Trends in Cancer

Special Issue: Quantitative Cancer Biology

Opinion

A Quantitative Paradigm for Decision-Making in Precision Oncology

Dalit Engelhardt[1,2,3,4,*] and Franziska Michor[1,2,3,4,5,6]

The complexity and variability of cancer progression necessitate a quantitative paradigm for therapeutic decision-making that is dynamic, personalized, and capable of identifying optimal treatment strategies for individual patients under substantial uncertainty. Here, we discuss the core components and challenges of such an approach and highlight the need for comprehensive longitudinal clinical and molecular data integration in its development. We describe the complementary and varied roles of mathematical modeling and machine learning in constructing dynamic optimal cancer treatment strategies and highlight the potential of reinforcement learning approaches in this endeavor.

New Frontiers for Decision-Making in Precision Oncology

Clinical oncology has undergone a period of revolutionary transformation owing to rapid advances in precision therapeutics and personalized medicine. Progress in diagnostic resolution and computational methods has yielded unprecedented insights into the complexity, heterogeneity, and dynamics of cancer and has led to the development of new therapeutic approaches capable of delivering more effective treatments [1]. Yet, the variety of clinical options and the large amounts of data that can be collected present new challenges for decision-making in oncology. Clinicians must determine which therapeutic option should be chosen for a particular patient at a particular time point while contending with uncertainty regarding the current state and future course of disease. These options must take into account the long-term effects of drugs administered early during treatment on the clonal tumor composition later during treatment or upon recurrence, and they need to balance efficacy and side effects throughout treatments that can last for decades. Clinical decisions need to account for all of these considerations in order to maximize the potential of available therapeutics and the success of new ones.

In the face of these opportunities and challenges, a comprehensive quantitative approach to personalized decision-making in precision oncology has become a practical imperative. Core to this endeavor is the ability to match the dynamic and potentially recurrent nature of cancer with a cohesive and farsighted treatment strategy – or sequence of treatment decisions – that is capable of robustly and dynamically steering the temporally variable and stochastic progression of cancer toward a desired path. Dynamic decision rules for medical interventions constitute an active area of research that has seen growing interest with the broadening of applications of personalized medicine. Generally known as dynamic treatment regimes, treatment policies, or adaptive treatment strategies [2–10], they consist of sequences of decisions in which observations of the state of the disease are periodically made; this feedback is used to inform the course of treatment until the next decision point at which new observations are made. These strategies adapt to the current state of the disease subject to certain goals such as maintenance or progression milestones and manageability of side effects throughout treatment (e.g., by avoiding spikes in drugs through pre-emptive and timely interventions). In particular, they can be made to be farsighted, with current decisions informed not only by their immediate expected outcome (e.g., the

Highlights

Tumor heterogeneity, stochasticity in tumor progression, and a limited detection threshold of drug-resistant variants necessitate a dynamic and farsighted approach to personalized therapeutic interventions that incorporates considerations of long-term efficacy and toxicity.

A clear working framework for the formulation and implementation of dynamic decision-making in cancer can drive and facilitate multidisciplinary clinical and research collaborations in treatment optimization.

A dynamic decision-making algorithm should include a specification of optimization criteria, structural constraints, mechanisms for projections of sequences of outcomes, and an optimization mechanism.

Longitudinal integrative data – genomic, imaging, clinical, and others – collected from patients is essential for the personalized optimization of dynamic treatment strategies.

The synthesis of mathematical modeling of tumor dynamics and advances in machine learning approaches to dynamic decision-making provides a promising avenue for addressing the challenges of developing personalized interventions in oncology.

The incorporation of deep learning into reinforcement learning algorithms has significant potential in optimization across variable datasets with a large number of features, and future work in this field should consider how algorithmic improvements can best serve the challenges of decision-making optimization in cancer.

[1]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA

expected range of tumor burden reduction in 3 months) but also by a long-term outcome (e.g., extending progression-free survival and avoiding multidrug-resistant recurrences). The translation of short-term predictions into flexible, farsighted, and long-term optimal sequences of decisions is done through mathematical and computational mechanisms that have been developed for the optimization of these dynamic decision processes [11].

