Challenges to Identification of Cancer Chemotherapy Regimens and Patient Cohorts in Administrative Claims

Catherine M. Lockhart[1], Cara L. McDermott[1], Gabriela Vazquez Benitez[2], Terese DeFor[2], Aaron B. Mendelsohn[3], James Marshall[3], Ali McBride[4], Érick Moyneur[5], Pamala A. Pawloski[2]
On behalf of the BBCIC G-CSF Comparative Effectiveness Research Team

1. Biologics and Biosimilars Collective Intelligence Consortium, Alexandria VA; 2. HealthPartners Institute, Minneapolis, MN; 3. Harvard Pilgrim Healthcare Institute, Boston, MA; 4. University of Arizona, Tucson, AZ; 5. StatLog, Montréal, Canada

BBCIC Biologics & Biosimilars Collective Intelligence Consortium
AMCP BBCIC, LLC

BACKGROUND and OBJECTIVE

Observational research and real-world evidence (RWE) have gained traction in the United States (US) as valuable sources of information for stakeholders and decision-makers in the healthcare system, including clinicians, patients, payers, and regulatory authorities.[1,2] The Biologics and Biosimilars Collective Intelligence Consortium (BBCIC) is a non-profit, multi-stakeholder, scientific collaborative with a mission to generate reliable RWE that examines the safety and effectiveness of biologics in order to improve public health.[3]

Objective: To explore strategies for optimizing identification of patient cohorts and multi-agent chemotherapy regimens using administrative claims data.

METHODS

In a study to compare the safety and effectiveness of granulocyte colony stimulating factor (G-CSF) products (filgrastim, pegfilgrastim) with their biosimilars. Using the BBCIC Distributed Research Network, we included adults aged >=20 years with insurance claims for any G-CSF originator product or biosimilar as febrile neutropenia (FN) prophylaxis following the first cycle of high or intermediate FN risk chemotherapy from 3/1/2015 to 12/31/2019.

We faced challenges in accurately identifying appropriate patient cohorts with cancer diagnoses of interest(lung, breast, colon, ovarian, pancreatic, testicular, cervical, uterine, Non-Hodgkin's lymphoma) receiving multiple-agent chemotherapy regimens. This work demonstrates the challenges and solutions identified for generating appropriate patient cohorts across multiple data sources.

RESULTS

FINDING THE PROBLEM

- Initial inclusion/exclusion criteria resulted in a cohort of only 420 patients (**Table 3**).
- Locating cancer cases followed by chemotherapy exposure, resulted in fewer cases than expected
  - May include people screened for cancer who did not have cancer.

Table 1. Initial Chemotherapy Rate by Diagnosis

Cancer Type	% w/ Chemo	% Expected (Stage I/II only)
Breast	15%	>17%[4]
Cervical	<1%	Need Ref
Colorectal	1%	>9%[5]
Lung	20%	>18%[4]
NHL	16%	>82%[4]
Ovarian	33%	>63%[6]
Pancreatic	<1%	Need Ref
Testicular	13%	>18%[4]
Uterine	<1%	>11%[4]

- Low numbers of patients (12% overall) were initially identified as having received chemotherapy.
- *Rate of chemotherapy treatment was lower than expected for most cancer types compared to published literature and prior experience* (**Table 1**).
- Billing claims for chemotherapy drugs with multiple doses in a cycle rarely occur according to treatment schedule (**Table 2**).
- For example: *98% of claims for 5-FU occurred on Day 1 of treatment despite standard dosing of an infusion on Day 1 and a bolus on Day 2.*

Table 2. 5-fluorouracil (5-FU) by day of claim for colorectal cancer patients receiving FOLFOX regimens.

