

1 **Title:** Determining Line of Therapy from Real-World Data in Non-Small Cell Lung Cancer

2

3 **Authors:** Connor B. Grady¹, MPH; Wei-Ting Hwang¹, PhD; Joshua E. Reuss², MD; Wade
4 Iams³, MD, MSCI; Amanda Cass³, PharmD; Geoffrey Liu⁴, MD; Devalben Patel⁴, MD; Stephen
5 V. Liu², MD; Gabriela Liliana Bravo Montenegro², MD; Tejas Patil⁵, MD; Jorge J. Nieva⁶, MD;
6 Amanda Herrmann⁶, MD; Kristen A. Marrone⁷, MD; Vincent K. Lam⁷, MD; William
7 Schwartzman⁸, MD; Jonathan Dowell⁸, MD; Liza C Villaruz⁹, MD; Kelsey Leigh Miller¹⁰, MD;
8 Jared Weiss¹⁰, MD; Fangdi Sun¹¹, MD; Vamsidhar Velcheti, MD¹²; D. Ross Camidge⁵, MD,
9 PhD; Charu Aggarwal¹³, MD, MPH; Lova Sun¹³, MD, MSCE; Melina E. Marmarelis¹³, MD,
10 MSCE

11

12 **Affiliations:**

13 ¹ Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, 423
14 Guardian Drive, Philadelphia, PA 19104

15 ² Department of Medicine Georgetown University, 3800 Reservoir Road NW, Washington, DC
16 20007

17 ³ Division of Hematology/Oncology, Department of Medicine, Vanderbilt University Medical
18 Center, 2220 Pierce Ave, 780 PRB, Nashville, TN 37232

19 ⁴ Princess Margaret Cancer Centre, 610 University Ave, Toronto, ON M5G 2C4, Canada

20 ⁵ University of Colorado Cancer Center, 1665 Aurora Court, MS F704, Aurora, CO 80045

21 ⁶ University of Southern California/Norris Comprehensive Cancer Center, 1441 Eastlake Avenue
22 NOR 3400, Los Angeles, CA 90033

23 ⁷ Johns Hopkins University School of Medicine, 300 Mason Lord Drive, Baltimore, MD 21224

24 ⁸ Harold C Simmons Comprehensive Cancer Center, UT Southwestern, 6202 Harry Hines, Blvd,
25 9th Floor, Dallas Texas, 75235

26 ⁹ UPMC Hillman Cancer Center, 5150 Centre Avenue, 5th Floor, Pittsburgh, PA 15232

27 ¹⁰ Lineberger Comprehensive Cancer Center, 170 Manning Drive, Room 3115, Chapel Hill, NC
28 27599

29 ¹¹ UCSF School of Medicine, University of California San Francisco, 533 Parnassus Ave, San
30 Francisco, CA 94143

31 ¹² NYU Grossman School of Medicine, 550 1st Ave, New York, NY 10016

32 ¹³ Department of Medicine, Division of Hematology and Oncology, Perelman School of
33 Medicine, University of Pennsylvania, 3400 Civic Center Blvd, 10th Floor, South Tower,
34 Philadelphia, PA 19104.