This pan-treatment strategic view is of critical importance in oncology, where treatments can be long term, and recurrences and resistance to treatment can go undetected for long periods of time due to the variant detection limitations of even the highest-resolution methods [12–16]. Yet, despite substantial theoretical and computational research on clinical dynamic treatment policies, practical applications of such approaches in oncology are rare. Cancer presents a particularly complex treatment environment with high interpatient variability, and our current understanding of its mechanisms and progression is the result of multifaceted efforts in data collection and analysis, laboratory experiments, and mathematical modeling. Implementation of a quantitative and robust dynamic decision-making approach requires distilling this complex space into the mathematical structures that define a decision process in an integrative and streamlined way that can facilitate collaboration between the different communities involved in cancer research and care. To this end, we have set here to describe this process. In outlining the components and development of a dynamic personalized decision-making model for cancer treatment (Figure 1), we hope to lay the foundation for a common working framework for clinical oncologists, biologists, and practitioners of decision theory to collaborate on the personalized optimization of therapeutic decision-making in different areas of oncology.

Dynamic Decision-Making in Cancer Treatments

A dynamic sequential decision-making model (Figure 2) aims to select, as part of a sequential strategy, the best action – here, treatment – for any observed state of the disease, subject to specified optimization criteria and under structural constraints that define the decision process, such as how often we are given information on a state, how often we can take an action, and what types of actions are possible. Establishing these optimization criteria and structural

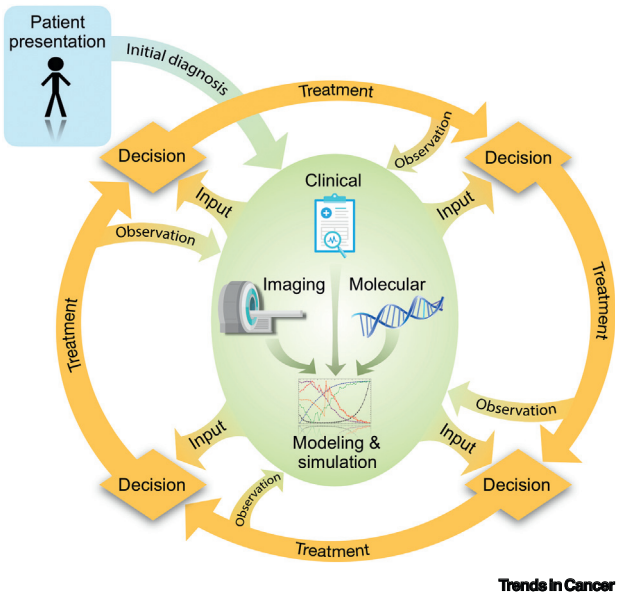

Figure 1. Depiction of an Integrative Dynamic Decision-Making Process in Cancer Treatment. Data collected at patient presentation is supplied to a trained decision-making model and used to inform treatment decisions until subsequent observations are supplied to the model to inform the next therapeutic decision. Molecular and imaging data can be used to parameterize mathematical models of tumor dynamics in order to supplement and complement the collected patient data; not all data types may be available at each decision step.

[2]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA
[3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA
[4]Center for Cancer Evolution, Dana–Farber Cancer Institute, Boston, MA, USA
[5]The Broad Institute of Harvard and MIT, Cambridge, MA, USA
[6]The Ludwig Center at Harvard, Boston, MA, USA

*Correspondence:
dengelhardt@ds.dfci.harvard.edu
(D. Engelhardt).

