Poster #MSR107

Estimating Optimal Personalized Treatment Sequencing for Patients with Multiple Myeloma Using Reinforcement Learning

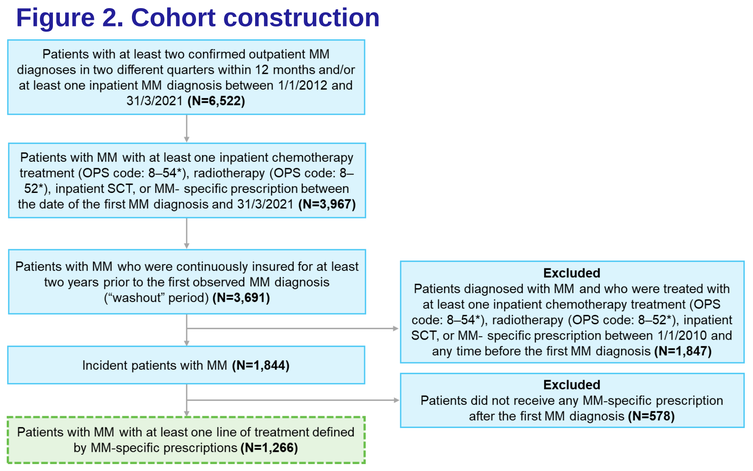
Fei Tang[1], Mahroz Haider[2], Axel Dittmar[3], Jason Simeone[1], Evie Merinopoulou[4], Alind Gupta[2], Andreas Fuchs[5]
[1]Cytel, Inc., Waltham, MA, US; [2]Cytel, Inc., Toronto, ON, Canada; [3]IPAM, Wismar, Germany; [4]Cytel, Inc., London, UK; [5]AOK PLUS, Dresden, Germany

A reinforcement learning (RL) approach predicted different optimal treatment regimens that prolonged survival and time to relapse in patients with multiple myeloma (MM) compared with observed regimens in a German claims setting.

Background

- MM is a hematological malignancy that arises from plasma cells in the bone marrow. This highly heterogeneous disease varies widely in clinical presentation and prognosis.[1]
- Patients with MM can relapse and/or become refractory (i.e., non-responsive) to therapy.
- The optimal sequence of treatments can depend on various factors, such as disease stage, age, and overall health status, highlighting the need for personalized treatment plans and ongoing monitoring of response.
- RL is a branch of machine learning (Figure 1) that involves training an agent to interact with an environment to maximize a reward signal. RL has shown promise in clinical decision-making, including the determination of optimal treatment sequences.[2]
- While RL has been suggested for identifying optimal treatment sequences by learning from real-world data, its application in administrative claims research is limited.

Figure 1. Main types of machine learning



Source: Peng et al., 2021[3]

Objective

- This study sought to use RL to identify patient-specific optimal treatment sequences that maximize the time to death or relapsed/refractory status among patients with MM using German claims data.

Methods

- This retrospective, observational cohort study used claims data from the German AOK PLUS health insurance fund (January 1, 2010 to March 31, 2022).
- Detailed inclusion and exclusion criteria are shown in Figure 2.

Figure 2. Cohort construction

Abbreviations, MM, multiple myeloma; OPS, operation and procedure; SCT, stem cell transplant

- New lines of treatment were defined as the introduction of a new agent not part of the prior line.
- Treatment sequences included up to two lines of treatment regimens.
- The outcome of interest was the duration between the initiation of the first-line (1L) treatment to either death or relapsed/refractory status, which was defined as the initiation of the third line of therapy.
- An RL framework with generalized survival random forest as an estimator for survival outcomes[4] was used to identify optimal treatment sequences that maximize the time to death or relapsed/refractory status, adjusting for relevant demographic and clinical characteristics during baseline (within 12 months prior to the initiation of 1L treatment).
- To ensure the model could converge, 297 (23%) patients who initiated rare regimens were excluded, leaving 969 patients in the treatment sequence analysis.

Results

Patient selection

- In total, 1,266 patients with MM who received at least 1L treatment were included in the study. (Figure 2)

Baseline characteristics (results not shown)

- The average age at diagnosis was 71 years old, and slightly more than half of patients were male (52%).
- Median Charlson Comorbidity Index score was 6.
- A substantial number of patients experienced hypertension (83%), bone-related disorders (54%), moderate or severe renal disease (49%), and neurological disease (46%) during the baseline period.