35

36 **Corresponding author:**

37 Melina E. Marmarelis, MD, MSCE

38 Department of Medicine, Division of Hematology & Oncology

39 Perelman School of Medicine

40 3400 Civic Center Blvd, 10th Floor, South Tower

41 Philadelphia, PA 19104

42 Email: Melina.Marmarelis@pennmedicine.upenn.edu

43

44 **Conflict of Interest Statements:**

45 CBG, WH, AC, GL, DP, GLBM, AH, WS, JD, KLM, VV, DRC: No conflicts of interest
46 associated with current work.

47 JER: Advisory Board / Consultant: Genentech/Roche, Sanofi/Genzyme, Personalis, Guardant,

48 AstraZeneca, BMS, Arcus, Abbvie, Daiichi Sankyo. Research Funding (to institution):

49 Genentech/Roche, Verastem, Nuvalent, Mesothelioma Applied Research Foundation,

50 LUNGeivity Foundation. Honoraria/Speaking Fees: AstraZeneca, Merck

51 WI: Consultant for OncLive, Clinical Care Options, Chardan, Cello Health, Curio Science.

52 Advisory Board/DSMC for Genentech, Mirati, Outcomes Insights, Jazz Pharmaceuticals, GI

53 Therapeutics, Takeda, AstraZeneca, Sanofi, Janssen, Amgen, Bristol Myers Squibb, NovoCure.

54 SVL: Advisory Board / Consultant for Abbvie, Amgen, AstraZeneca, Boehringer Ingelheim,

55 Bristol-Myers Squibb, Catalyst, Daiichi Sankyo, Eisai, Elevation Oncology, Genentech/Roche,

56 Gilead, Guardant Health, Janssen, Jazz Pharmaceuticals, Merck, Merus, Mirati, Novartis,

57 Regeneron, Sanofi, Takeda, Turning Point Therapeutics. Research grant (to institution)

58 from Alkermes, Elevation Oncology, Genentech, Gilead, Merck, Merus, Nuvalent, RAPT,

59 Turning Point Therapeutics. Data Safety Monitoring Board for Candel Therapeutics.

60 TP: Advisory Role (advisory boards or consultations) in last 3 years: AstraZeneca, Biocept,

61 Boehringer Ingelheim, Bristol-Myers Squibb, Bicara, Caris, Guardant Health, Guidepoint, EMD

62 Soreno, Janssen, Jazz Pharmaceuticals, Mirati Therapeutics, Natera, Pfizer, Sanofi, Regeneron,

63 Roche/Genentech, Takeda. Advisory Committees: Elevation Oncology (DSMB). Research
64 Funding: EMD Soreno, Janssen, Gilead
65 JN: Stock/Ownership interests in Epic Sciences, Cansera, Quantagene, Indee P/L. Consulting for
66 AstraZeneca, Naveris, Aadi biosciences, Bioatla, Mindmed, ANP Technologies. Research
67 funding to institution from Merck, Genentech. Patent pending on movement and unexpected
68 healthcare encounters.
69 KAM has received consulting/advisory fees from AstraZeneca, Amgen, Janssen, Mirati
70 Therapeutics, Daiichi Sankyo/Lilly and Puma Biotechnology, as well as Honoraria from
71 AstraZeneca. KAM receives research funding to Johns Hopkins University from Bristol-Myers
72 Squibb and Mirati Therapeutics.
73 V.K.L. reports serving in a consultant/advisory role for Seattle Genetics, Bristol-Myers Squibb,
74 AstraZeneca, Guardant Health, Takeda, and Anheart Therapeutics, and has received research
75 funding from BMS, Merck, Seattle Genetics, Astra Zeneca
76 LCV reports researching funding from Janssen, BMS, Merck, Regeneron, GSK, AstraZeneca,
77 BioAtla, Black Daimon Therapeutics, Jazz, Genentech, Beigene. Consulting fees from Takeda,
78 Janssen, Intervenn Biosciences, Sanofi, Daiichi Sankyo, Jazz, BMS, Gilead.
79 CA reports receiving institutional research funding from AstraZeneca, Genentech, Incyte,
80 Macrogenics, Medimmune, and Merck Sharp & Dohme, and receiving consultation fees from
81 Genentech, Lilly, Celgene Merck, AstraZeneca, Blueprint Genetics, Shionogi, Daiichi Sankyo/
82 Astra Zeneca, Regeneron/ Sanofi, Eisai, BeiGene, Turning Point, Pfizer, Janssen, Boehringer
83 Ingelheim.
84 LS reports research funding from Blueprint (Inst), Seagen (Inst), IO Biotech (Inst), Erasca (Inst).
85 Consulting with Sanofi Genzyme, Regeneron, GenMab, Seagen, and Bayer.
86 MEM reports researching funding from Eli Lilly (Inst), AstraZeneca (Inst), Merck (Inst),
87 Genentech (Inst) consulting role with Astra Zeneca, Novocure, Boehringer Ingelheim, Janssen,
88 Takeda, Blueprint Pharmaceuticals, Bayer, Bristol Myers Squibb, Ikena; Honorarium from
89 Thermo Fisher; stock in Merck, Johnson & Johnson.

90

91 **Funding:** This work is supported in part by grants from Takeda, AstraZeneca and LUNgevity.