Day 1 (# claims)	Day 2 (# claims)	% patients
2	0	51%
1	0	47%
1	1	2%

Table 3. Initial Inclusion and Exclusion Criteria to Define the Cohort

	INCL/EXCL Criteria	Description	Count
1	INCL – DIAGNOSIS	First occurrence of 1 inpatient or 2 outpatient diagnoses of cancer of interest ≥30 days apart and within a 12-month period	22,859
2	INCL – CHEMO	Receipt of high- or intermediate-FN risk myelosuppressive chemotherapy regimens with high or intermediate FN risk	2,814
3	INCL – GCSF TREATMENT	G-CSF claim ≤7 days after completion of first chemotherapy administration	1,318
4	INCL - ENROLLMENT	≥365 days of health plan enrollment (≤45-day gap allowed prior to Index), and ≥90 days continuous enrollment after Index (or ≥30 days after G-CSF claim)	843
5	INCL – AGE 20+	Age ≥20 years Index	842
6	EXCL – PRIOR CANCER	1 inpatient or 2 outpatient cancer diagnoses ≥30 days apart ≤183 days prior to Index for cancers that differ from the enrolling diagnosis	781
7	EXCL – PRIOR CHEMO	Receipt of any chemotherapy ≤183 days prior to Index	631
8	EXCL – PRIOR GCSF	Any G-CSF claim ≤183 days prior to Index	610
9	EXCL – PRIOR SNF/HOSPICE	≥2 medical claims ≥30 days apart for SNF or hospice care ≤183 days prior to Index	610
10	EXCL – CLINICAL EXCLUSIONS	≥2 medical claims ≥1 day apart for cancer-related radiotherapy, bone marrow or stem cell transplant, presence of indicators of metastatic disease, receipt of bone-targeted agents, diagnosis of HIV/AIDS, severe hepatic disease, chronic kidney disease, or any non-oncology related neutropenia ≤183 days prior to Index	420

A series of problem-solving revisions led to a robust algorithm to identify and classify patient cohorts (**Table 4**)

Table 4. Change in cohort size after algorithm revisions.

	INCL/EXCL Description	Rev 1	Rev 2	Rev 3	Rev 4
1	INCL – CHEMO	5,320	5,182	13,600	13,600
2	INCL – DIAGNOSIS	3,645	3,396	6,010	6,010
3	INCL – GCSF TREATMENT	1,501	1,476	1,779	1,801
4	INCL – ENROLLMENT	931	916	1,093	1,116
5	INCL – AGE 20+	930	915	1,086	1,109
	EXCL – PRIOR RADIATION	Not Excluded	Not Excluded	1,081	1,068 (1 code)
6	EXCL – PRIOR CANCER	853	Not Excluded	Not Excluded	Not Excluded
7	EXCL – PRIOR CHEMO	702	727	Not Excluded	Not Excluded
8	EXCL – PRIOR GCSF	668	689	962	958
9	EXCL – PRIOR SNF/HOSPICE	668	689	962	958
10	EXCL – CLINICAL EXCLUSIONS	449	462	620	886
	FINAL COHORT	**449**	**462**	**620**	**886**

Index Date: First chemotherapy administration during the study period

Revision 1:
- Reverse order of INCL - DIAGNOSIS and INCL - CHEMO
- EXCL – PRIOR CHEMO: Ignore prednisone, methotrexate
- Patients receiving ≥1 day of chemo assumed to receive entire regimen
- Remove cisplatin + vinorelbine regimen in lung cancer
- Remove BEP regimen in testicular cancer

Revision 2:
- INCL – GCSF TREATMENT: claims -7 days ≤ Index ≥ +7 days
- EXCL – PRIOR CANCER: remove
- Categorize according to latest cancer diagnosis (closest to regimen) instead of first (earliest) diagnosis

Revision 3:
- INCL 1: Use all chemo records within 5-day window to classify regimen
- EXCL 7: remove
- ADD: EXCL prior radiation
- Include paclitaxel regimens for lung cancer

Revision 4:
- EXCL 10: do not exclude metastatic disease or bone-targeting agents ≤183 days prior to Index. Manually exclude later.