Treatment patterns (Table 1)

Observed patterns in German claims:

- The most common 1L regimen was bortezomib (BORT) monotherapy.
- The most common second-line (2L) regimen was combination therapy of lenalidomide (LEN) + dexamethasone (DEX).

Patterns predicted by the RL algorithm:

- The most frequently predicted optimal 1L regimen was BORT + cyclophosphamide + DEX (CyBorD).
- The most frequently predicted optimal 2L regimen was LEN + daratumumab (DARA) + DEX + prednisolone.
- Disparities were observed between commonly prescribed 1L and 2L regimens and most frequently predicted optimal 1L and 2L regimens. Higher proportions of combination therapy were recommended by the algorithm as optimal 1L (85% vs. 64%) and 2L (95% vs. 80%) regimens, compared with the actual regimens received by patients.

Commonly observed treatment sequences (results not shown)

- The utilization of BORT as 1L treatment was the most observed treatment sequence (n=139; 11%).
- The utilization of BORT as 1L treatment and LEN + DEX as 2L treatment was the second-most observed treatment sequence (n=63; 5%).

Table 1. Most frequently observed and predicted 1L and 2L regimens

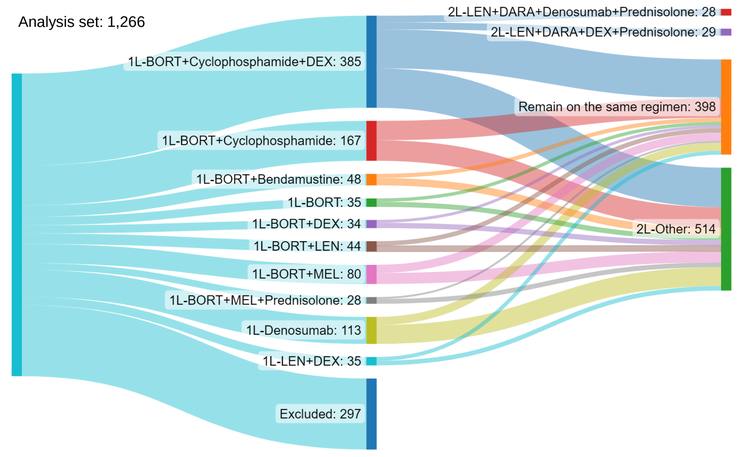
Most frequently observed regimens	N (%)	Most frequently predicted optimal regimens	N (%)
1L			
Total with 1L therapy	1,266 (100)	Total with 1L therapy	969 (100)
BORT	362 (28.59)	CyBorD	385 (39.73)
BORT + DEX	119 (9.40)	BORT + cyclophosphamide	167 (17.23)
BORT + MEL + prednisolone	80 (6.32)	Denosumab	113 (11.66)
BORT + MEL	77 (6.08)	BORT + MEL	80 (8.26)
BORT + cyclophosphamide	75 (5.92)	BORT + bendamustine	48 (4.95)
Combination therapy	**808 (63.82)**	**Combination therapy**	**821 (84.73)**
Monotherapy	**458 (36.18)**	**Monotherapy**	**148 (15.27)**
2L			
Total with 2L therapy	697 (100)	Total with 2L therapy	571 (100)
LEN + DEX	147 (21.09)	LEN + DARA + DEX+ prednisolone	61 (10.68)
LEN	62 (8.9)	LEN + DARA + denosumab + prednisolone	58 (10.16)
BORT + DEX	45 (6.46)	Elotuzumab + pomalidomide	30 (5.25)
LEN + carfilzomib + DEX	31 (4.45)	CyBorD	28 (4.90)
BORT	29 (4.16)	Isatuximab + carfilzomib + prednisolone	26 (4.55)
Combination therapy	**558 (80.06)**	**Combination therapy**	**540 (94.57)**
Monotherapy	**139 (19.94)**	**Monotherapy**	**31 (5.43)**

Abbreviations: 1L, first line; 2L, second line; BORT, bortezomib; CyBorD, bortezomib + cyclophosphamide + dexamethasone; DARA, daratumumab; DEX, dexamethasone; LEN, lenalidomide; MEL, melphalan

Frequently predicted optimal treatment sequences

- The Sankey diagram (Figure 3) illustrates the treatment sequences of the predicted optimal regimens which prolong time to death or relapse/refractory status among patients with MM. The utilization of CyBorD as 1L treatment was the optimal treatment sequence most frequently recommended by the algorithm (n=163; 17%). However, only 34 (3%) patients with MM were prescribed this treatment sequence in the real world.