92 **Keywords:** Non-small cell lung cancer, EGFR, ALK, ROS1, real-world data

93 **Word Count:** 1196

94 **Abstract**

95 **Introduction:** Determining lines of therapy (LOT) using real-world data is crucial to inform
96 clinical decisions and support clinical research. Existing rules for determining LOT in patients
97 with metastatic non-small cell lung cancer (mNSCLC) do not incorporate the growing number of
98 targeted therapies used in treatment today. Therefore, we propose rules for determining LOT from
99 real-world data of patients with mNSCLC treated with targeted therapies.

100 **Methods:** LOT rules were developed through expert consensus using a real-world cohort of 550
101 patients with *ALK*⁺ or *ROS1*⁺ mNSCLC in the multi-institutional, electronic medical record-based
102 Academic Thoracic Oncology Medical Investigators Consortium's (ATOMIC) Driver Mutation
103 Registry. Rules were subsequently modified based on a review of appropriate LOT determination.
104 These resulting rules were then applied to an independent cohort of patients with *EGFR*⁺
105 mNSCLC to illustrate their use.

106 **Results:** Six rules for determining LOTs were developed. Among 1133 patients with *EGFR*
107 mutations and mNSCLC, a total of 3168 regimens were recorded with a median of 2 regimens per
108 patient (IQR, 1-4; range, 1-13). After applying our rules, there were 2834 total LOTs with a median
109 of 2 LOTs per patient (IQR, 1-3; range, 1-11). Rules 1-3 kept 11% of regimen changes from
110 advancing the LOT. When compared to previously published rules, LOT assignments differed
111 5.7% of the time, mostly in LOTs with targeted therapy.

112 **Conclusion:** These rules provide an updated framework to evaluate current treatment patterns,
113 accounting for the increased use of targeted therapies in patients with mNSCLC and promote
114 standardization of methods for determining LOT from real-world data.

115

116 **Key Points**

- 117 • Use of targeted therapy to treat patients with mNSCLC is growing
- 118 • Determining lines of therapy from real-world data is crucial for clinical research
- 119 • Our rules aim to advance the line of therapy to reflect changes in clinical status
- 120 • Using these rules can lead to better method harmonization in mNSCLC research

121 **Introduction**

122 When using real-world data to answer clinical questions in oncology, accurate
123 determination of lines of therapy (LOT) using treatment records is crucial to conducting
124 meaningful analyses. In the absence of detailed information on progression and toxicity,
125 investigators rely on treatment records alone to infer changes in clinical disease status and
126 determine LOTs. This non-uniform methodology may cause significant variation in how LOTs are
127 defined depending on malignancy type, data source, and treatment patterns. Standardized rules and
128 algorithms for determining LOTs are needed to improve the reproducibility and comparability of
129 real-world data analyses.

130 There is tremendous opportunity for impactful analysis of real-world data from patients
131 with metastatic non-small cell lung cancer (mNSCLC), but standardized LOT determination has
132 been difficult. Systemic therapies for mNSCLC now include immunotherapy, chemotherapy, and
133 targeted therapies, which can be used alone or in combination. Because of this, treatment patterns
134 and sequencing are complex and heterogeneous¹. This is particularly true for the subset of patients
135 with mNSCLC and rare molecular characteristics and patients with co-morbidities where complex
136 treatment patterns and outcomes are only captured outside of clinical trials.

137 Previously, others have proposed algorithms for determining LOTs for patients with solid
138 tumors, lung cancer, and mNSCLC.² However, these algorithms do not specifically incorporate
139 common treatment patterns for patients with targetable alterations, such as *EGFR*, *ALK*, and *ROS1*
140 alterations. In this report, we aim to propose new rules for determining LOT from real-world data
141 of patients with mNSCLC, focusing on treatment patterns among patients receiving targeted
142 therapies.

