10-14	2005-10-14
FOLFOX	35804199	2	2005-11-14	2005-11-14
FOLFOX	35804199	3	2005-12-12	2005-12-12

We generated an **R package for extract chemotherapy regimen records** and generate episode / episode_event table from single medication records in the CDM database.

In this package, we used **HemOnc database** as a reference of the regimen and parameterized the regimen descriptions as a JSON (JavaScript Object Notation) to leverage in the package.

Study Methods



Episode

Regimen	Start date	End date	Cycle	Type
FOLFOX	2020-01-11	2020-01-16	1st	Cycle
FOLFOX	2020-02-02	2020-02-04	2nd	Cycle
FOLFOX	2020-01-11	2020-02-04	NULL	Treatment-line

Based on parameterized JSON regimen descriptions, R package extract chemotherapy records considering:

- After the dispense date of **'primary drug'** indexing **the first date of chemotherapy**, whether the **combination drugs of the regimen are dispensed or not**.
- When each combination of prescribed dates is regard as one cycle, whether the cycles are in the **identical treatment line** or different treatment line.

Study Methods

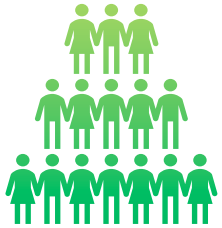
We validated this algorithm by manual reviewing the clinical notes of 300 patients with colorectal cancer.

Targeted regimens were FOLFOX, FOLFIRI, XELODA .

We confirmed an accuracy for :

- The types of regimen are matched to clinical notes.
- The repeated cycle number of each patient is matched to clinical notes.

Study Methods



Malignant Neoplasm Patients
With Chemotherapy
(1994 ~ 2018)

Colorectal Cancer (n = 3,217)
Breast Cancer (n = 5,924)

Visualization

1. The annual usage change of the chemotherapy regimen in each cancer types (2010~2018).
2. The trends of the iteration number for the chemotherapy compared to recommended number of cycle in HemOnc database.
3. Treatment pathway of the cohorts including surgery and chemotherapy.

Study Methods

Chemotherapy treatment

"Neutropenia"

- First incidence of the Grade 1 Chemotherapy Induced Neutropenia

(1500 mm^3 < Absolute Neutrophil Count < 2000 mm^3)

As a proof-of-concept study, we demonstrated the onset timing of the neutropenia.

- First, the number of neutropenia patients on each **cycle number**
- Second, the neutropenia **onset dates** after the first chemotherapy

Study Results

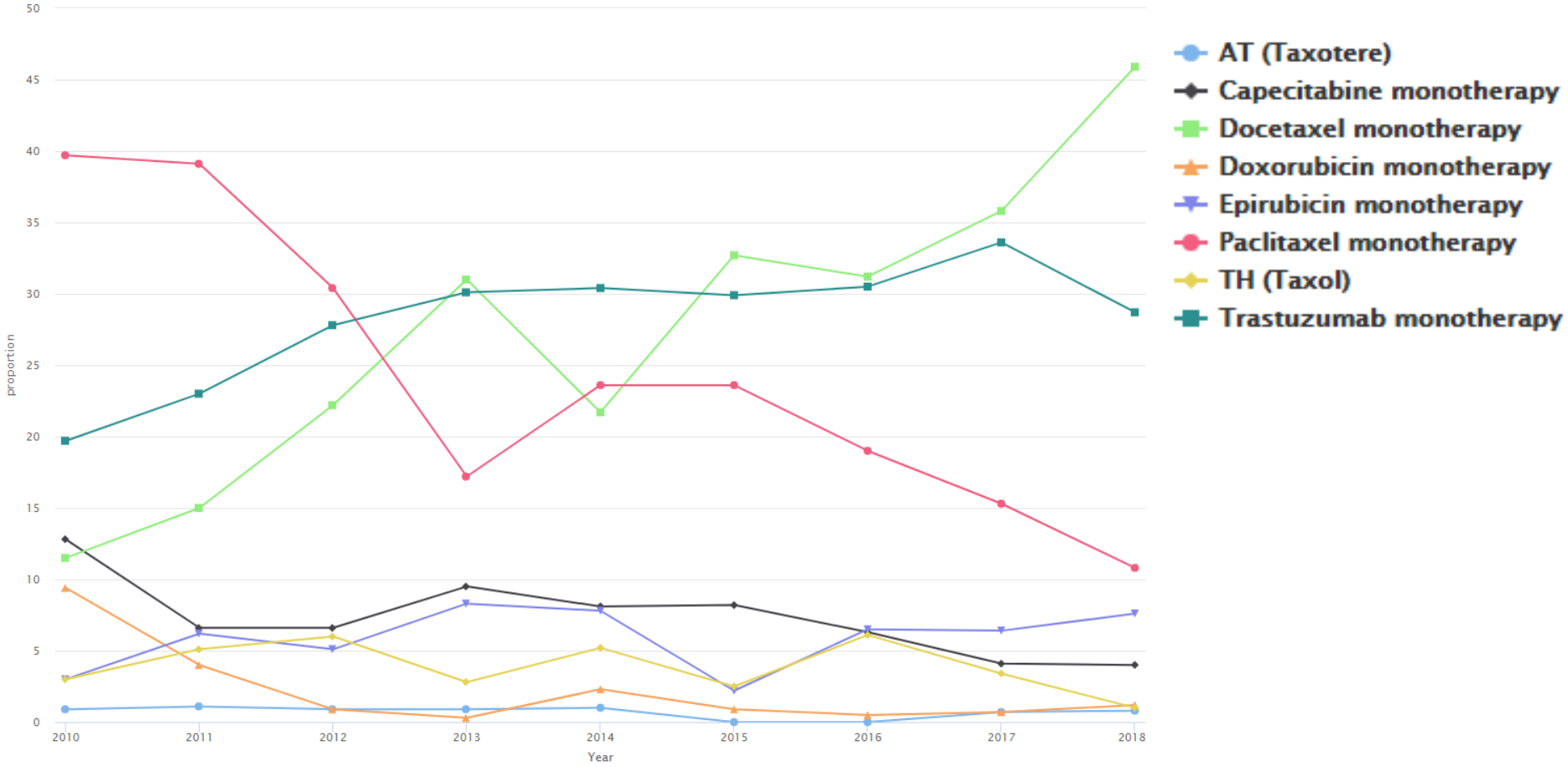
Validation of the chemotherapy regimen extraction accuracy