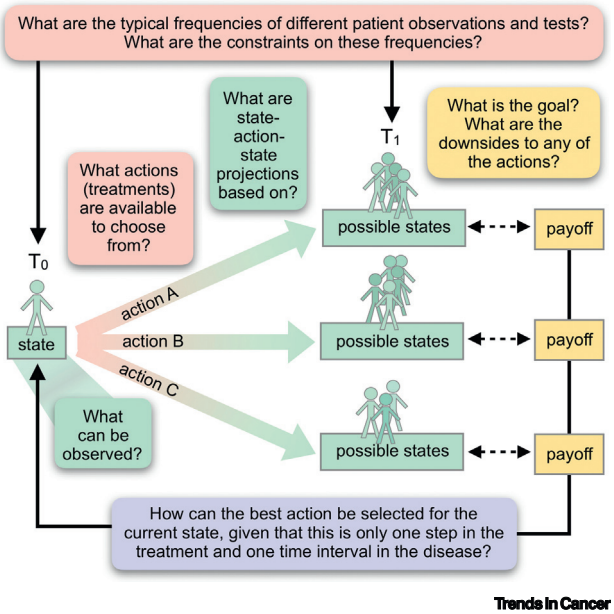

Figure 2. Optimization of a Dynamic Decision-Making Process. This process involves determination of what action should be taken in a given state subject to basic structural constraints (pink) and optimization criteria (yellow). This determination is implemented through a decision optimization mechanism (purple), which relies on recorded sequences of events or projections based on data and/or models of the likely outcomes of taking any of the available actions given the features of the current state (green).

constraints comprises the first step in the formulation of a decision-making model. Structural constraints include the typical frequency of appointments, the range of reasonable frequency of appointments, the types of observations and tests (e.g., clinical examination, specimen collection, and biopsy) that can be collected and their acceptable frequency ranges, and the full space of actions: which types of therapies or procedures may be available, including whether these treatments are authorized for use only at certain times or based on certain patient covariates within the diagnostic class. The optimization criteria include a specification of the general goal of the decision process: does the treatment have a curative intent? What type of balance between quality of life and tumor burden reduction do we seek at different stages of treatment? Are treatment cost and hospital resources major limiting factors? On a more specific level, it can include preferences for specific treatments based on toxicity, cost, or other factors. In the decision model, these criteria are captured as payoffs: numerical representations of these preferences that determine the optimization of decision sequences. Ultimately, when the development of a decision-making model is complete, simulation studies should be conducted that explore different representations of optimization criteria. The results of these studies – simulated treatment trajectories under different payoff assignments – should be presented to healthcare teams for feedback prior to final model selection.

The central object of the decision-making model is the disease state. The true disease state is not fully observable: it is comprised of every mechanism, interaction, and process taking place within the patient. Instead, in defining the state in the model, we select from the set of features that a decision-maker can access – either through observations or modeling – and on which we have enough information to generate projections for the likelihood of transitioning from one set of observed features to another under the actions considered in the decision-making process. Features should therefore be derived from longitudinal patient observations, although not all features must be known at every point of observation; for example, biopsies of solid malignancies might be taken only initially and upon a recurrence. Given the substantial interpatient heterogeneity and the complexity of cancer treatments, longitudinal observations should also be horizontally

integrated, incorporating diverse information from genomics, imaging, and clinical data into the definition of the disease state (Figure 1). Individual features can range from high-level preidentified biomarkers to raw inputs, with the caveat that while a large number of features will result in more patient-specific information, model training will necessitate a larger amount of data. In addition to greater personalization, horizontal integration can enhance the ability of the model to generalize to states not seen in the training data by identifying how distances, or extent of similarity, in this high-dimensional feature space may correspond to similarity in treatment response. Due to the crossdisciplinary challenges of data integration, successful construction of a disease state would strongly benefit from collaboration between researchers familiar with different types of data and models in bringing these efforts to fruition.

Projections of how a disease state will change under a certain action may be constructed from raw data, distributions derived from data, or simulations based on models and data. Sequential Multiple Assignment Randomized Trials (SMARTs) [2,7,17], which consist of multiple sequential steps of participant randomization to different treatment options, provide a preferred mechanism for generating the data used for estimating adaptive treatment strategies. SMART or SMART-precursor designs have been implemented in several cancer studies [18–27], but the design and use of cancer SMARTs for the purpose of analyzing adaptive treatment strategies has been limited [28,29]. At present, most data available from cancer patients are acquired either through more standard clinical trials or outside of trials, and hence our focus here is on this type of data; for example, retrospective observational studies across multiple longitudinal datasets or predictive analyses based on specific clinical trial designs.