144 **Methods**

145 We defined a regimen as any systemic anticancer agent (i.e., chemotherapy,
146 immunotherapy, or targeted therapy) or combination of systemic agents administered at the same
147 time. First, we developed rules for determining LOT (“LOT rules”) by expert consensus and
148 review of existing literature. Next, these rules were applied to treatment records from 550 patients
149 with mNSCLC and sensitizing *ALK* or *ROS1* fusions in the ATOMIC Driver Mutation Registry.
150 ATOMIC is a multi-institutional collaborative of twelve academic centers across North America.
151 Trained data abstractors used a standardized digital form to record patients’ regimen agent(s) and

152 start/end dates. Finally, thoracic oncologist members of ATOMIC reviewed the LOT
153 determination and revised the rules to reflect practice patterns most indicative of changes to clinical
154 disease status. These modified rules were then applied to a separate ATOMIC cohort of 1,133
155 patients with mNSCLC and *EGFR* sensitizing mutations. Rules were also developed to identify
156 LOTs that are most likely applicable to early-stage disease (“ESD rules”) since data sources of
157 patients with mNSCLC may inadvertently include treatments for this disease stage.

158 In this report, we demonstrate the use of LOT and ESD rules on treatment records from the
159 ATOMIC *EGFR* cohort. We compared our LOT determination with those of an existing algorithm
160 for mNSCLC proposed by Hess et al using the SAS macro %mnsclc_lot². Analyses were
161 performed in R version 4.2.1³ and SAS software version 9.4. This study was approved by the
162 University of Pennsylvania Institutional Review Board (IRB #829009).

163

164 **Results**

165 We propose six rules for determining LOT in real-world data of patients with mNSCLC
166 (**Table 1, Figure 1**). Among 1,133 patients in ATOMIC with *EGFR* alterations and mNSCLC, a
167 total of 3168 regimens were recorded with a median of 2 regimens per patient (IQR, 1-4; range, 1-
168 13). After applying our rules, there were 2834 total LOTs with a median of 2 LOTs per patient
169 (IQR, 1-3; range, 1-11). Rules 1-3 prevented 334 regimen changes from advancing the LOT (**Table**
170 **1**). Thirty-seven patients switched regimens within 28 days of the start of the first LOT and had a
171 line 0.5 recorded.

172 Comparatively, after applying an algorithm proposed by Hess et al. to the same treatment
173 records, there were 2771 total LOTs with a median of 2 per patient (IQR 1-3; range 1-9). Among
174 our method’s 3999 agent-LOT assignments (e.g., osimertinib as 1L, carboplatin as 2L, etc.), ours
175 differ from the Hess method in 5.7% of assignments (Table 2).

176 Additionally, we developed two rules to remove systemic therapy agents likely used in the
177 treatment of early-stage disease (ESD) from the LOT determination. (1) If a LOT was durvalumab
178 monotherapy administered prior to November 11, 2022 (i.e., prior to Food and Drug
179 Administration approval for durvalumab in stage IV NSCLC)⁴ or (2) if the LOT was chemotherapy
180 with a cumulative cycle duration of 30-140 days and no subsequent LOT for at least four months
181 of follow-up, the LOT should be considered as a treatment for early-stage disease and excluded
182 from the patient’s LOT enumeration for mNSCLC. Among patients in ATOMIC with *EGFR*

183 alterations and mNSCLC, the first LOT for 81 patients met the criteria for early-stage disease
184 treatment (ESD Rule 1 met for two patients and ESD Rule 2 met for 79). These LOTs were
185 removed, and the remaining LOTs were renumbered as LOTs for mNSCLC.

186

187 **Discussion**

188 We propose six simple rules to determine whether a change in treatment for mNSCLC
189 should or should not advance the LOT in real-world data. Our LOT assignments led to different
190 classifications from previous work in 5.7% of cases, primarily in characterizing targeted therapy-
191 containing regimens. Since LOT assignment can affect outcome measures such as real-world
192 progression-free survival or time to treatment discontinuation, even small differences may lead to
193 different conclusions in an observational study. Therefore, it is critical that the LOT determination
194 accurately reflects clinical practice.