Treatment Regimen	No Information in Clinical Note	Regimen type Positive Predictive Value	Cycle Number Accuracy	Cycle Difference Average	RMSE
FOLFOX	8	100%	87%	0.3	0.64
FOLFIRI	21	100%	88.60%	0.4	1.36
XELODA	65	100%	69%	0.65	1.5

- In all cases, the extracted chemotherapy episode records were matched in regimen types.
- The repeated cycle number of the patients were matched with note in average 81.5% cases.
- There were only 1 or 2 cycle differences between clinical notes and algorithm results.

Study Results

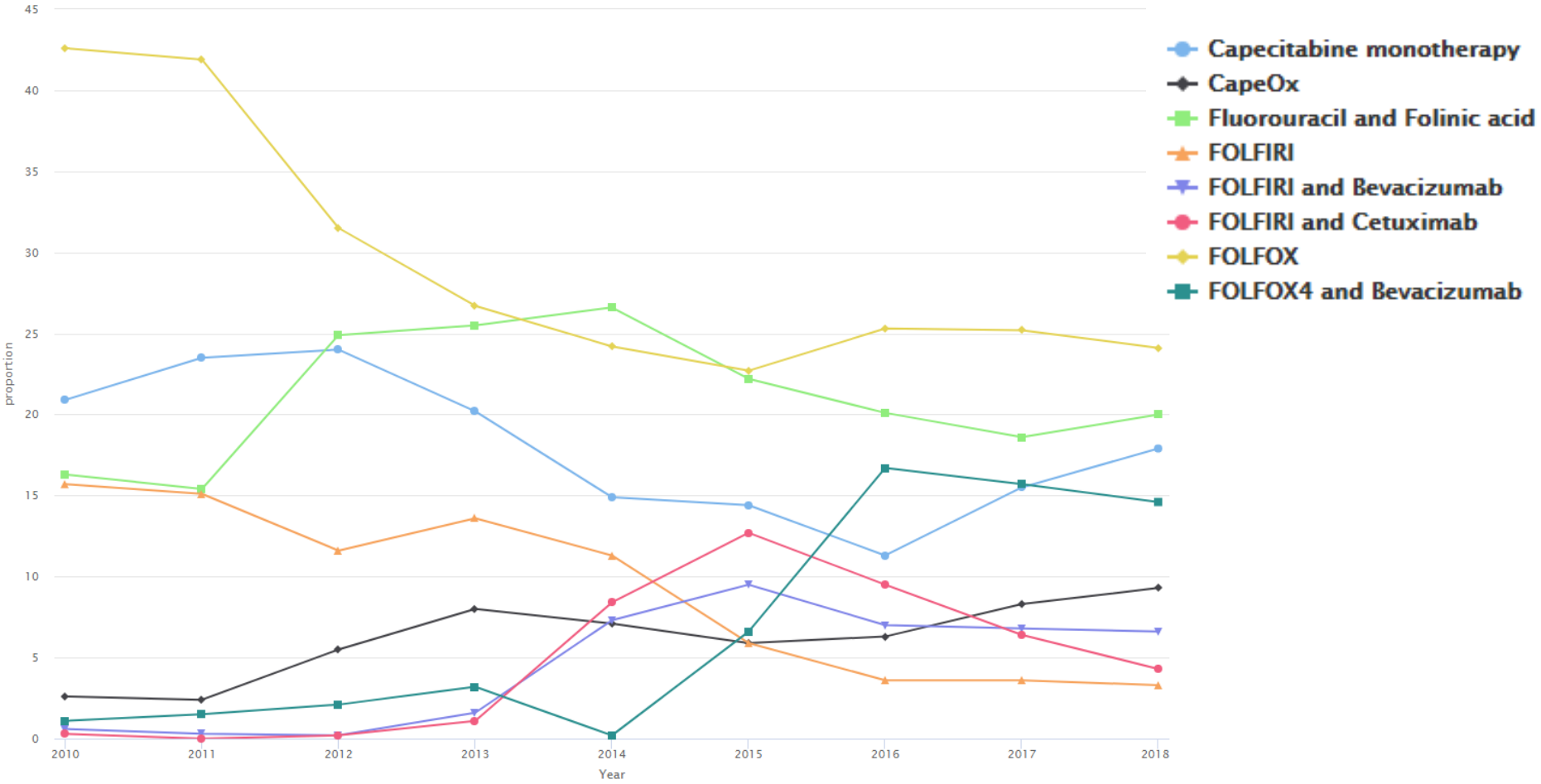
The annual usage proportion change of the chemotherapy regimen in breast cancer