Figure 3. Sankey diagram of the predicted optimal treatment sequences

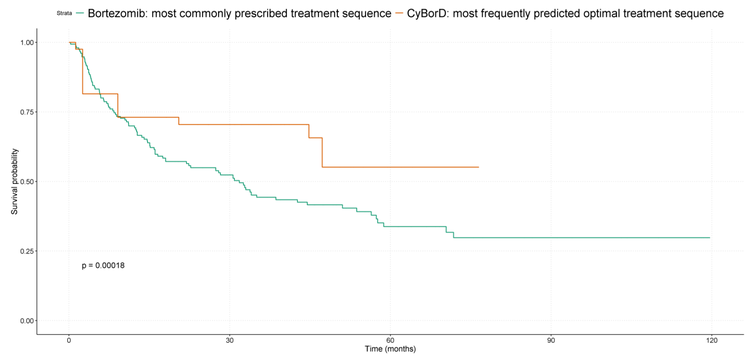


Regimens with fewer than 20 patients are subsumed as "other."
Abbreviations: BORT, bortezomib; DEX, dexamethasone; LEN, lenalidomide; MEL, melphalan

Prognosis among patients receiving CyBorD as 1L treatment vs. patients receiving bortezomib as 1L treatment

- Results of the Kaplan-Meier survival curves (Figure 4) comparing patients receiving CyBorD as 1L treatment (the optimal treatment sequence most frequently recommended by the RL algorithm) with patients receiving bortezomib as 1L treatment (the treatment sequence most frequently prescribed in the real world) indicated better prognosis among patients with CyBorD.

Figure 4. Prognosis among patients receiving CyBorD as 1L treatment vs. patients receiving bortezomib as 1L treatment, adjusted for relevant baseline characteristics



Abbreviation: CyBorD, bortezomib + cyclophosphamide + dexamethasone

Conclusions

- Personalized optimal treatment sequences for patients with MM were identified using an RL algorithm.
- Utilization of CyBorD as 1L treatment was the optimal regimen most frequently recommended by the algorithm, which aligns with findings from previous studies that CyBorD is a highly responsive regimen in newly diagnosed patients with MM.[5]
- The optimal 1L and 2L treatment regimens recommended by the algorithm were different from real-world treatment patterns; in particular, the use of combination therapy was more recommended in the first two lines.
- Patients who received 1L CyBorD exhibited longer time to relapse/refractory status or death, suggesting better disease prognosis among these patients.
- The established framework may provide a robust and clinically meaningful approach for estimating personalized treatment regimens that can improve patient outcomes and address heterogeneity in MM and similar disease settings using real-world claims data.

Future extension

- Inclusion of time-varying features may be crucial when predicting optimal treatment sequences, which will be assessed in the subsequent analytic stage.
- Associations of model features and the predicted optimal treatment sequences will also be evaluated.
- Validation/test datasets are generally not used in RL, which differs from traditional machine learning algorithms. An independent dataset of patients with MM will be used to further validate the established framework.

References

1. Cowan AJ, Green DJ, Kwok M, et al. Diagnosis and management of multiple myeloma: a review. Jama. 2022;327(5):464-477; 2. Coronato A, Naeem M, De Pietro G, Paragliola G. Reinforcement learning for intelligent healthcare applications: A survey. Artificial Intelligence in Medicine. 2020;109:101964; 3. Peng J, Jury EC, Dönnes P, Ciurtin C. Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: applications and challenges. Frontiers in pharmacology. 2021;12:720694; 4. Cho H, Holloway ST, Couper DJ, Kosorok MR. Multi-stage optimal dynamic treatment regimes for survival outcomes with dependent censoring. Biometrika. 2023;110(2):395-410; 5. Reeder CB, Reece DE, Kukreti V, Chen C, Trudel S, Hentz J, et al. Cyclophosphamide, bortezomib and dexamethasone induction for newly diagnosed multiple myeloma: high response rates in a phase II clinical trial. Leukemia. 2009;23(7):1337-41.

Presented at ISPOR EU | Copenhagen, Denmark | November 12–15, 2023

Contact information: www.cytel.com