195 There are several ways that our proposed rules better characterize LOTs in mNSCLC
196 populations treated with targeted therapy compared to previously reported methods
197 (**Supplemental table 1**).^{2,4-7} First, we distinguish between biologics and therapies targeting
198 oncogenic drivers in mNSCLC. This distinction is important because biologics, such as
199 angiogenesis inhibitors, may have little anti-tumor activity alone and are often added or removed
200 from treatment regimens without a clear change in disease status, while targeted therapies have
201 tremendous anti-tumor activity and are used in the setting of a change in disease status. Second,
202 there are instances where targeted therapies are added to or replace a previous regimen because of
203 a newly identified targetable mutation and not because of a change in disease status. The rules
204 proposed allow for this approach by allowing for 1 cycle of chemotherapy to be given while
205 waiting for molecular testing results prior to switching to a targeted therapy without advancing the
206 LOT. Third, we allow for a treatment pause and re-initiation within 60 days to be considered the
207 same LOT, which may be more appropriate for targeted therapies typically taken daily by mouth
208 that are paused and restarted for toxicity.⁵⁻⁷

209 Retrospective determination of whether a treatment was given in the setting of metastatic
210 disease versus early-stage disease is critical to determining real-world clinical endpoints. In this
211 historical dataset, we used rules based on the approved therapies at the time of the dataset, but even
212 while completing the analysis for this study the indications for several systemic therapies expanded
213 to include early-stage disease, making the agent itself an unreliable indicator of early-stage

214 treatment. Given this complexity, extracting the date of metastatic diagnosis at the time of data
215 abstraction will become even more important to evaluate the practice patterns contemporary to the
216 time of the dataset.

217 We believe wider use of this framework for LOT determination could lead to better method
218 harmonization of observational studies in mNSCLC and allow for comparison and aggregation
219 across datasets in the future.

220

221 **Acknowledgements**

222 This work is supported in part by grants from Takeda, AstraZeneca and LUNGeivity.

223

224 **References**

- 225 1. Hess LM, Cui ZL, Li XI, Molife C, Oton AB. Treatment sequencing for the care of
226 patients with advanced or metastatic non-small cell lung cancer in the United States. *Curr Med*
227 *Res Opin.* 2021;37(3):469-476. doi:10.1080/03007995.2020.1866516
- 228 2. Hess LM, Li X, Wu Y, Goodloe RJ, Cui ZL. Defining treatment regimens and lines of
229 therapy using real-world data in oncology. *Future Oncol Lond Engl.* 2021;17(15):1865-1877.
230 doi:10.2217/fon-2020-1041
- 231 3. R Core Team. R: A language and environment for statistical computing. Published online
232 2022. <https://www.R-project.org/>.
- 233 4. Food and Drug Administration. IMFINZI® (durvalumab) injection, for intravenous use.
234 Published online November 1, 2022. Accessed March 28, 2023.
235 https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/761069s033lbl.pdf
- 236 5. Abernethy AP, Arunachalam A, Burke T, et al. Real-world first-line treatment and
237 overall survival in non-small cell lung cancer without known EGFR mutations or ALK
238 rearrangements in US community oncology setting. *PLoS ONE.* 2017;12(6):e0178420.
239 doi:10.1371/journal.pone.0178420
- 240 6. Johnston S, Wilson K, Varker H, et al. Real-world Direct Health Care Costs for
241 Metastatic Colorectal Cancer Patients Treated With Cetuximab or Bevacizumab-containing
242 Regimens in First-line or First-line Through Second-line Therapy. *Clin Colorectal Cancer.*
243 2017;16(4):386-396.e1. doi:10.1016/j.clcc.2017.03.014

244 7. Meng W, Mosesso KM, Lane KA, et al. An Automated Line-of-Therapy Algorithm for
245 Adults With Metastatic Non-Small Cell Lung Cancer: Validation Study Using Blinded Manual
246 Chart Review. *JMIR Med Inform.* 2021;9(10):e29017. doi:10.2196/29017

247

248 Tables and figures

249 **Table 1:** Proposed rules for line of therapy determination from real-world treatment regimens.