- The graph shows the proportion of each regimen usage for total chemotherapy records.
- The proportion of the Docetaxel monotherapy usage in breast cancer patients is increasing and Paclitaxel monotherapy usage is decreasing

Study Results

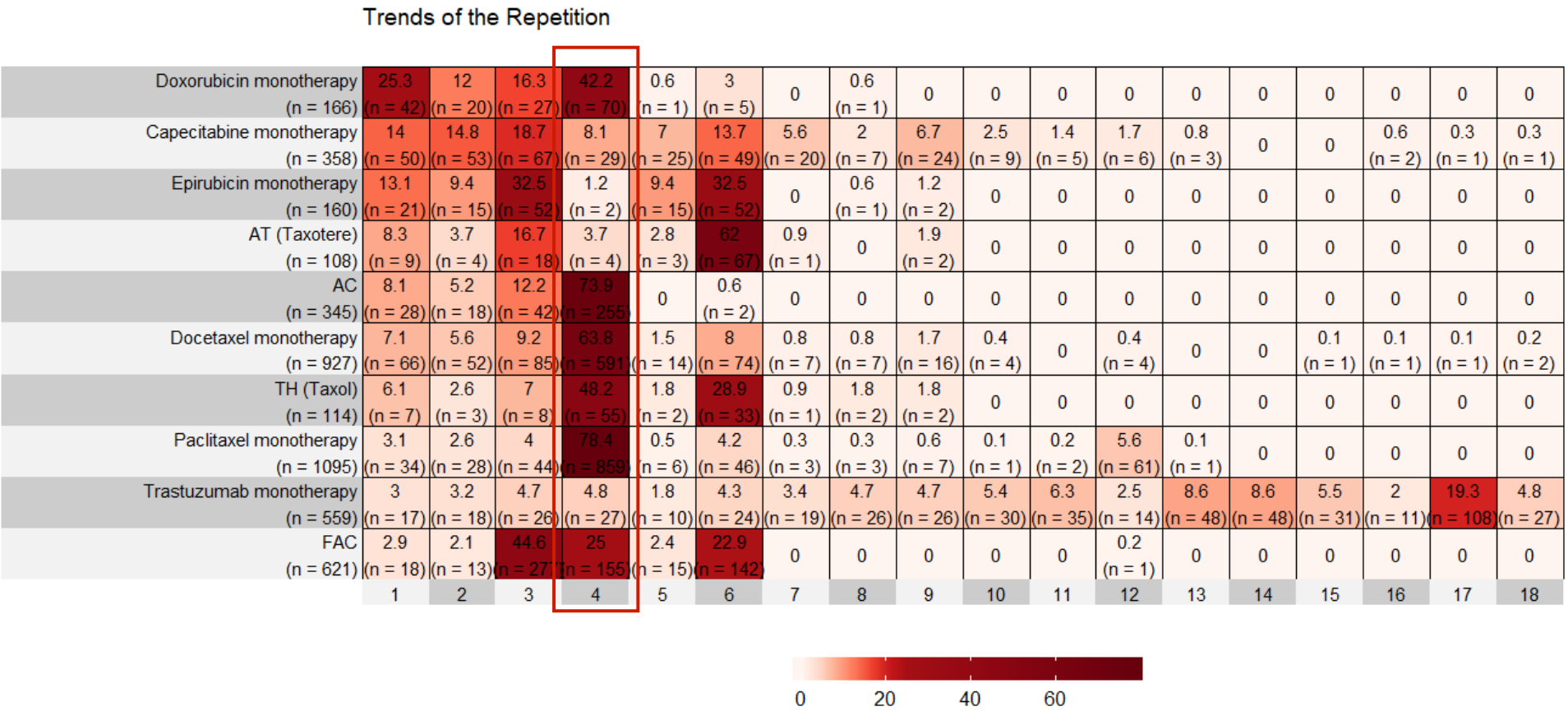
The annual proportion change of the chemotherapy regimen in colorectal cancer