Complementing Data-Driven Optimization with Mathematical Modeling

An important additional channel for addressing the limitations of observational studies in treatment exploration is afforded by mathematical modeling and simulation of tumor dynamics [30,31]. Modeling and simulation can be used to identify relevant features for use in optimization, investigate the ability of a decision-making model to generalize its performance to wide parameter ranges not seen during training [32], and simulate and explore *in silico* the effects of potential treatment sequences prior to clinical trials. For instance, mathematical models of cancer evolution have suggested dosing schedules aimed at addressing the problem of acquired resistance to tyrosine kinase inhibitors (TKIs) in epidermal growth factor receptor (EGFR)-mutant non-small cell lung cancer (NSCLC) patients: high-dose pulses in combination with low-dose continuous dosing of erlotinib was predicted to maximize the delay of onset of resistance [33], and such schedules were subsequently tested in a clinical trial [34]. Another modeling approach [35] recommended against prolonged pre-chemoradiotherapy use of TKIs due to predictions of reduced efficacy of adjuvant TKI maintenance therapy. Among many other examples, mathematical models have also been used to address the development of resistance during prostate cancer therapy by capitalizing on longitudinal prostate-specific antigen (PSA) levels: models were used to support clinical trials [36] of abiraterone adaptive therapy [37] cycles in metastatic prostate cancer, and to predict individual patient response during intermittent androgen deprivation therapy of biochemically recurrent prostate cancer to subsequent treatment cycles [38]. Imaging provides an additional channel for non-invasive longitudinal monitoring of tumor progression, based on which simulations of different treatment schedules can be constructed for *in silico* treatment optimization [39].

Serial collection of genomic patient data can form the basis for predictive modeling of tumor evolutionary trajectories [40–42]. While the invasiveness of tumor biopsies limits the availability and frequency of longitudinal genomic datasets in solid malignancies, genomic analysis of cell-free circulating tumor DNA is emerging as a promising and non-invasive avenue for detecting

longitudinal tumor changes [43,44] and may present certain advantages, such as the ability to capture greater spatial heterogeneity, over solid tumor samples, in particular in the detection of acquired resistance [45]. For instance, a recent approach [40] combined serial sampling and genomic profiling of cell-free DNA with tumor burden measurements and population dynamics models for sensitive and resistant cell populations to predict time to treatment failure in a trial cohort of colorectal cancer patients undergoing treatment with cetuximab.

Deep Reinforcement Learning for Decision Optimization in Cancer

We have described the optimization criteria, structural constraints, selection of features defining the disease state, and mechanisms for generating projections that underly a dynamic decision-making model. A key piece of machinery in such a model is the mechanism for determining how an action should be optimally chosen in each disease state as part of a cohesive strategy of sequential choices. In general, the methods that have been developed for optimizing treatment policies span the spectrum from statistical causal inference approaches to computer science-based reinforcement learning (RL) approaches [11]. In developing cancer treatment strategies, the variability between molecular, cellular, and physiological data as well as simulation-based synthetic data brings into question how such diverse sources of information should be combined in an optimization algorithm (Figure 2). Deep learning approaches have become pervasive in biomedical applications due to their ability to learn from complex, high-dimensional data such as a large number of patient covariates or information-dense imaging data; however, in a decision-making model, deep learning algorithms must be used in conjunction with a mechanism to optimize a sequence of decisions based on the input data. Advances in recent years in the incorporation of deep neural networks into RL [46,47] – a rapidly evolving area of machine learning approaches aimed at decision-making and control – provide mechanisms for combining the advantages and successes of RL and deep learning in training on variable and complex input data. RL algorithms are based on the idea that optimal decision-making strategies – which action should be picked in a given state – can be learned through automated feedback (the payoffs, or 'rewards', defined through the optimization criteria) on sequences of states directly experienced during training. The reliance on optimality-based feedback sets RL apart from supervised learning, in which ground truth labels must be provided, and unsupervised learning, which is aimed at finding structure in unlabeled data. The potential of deep RL for complex treatment optimization problems in cancer has been gaining attention in recent years. For instance, deep RL methods were used [48] to generate automated adaptive radiation protocols for NSCLC patients by incorporating molecular and imaging patient data into optimization. Another approach using deep RL was developed [49] for the prevention and treatment of graft versus host disease in leukemia patients following allogeneic hematopoietic cell transplantation using longitudinal transplant patient data. Deep RL was used for simulation-based optimization of clinical trial chemotherapy and radiotherapy administration [50]. Deep RL was also combined with evolutionary modeling [32] in a simulation-based study of drug resistance suppression via optimal dosing in bacterial populations.