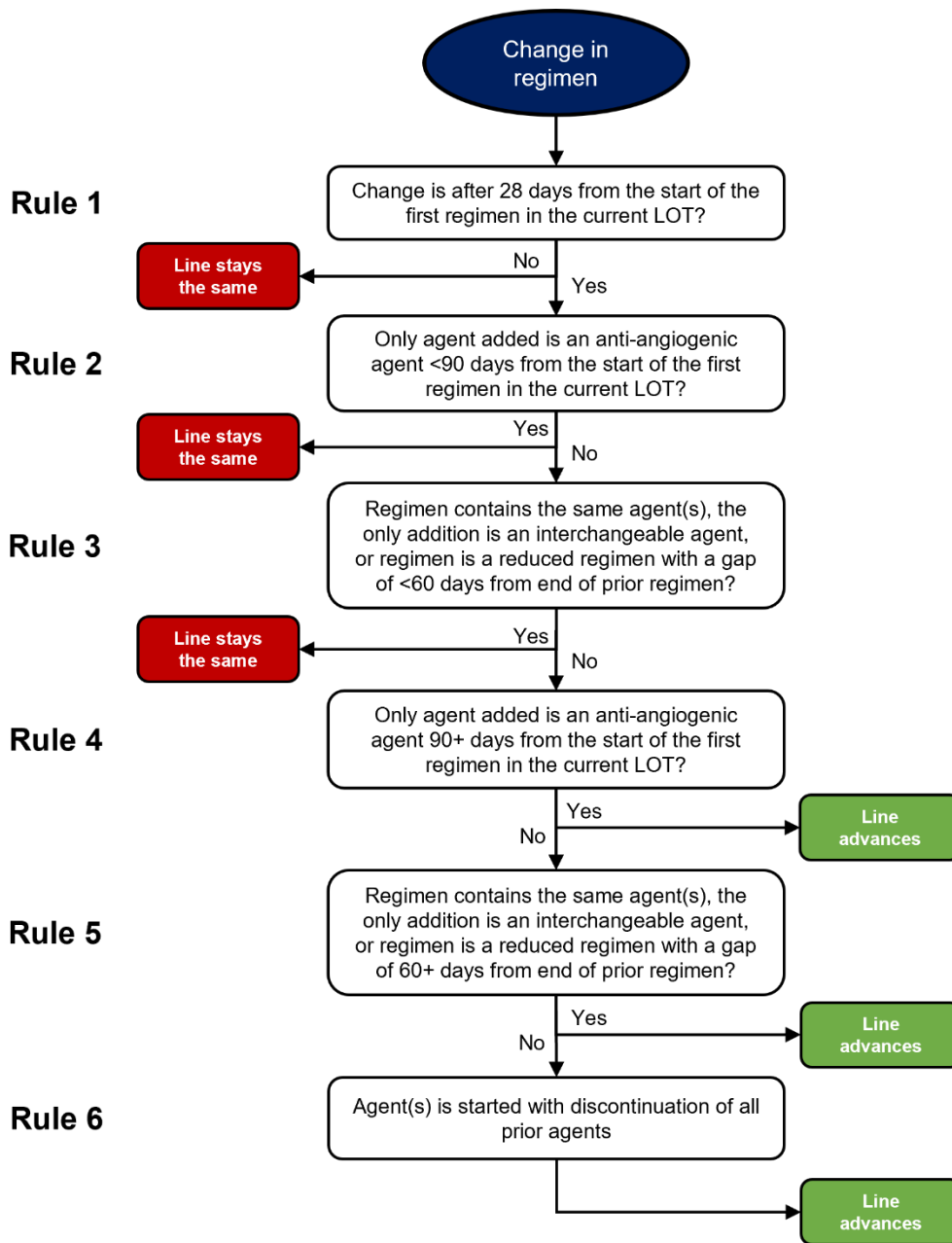
Rule	Line of therapy stays the same:	Example regimen change	Example LOT determination	Frequency rule applied^a
1	If any agent is introduced within 28 days of the start of the first regimen of the current LOT.	Day 0-21: Carboplatin, Pemetrexed Day 14-400: Brigatinib	0.5L ^b (Day 0-21): Carboplatin + Pemetrexed 1L (Day 14-400): Brigatinib	109
2	If an angiogenesis inhibitor (e.g., bevacizumab, ramucirumab) is introduced within 90 days from the start of the first regimen of the current LOT.	Day 0-80: Osimertinib Day 80-200: Osimertinib, Bevacizumab	1L (Day 0-200): Osimertinib + Bevacizumab	1
3	If the gap between the end of the prior and the start of the current regimens is ≤ 60 days and the current regimen uses a. the same agent(s), b. an exchange of interchangeable agents (e.g., Carboplatin/Cisplatin, Paclitaxel/Nab-paclitaxel), or c. a reduced regimen	Day 0-30: Entrectinib Day 80-110: Entrectinib	1L (Day 0-110): Entrectinib	224
Rule	Line of therapy advances:	Example regimen change	Example LOT determination	Frequency rule applied^a
4	If an angiogenesis inhibitor (e.g., bevacizumab, ramucirumab) is introduced >90 days after the start of the first regimen of the current LOT.	Day 1-120: Osimertinib Day 121-160: Osimertinib + ramucirumab	1L (Day 1-120): Osimertinib 2L (Day 121-160) Osimertinib + ramucirumab	22
5	If the gap between the end of the prior and the start of the current regimens is >60 days and the current regimen uses a. the same agent(s), b. an exchange of interchangeable agents (e.g., Carboplatin/Cisplatin, Paclitaxel/Nab-paclitaxel), c. a reduced regimen	Day 1-200: Alectinib Day 300- 400: Alectinib	1L (Day 1-200): Alectinib 2L (Day 300-400): Alectinib	34