- The FOLFOX is the prevalently used regimen for the most part of the years in colorectal cancer, but the proportion of the regimen including biological agents are increasing gradually.

Study Results

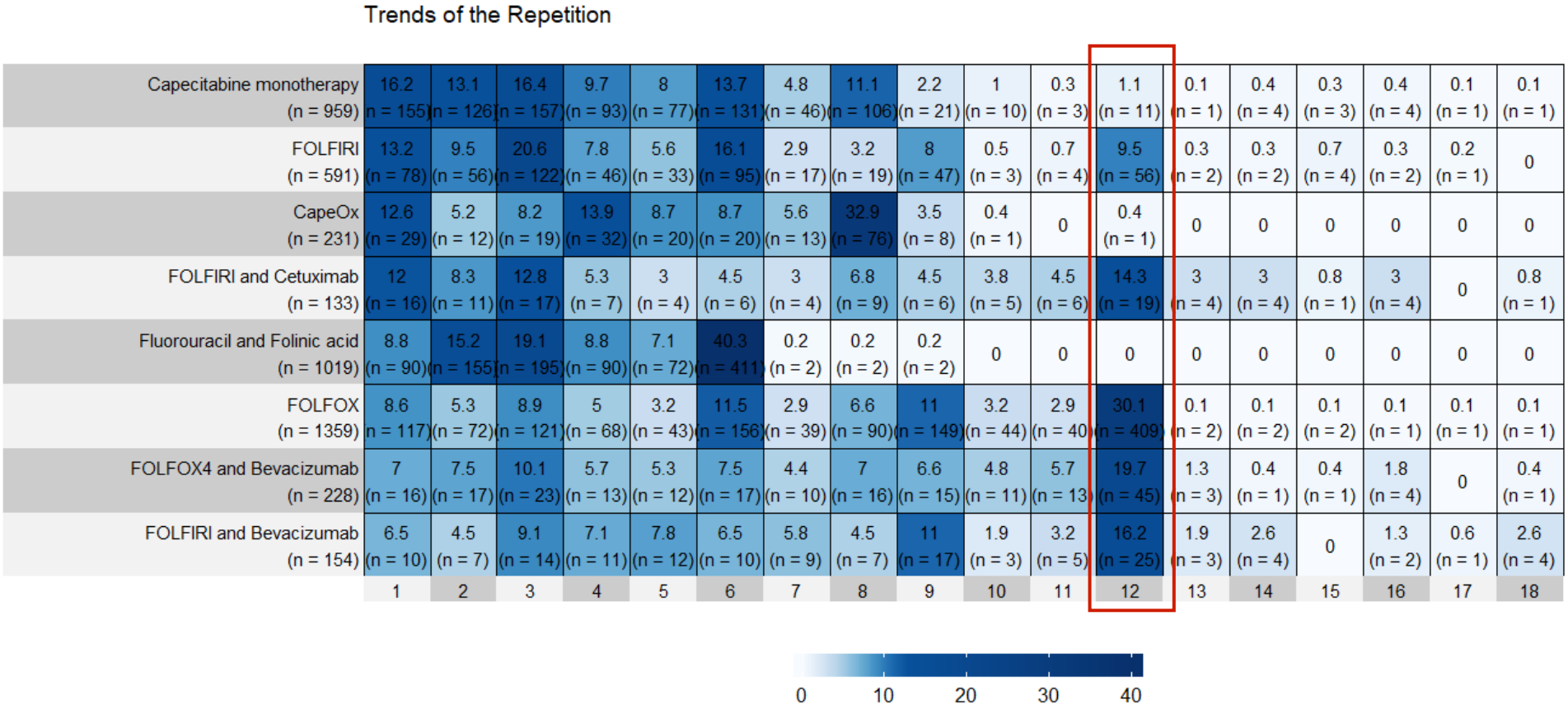
The distribution of the patients in iteration number of breast cancer chemotherapy