SOLVING THE PROBLEM

	Description	Rationale
Revision 1	Reverse order of first two inclusion criteria	• Eliminated patients screened but who do not have cancer • Retained patients with initial cancer diagnosis outside 12-month look-back period
	Chemotherapy claim on ≥1 day assumed to receive all doses	• Timing of claims do not necessarily reflect clinical care (e.g., colorectal patients with claim for 5-FU on ≥1 day assumed to have received 2-day regimen in FOLFOX (Table 3))
	Oral predisone and methotrexate not used for case identification or exclusion	• Multi-indication treatments that did not influence FN risk were not included • Patients may fill oral prescription days to weeks prior to infusion
	Remove cisplatin+vinorelbine in lung cancer	• Removed - not commonly used in clinical practice per clinician advisement
	Remove BEP regimen in testicular cancer	• Removed - potential interaction due to bleomycin and G-CSF per clinician advisement
Revision 2	G-CSF treatment window	• Patients may fill G-CSF prescriptions separately, so allowing 7 days before and after infusion allows for more accurate capture of G-CSF use
	Prior cancers exclusion	• This exclusion was unnecessarily narrow and eliminated patients even with early-stage or other cancers without G-CSF or chemotherapy exposure
	Categorize on most recent diagnosis	• To improve the likelihood that patients are placed in the disease cohort that most likely reflects their chemotherapy and G-CSF use
Revision 3	5-day window to identify chemotherapy drugs	• More accurately capture chemotherapy use, accounting for differences in timing of claims vs. clinical care
	Prior chemotherapy exclusion	• Exclusions done after initial data extraction
	Separate exclusion for prior radiation therapy	• Moved to earlier in exclusion process to streamline data extraction
	Include paclitaxel regimens for lung cancer	• Were not always identified; the issue was resolved when the 5-day window for regimen identification was implemented
Revision 4	Metastatic disease or bone-targeting agent exclusion ≤183 days	• Exclusions done after initial data extraction

LESSONS LEARNED

- Administrative data do not always reflect the relative timing of clinical care. It is important to identify those areas to develop solutions, for example, to accurately identify multi-drug treatment regimens.
  - Relative timing of medication exposure must be evaluated to appropriately classify treatment cohorts.
  - Appropriate treatment window for claims to be recorded.
- Medications used for multiple indications may confound, or at least complicate, identification of treatment regimens.
- Coding errors, or use of non-specific codes in claims is a risk that should be quantified.
- The order of inclusion and exclusion criteria can make a difference in identifying patients.
- Great care must be taken in designing clinically-meaningful inclusion and exclusion criteria.
- Development and testing of the algorithm prior to data extraction across multiple sites was instrumental in identifying and resolving data collection challenges.
- Ongoing iterative review between programmers and investigators was critical to algorithm development

ACKNOWLEDGEMENTS

BBCIC G-CSF Research Team:
Heidi Bailly, HealthPartners
Jaclyn Bosco, IQVIA
Maria Bottorff, HOPA
Kimberly Daniels, HealthCore
Audrey Djibo, CVS CTS
Aziza Jamal-Allial, HealthCore
Christopher Kim, Amgen
Annemarie Kline, CVS CTS
Edward Li, Sandoz
Sam Li, Univ Tennessee
Nancy Lin, IQVIA
Ali McBride, Univ. Arizona
Jennifer Schenfeld, Amgen
Richard Thielen, AbbVie
Cheryl Walraven, CVS CTS
Gary Yee, UNMC

Special thanks also to all BBCIC Sponsors and Research partners: AbbVie, Amgen, Boehringer-Ingelheim, CVS Health Clinical Trial Services (CVS CTS), Harvard Pilgrim Health Care Institute, HealthCore, HealthPartners Institute, Kaiser Permanente Washington Research Institute, Sandoz

REFERENCES

1. Beaulieu-Jones BK, Finlayson SG, Yuan W, et al. Examining the Use of Real-World Evidence in the Regulatory Process. Clin Pharmacol Ther 2020;107:843-52.
2. Blonde L, Khunti K, Harris SB, Meizinger C, Skolnik NS. Interpretation and Impact of Real-World Clinical Data for the Practicing Clinician. Adv Ther 2018;35:1763-74.
3. Biologics and Biosimilars Collective Intelligence Consortium. (Accessed May 14, 2021, at www.bbcic.org.)
4. Miller et al. Cancer Treatment & Survivorship Facts & Figures 2019-2021. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-treatment-and-survivorship-facts-and-figures/cancer-treatment-and-survivorship-facts-and-figures-2019-2021.pdf
5. Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin 2019;69:363-85
6. Cancer Trends Progress Report. National Cancer Institute, NIH, DHHS, Bethesda, MD, July 2021. https://progressreport.cancer.gov/ (Accessed 4/11/2022)

This study was supported by the Biologics and Biosimilars Collective Intelligence Consortium