Despite growing interest, work incorporating deep RL into cancer treatment optimization remains limited, in part due to the relative novelty of deep RL algorithms and the technical challenges involved in their implementation. Future work in this field will need to address specific challenges involved in using RL to learn from observational medical data [51,52], as well as general challenges involved in the learning of RL policies without interactive data collection during training [53]. There is a clinical need to address these challenges and apply such approaches to large-scale integrated patient data, and reciprocal efforts in data acquisition and algorithmic development between clinical researchers and machine learning practitioners will likely propel productive work in this area (see Outstanding Questions).

Concluding Remarks

Despite breakthrough advances in precision diagnostics and treatments, the therapeutic decision-making process in cancer remains highly complex and plagued by a number of challenges. We have argued here that a systematic, quantitative framework for developing dynamic decision-making models for cancer diagnoses stands to greatly advance personalized treatments in precision oncology. By outlining the main ingredients and the development process of such models, we hope to engender and facilitate collaborations between crossdisciplinary teams of clinicians and researchers on the formulation and development of dynamic decision-making models in oncology.

A major challenge that such endeavors face is data acquisition. Longitudinal datasets are of critical importance in optimizing a multistep decision process with potentially long-term consequences. These datasets should also be horizontally integrated across all data types (clinical examinations, biopsies, blood tests, etc.) collected during treatment. Although longitudinal data are routinely collected by clinics and hospitals, accessing such datasets for retrospective research is often fraught with challenges: they can be dispersed between different hospital data systems, retained by clinical groups using them for specific studies, present data curation challenges [54], or are otherwise not available in a form that makes it feasible with limited resources to longitudinally and horizontally combine and curate the treatment records of a patient cohort with a particular diagnosis. Decision-making collaborations in smaller healthcare facilities are further limited by small sample sizes.

Federated learning, a paradigm for collaboratively training models on data held securely in individual institutions, has been gaining attention in healthcare informatics as a way to address the need for large and diverse datasets while maintaining the security and privacy of patient information [55,56]. Until such efforts become widely implemented, data sharing agreements between healthcare systems with a streamlined, common data access application and approval process for health researchers can provide a much-needed expansion of the data available for health informatics and create increased incentives to standardize data collection across systems. When possible, prospective studies can be designed in collaboration with decision algorithm developers to implement an appropriate data recording scheme throughout the study. The focus of these studies should be on keeping clear longitudinal records – time-resolved sequences of clinical records and tests – of patients who are tracked as part of the study. Such data collection and organization would significantly facilitate subsequent work in treatment optimization; like retrospective studies, a large collaboration between multiple clinics would result in a larger dataset for the optimization analysis.

Ultimately, one of the greatest challenges in constructing decision-making models for cancer involves the formation of interdisciplinary collaborations necessary for the development and implementation of a decision-making model. The end-users of such models – clinicians – must be involved in the process of their development to ensure that a model is trusted and appropriately integrated in the clinical workflow [57]. For individual patients to fully benefit from the data and diagnostics available today, sequential decision-driven analysis that is comprehensive, integrative, and that leverages both data-driven and mathematical modeling approaches provides unparalleled opportunities for advancing treatment optimization efforts across the landscape of precision oncology.