- The heatmap shows the number of patients in each treatment iteration counts
- Most of the patients who received Doxorubicin monotherapy, AC and Paclitaxel monotherapy were treated 4 cycles of chemotherapy which is a standard regimen iteration number in HemOnc

Study Results

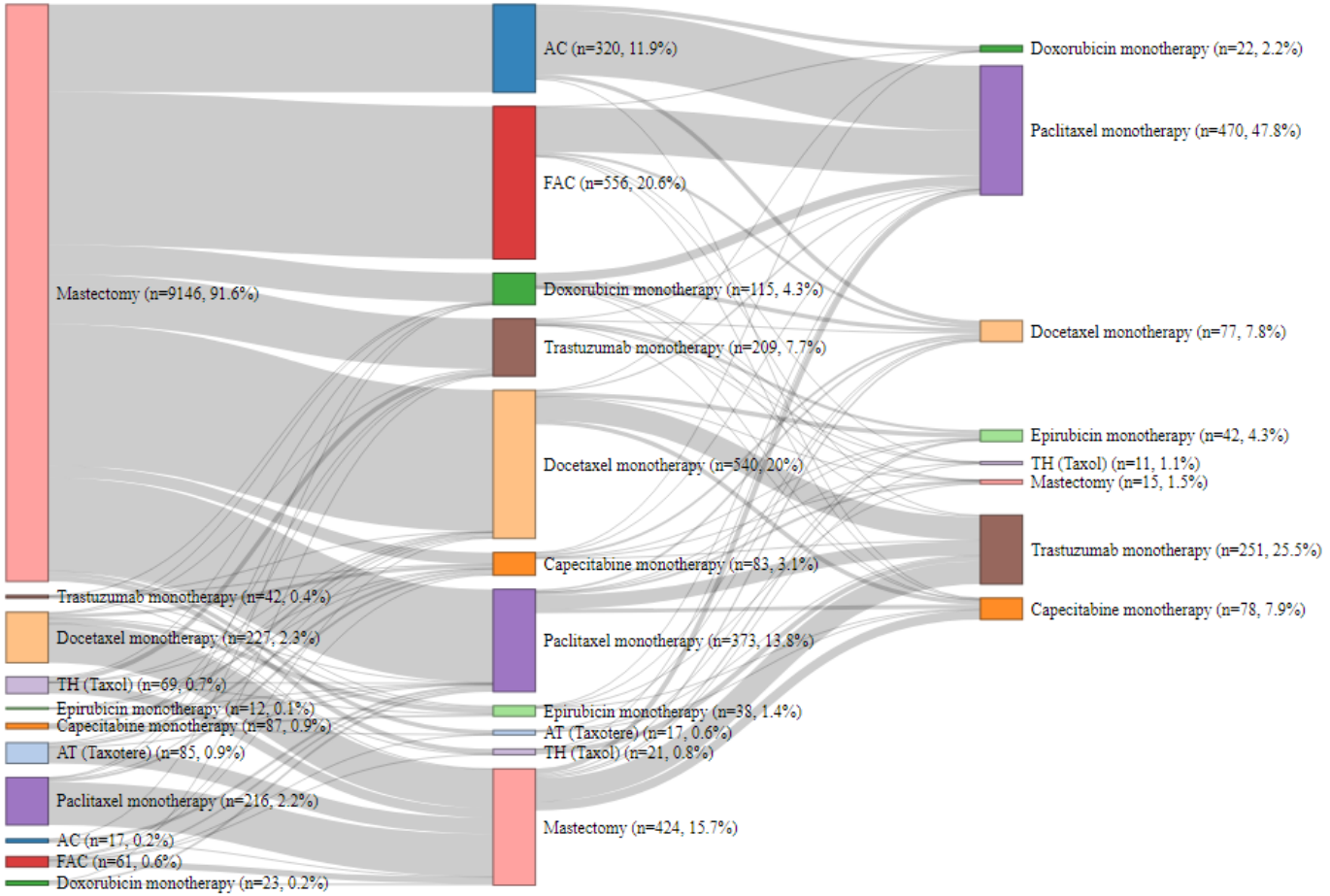
The distribution of the patients in iteration number of colorectal cancer chemotherapy



- Also, the iteration numbers of the colorectal cancer regimen are matched to HemOnc recommended cycle iteration number (FOLFOX and FOLFIRI are recommended to 12 cycles).