6	If a new non-interchangeable agent is added to a regimen or is introduced after the discontinuation of all prior agents >28 days after the start of the first regimen of the current LOT.	Day 1-42:	1L (Day 1-42):	1645
		Carboplatin + pemetrexed	Carboplatin + pemetrexed	
		Day 43-700:	2L (Day 43-700)	
		Lorlatinib	Lorlatinib	

250 ^a Number of times rule is applied in the ATOMIC *EGFR* cohort

251 ^b When a regimen is discontinued followed by the start of a new regimen all within 28 days of
252 the start of the first line of therapy, the first regimen is assigned a line number of 0.5 while still
253 being considered part of first-line therapy. LOT, line of therapy

254 **Figure 1:** Algorithm for line of therapy determination



256 **Table 2:** Example line of therapy determination differences between the proposed rules and
 257 other (Hess et al.) rules

Patient	Duration of regimen (days)	Regimen agent(s)	Proposed LOT	Hess et al. LOT
Patient A	0-150	Erlotinib	1	1
	151- 242	Osimertinib	2	2
	243-438	Bevacizumab, Carboplatin, Pemetrexed	3	3
	439-674	Pemetrexed	3	3
	675-702	Bevacizumab, Pemetrexed	4	3
	703-849	Pemetrexed	4	3
	881-ongoing	Bevacizumab, Erlotinib	5	3
Patient B	0-630	Bevacizumab, Erlotinib	1	1
	632-1036	Osimertinib	2	2
	1038-1091	Osimertinib, Crizotinib	3	2
Patient C	0	Bevacizumab, Carboplatin, Pemetrexed	1	1
	30-800	Erlotinib	2	2
	802-904	Bevacizumab, Carboplatin, Pemetrexed	3	3
	905-1010	Bevacizumab, Pemetrexed	3	3
	1205-1849	Osimertinib	4	4
	1851-2035	Osimertinib, Gefitinib	5	4

258 Highlighted cells indicate differences in how lines of therapy are determined between the
 259 proposed rules and other (Hess et al.) rules. LOT, line of therapy.

260 **Supplemental table 1:** Differences between proposed rules and Hess et al. rules

Scenario	Line of therapy advances	
	Proposed rules	Hess et al. rules
Targeted or immunotherapy is added to a regimen 28+ days after the start of the first regimen in the current LOT.	Yes	No
Targeted or immunotherapy is started with discontinuation of all agents in prior regimen.	Yes* *Except if started <28 days from start of LOT	Yes
An interchangeable agent is exchanged 60+ days after the end of the prior regimen.	Yes	No
Introduction of an anti-angiogenic agent <90 days from the start of any LOT, with or without discontinuation of the prior regimen.	No	Yes* *Except if pemetrexed is part of prior regimen and given with anti-angiogenic agent in the first LOT
A regimen is reduced but with a gap in treatment of 60+ days. (e.g., Day 0-60: Carboplatin + Alectinib, Day 140-ongoing: Alectinib)	Yes	Not addressed in rules

261 LOT, line of therapy