Study Results

Treatment patterns in sequential treatment line of the breast cancer

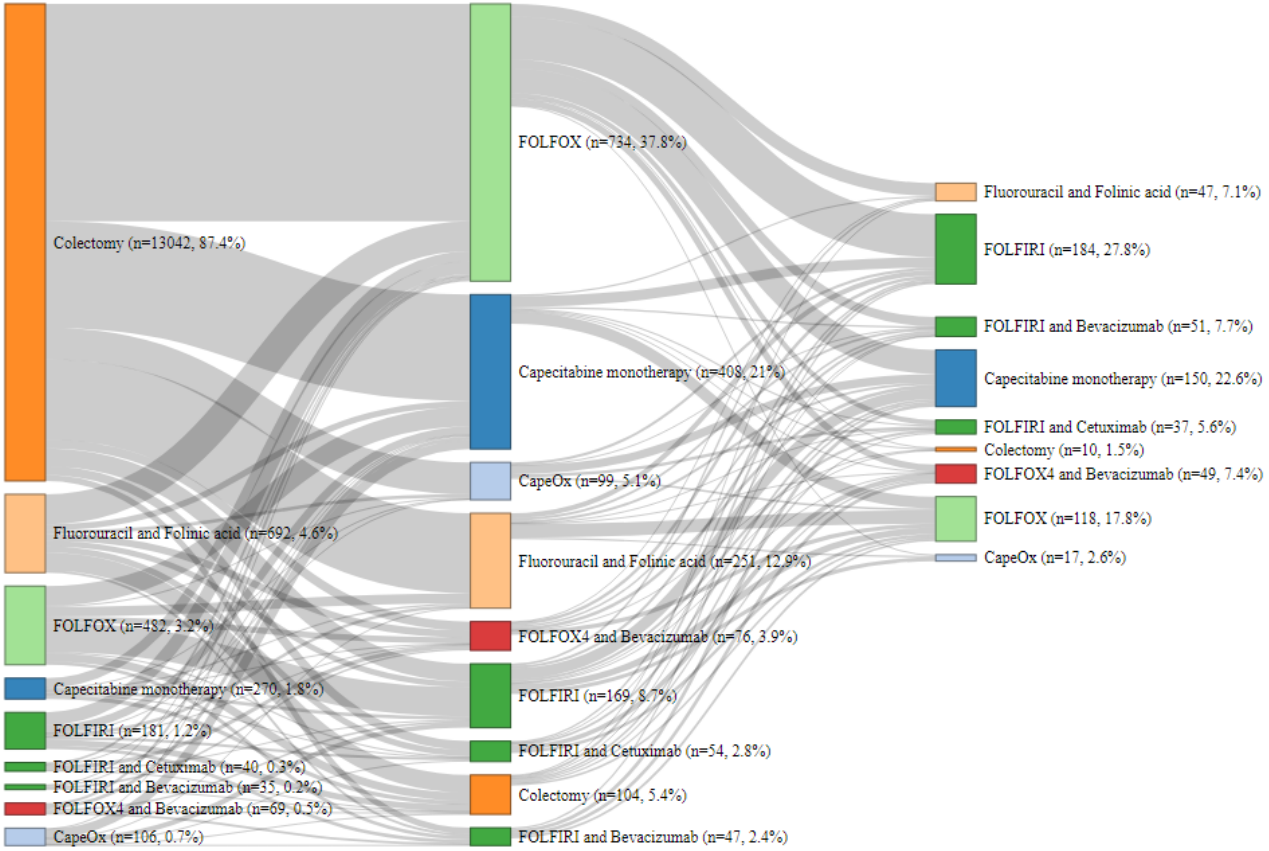


The flow chart shows which regimen is used as a before and after the surgery

- In breast cancer, FAC was frequently used as an **adjuvant chemotherapy**.
- Paclitaxel monotherapy was the most prevalently used **neoadjuvant chemotherapy.**

Study Results

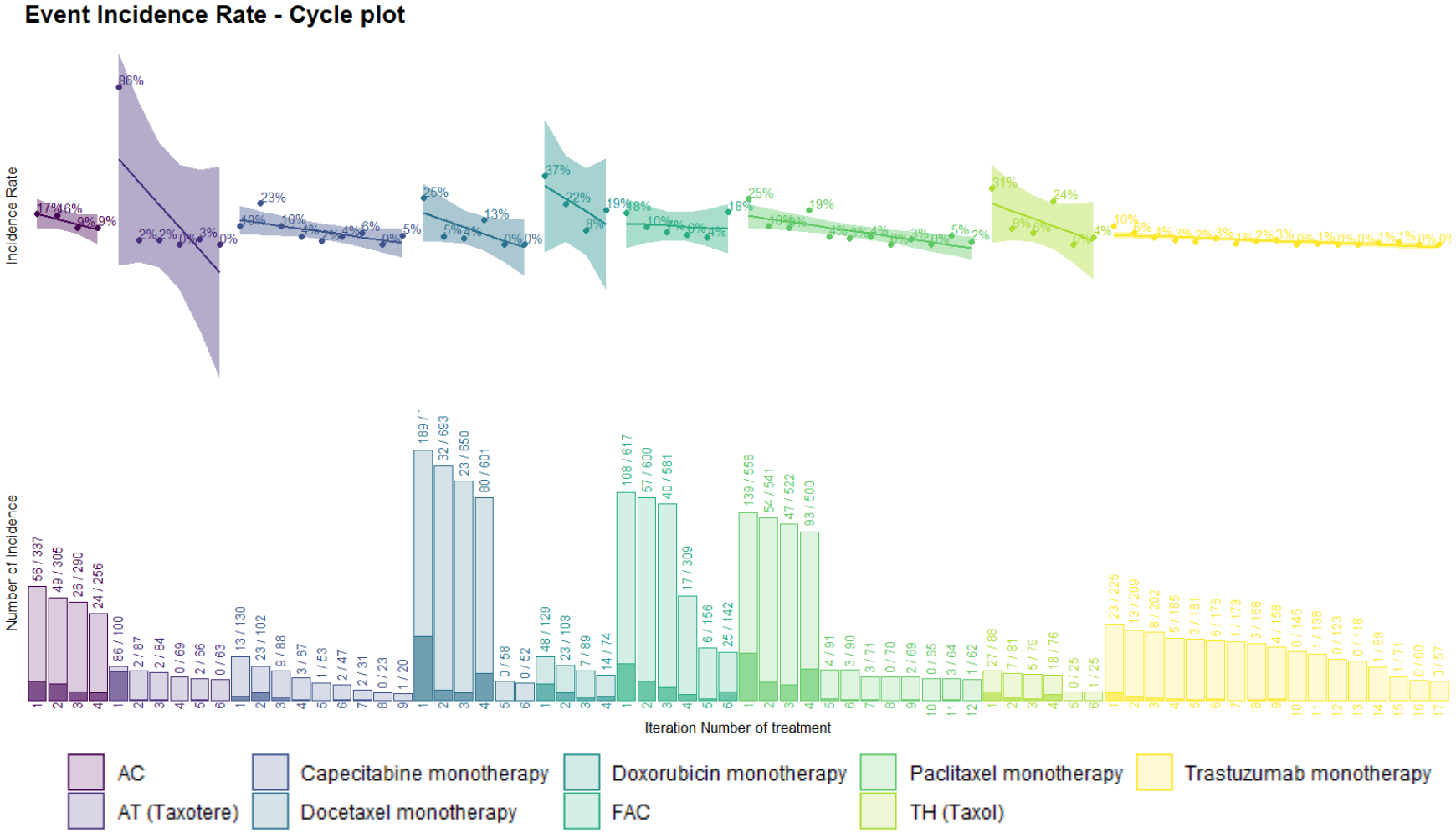
Treatment patterns in sequential treatment line of the colorectal cancer



- In colorectal cancer, the largest portion of the patients were treated FOLFOX regimen after the surgery
- However, patients received **Fluorouracil and Folinic acid** most frequently before the surgery, which is excluded oxaliplatin from FOLFOX regimen.

Study Results

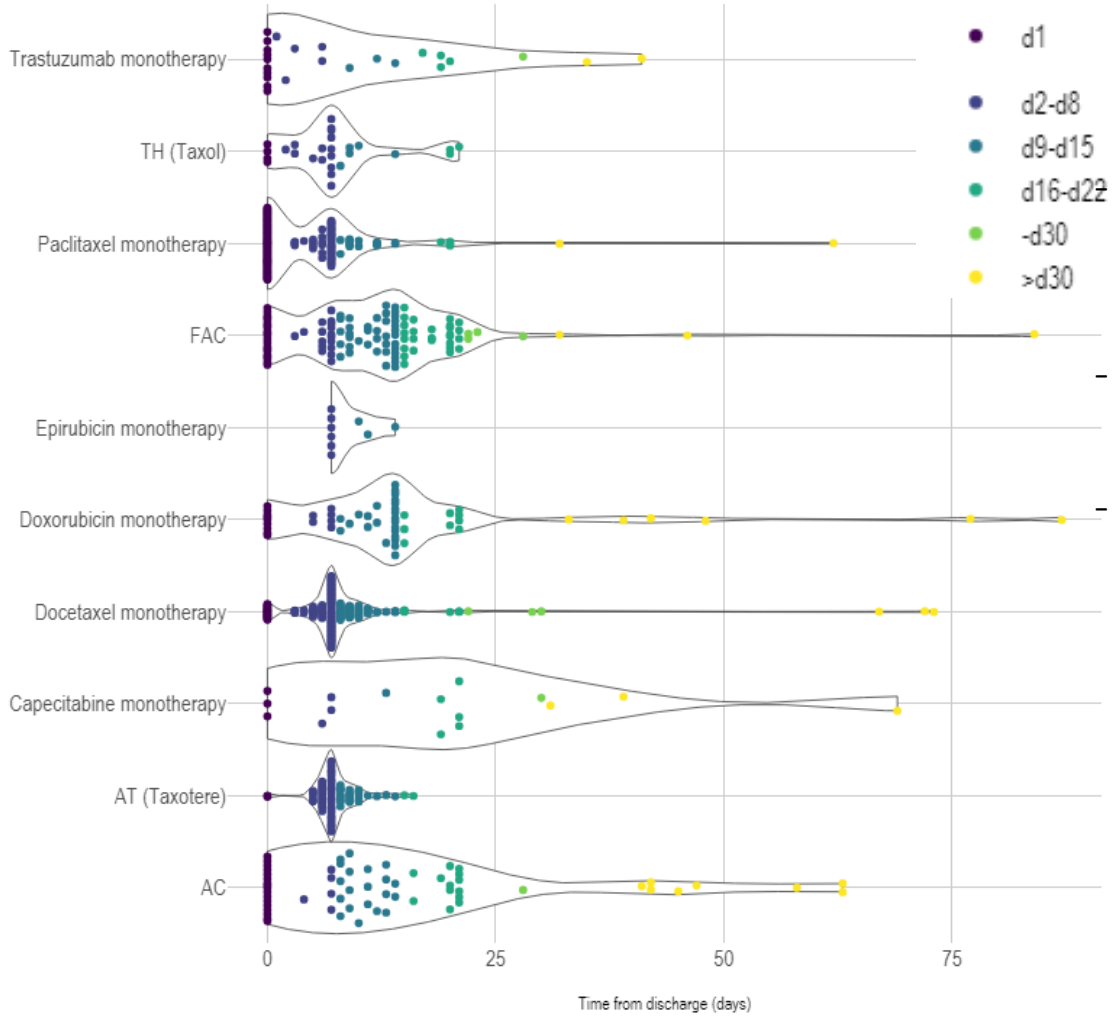
The incidence of first neutropenia in each repeated iteration number of the breast cancer



- The chart shows the distribution of the patients for **first neutropenia onset cycle of the regimen**
- The neutropenia was occurred in large portion after the first AT regimen cycle

Study Results

The distribution of the patients for neutropenia onset date after the first chemotherapy in breast cancer



- The chart shows the distribution of patients for first neutropenia onset date.
- The clustered points of the patients are described as a violin plot.
- Docetaxel and Paclitaxel included chemotherapy (TH, AT, Paclitaxel monotherapy, Docetaxel monotherapy) induced neutropenia not only D1 also D7.

Conclusions

- We suggested the **extraction method for chemotherapy regimen** from Electronic Health Records or the other database which is deficient in chemotherapy regimen data

- Based on extracted chemotherapy regimen data, we demonstrated the **treatment pathway and patterns** of the breast and colorectal cancer

- Also, we suggested the neutropenia onset timing analysis as a proof-of-concept study

- All the codes are uploaded on GitHub page (https://github.com/ABMI/CancerTxPathway/
- https://github.com/ohdsi-studies/CancerTxPathway).

How can you join in?

- Leadership group
  - Overseeing the whole project
  - Leads: current core group (m-gurley@northwestern.edu, rimma.belenkaya@gmail.com)
  - Meetings: monthly

- Research group
  - Defining use cases, cohorts, phenotypes, consulting other groups on clinical/research matters.
  - Leads: Andrew + Christian (awilliams15@tuftsmedicalcenter.org)
  - Meetings: monthly

- CDM/Vocabulary group
  - CDM, vocabulary, ETL conventions issues, Data Quality checks.
  - Leads: Rimma + Dima (rimma.belenkaya@gmail.com)
  - Meetings: weekly

- Development group
  - Implementation of ETL, DQ checks, cohorts, algorithms. Help all participating institutions to implement scripts locally.
  - Leads: Michael + Robert (m-gurley@northwestern.edu or shilpa.ratwani@iqvia.com)
  - Meetings: weekly

- Genomic group
  - Extending CDM/Vocab to support Genomic Oncology use cases
  - Leads: Meera + Chan (PatelM9@mskcc.org)
  - Meetings: weekly

- Outreach group
  - Communicating with standard organizations, collaborators, conferences/publications
  - Leads: Andrew + Christian (awilliams15@tuftsmedicalcenter.org)
  - Meetings: monthly

Thank you for listening !