Outstanding Questions

What should be done to foster the creation of multidisciplinary task forces dedicated to decision-making optimization with expertise in genomics, imaging data, electronic health records, modeling and simulation, and algorithmic development?

Can clinical data collection protocols be modified to emphasize large-scale collection of non-invasive serial diagnostics needed for sequential optimization studies?

What steps should be taken specifically in cancer treatment optimization to address the practical and theoretical challenges of reinforcement learning algorithms that use observational clinical data?

What steps can be taken in the near future in order to increase the data available for model training without compromising patient privacy and data security? How can we ensure that the benefits of data-driven treatment optimization extend to smaller healthcare facilities with limited data and collaboration resources?

How can decision-making algorithms best accommodate transparency and patient and clinician feedback throughout clinical implementation?

Acknowledgments

The authors gratefully acknowledge support from the Ludwig Center at Harvard, the Dana-Farber Cancer Institute Physical Sciences-Oncology Center (NIH U54CA193461; to F.M.) and the Helen Gurley Brown Presidential Initiative at Dana-Farber Cancer Institute (to both).

Declaration of Interests

D.E. is the inventor on a pending patent application on machine learning techniques for the determination of therapeutic agent dosages filed by Harvard University. F.M. is the co-founder of an oncology company.

References

1. DeVita, V.T. *et al.* (2018) *DeVita, Hellman, and Rosenberg's Cancer: Principles & Practice of Oncology* (11th edn), Lippincott Williams & Wilkins
2. Lavori, P.W. and Dawson, R. (2000) A design for testing clinical strategies: biased adaptive within-subject randomization. *J. R. Stat. Soc. Ser. A* 163, 29–38
3. Lavori, P.W. and Dawson, R. (2008) Adaptive treatment strategies in chronic disease. *Annu. Rev. Med.* 59, 443–453
4. Thall, P.F. *et al.* (2000) Evaluating multiple treatment courses in clinical trials. *Stat. Med.* 19, 1011–1028
5. Thall, P.F. *et al.* (2002) Selecting therapeutic strategies based on efficacy and death in multicourse clinical trials. *J. Am. Stat. Assoc.* 97, 29–39
6. Thall, P.F. *et al.* (2007) Bayesian and frequentist two-stage treatment strategies based on sequential failure times subject to interval censoring. *Stat. Med.* 26, 4687–4702
7. Murphy, S.A. (2005) An experimental design for the development of adaptive treatment strategies. *Stat. Med.* 24, 1455–1481
8. Lunceford, J.K. *et al.* (2002) Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials. *Biometrics* 58, 48–57
9. Wahed, A.S. and Tsiatis, A.A. (2004) Optimal estimator for the survival distribution and related quantities for treatment policies in two-stage randomization designs in clinical trials. *Biometrics* 60, 124–133
10. Wahed, A.S. and Tsiatis, A.A. (2006) Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data. *Biometrika* 93, 163–177
11. Chakraborty, B. and Moodie, E.E.M. (2013) *Statistical methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine*, Springer
12. Fox, E.J. *et al.* (2014) Accuracy of next generation sequencing platforms. *Next Gener. Seq. Appl.* 1, 1000106
13. Salk, J.J. *et al.* (2018) Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat. Rev. Genet.* 19, 269–285
14. Azuara, D. *et al.* (2012) Nanofluidic digital PCR for KRAS mutation detection and quantification in gastrointestinal cancer. *Clin. Chem.* 58, 1332–1341
15. Coccaro, N. *et al.* (2020) Digital PCR: a reliable tool for analyzing and monitoring hematologic malignancies. *Int. J. Mol. Sci.* 21, 3141
16. Cilloni, D. *et al.* (2019) Digital PCR in myeloid malignancies: ready to replace quantitative PCR? *Int. J. Mol. Sci.* 20, 2249
17. Lavori, P.W. and Dawson, R. (2004) Dynamic treatment regimes: practical design considerations. *Clin. Trials* 1, 9–20
18. Stone, R.M. *et al.* (1995) Granulocyte-macrophage colony-stimulating factor after initial chemotherapy for elderly patients with primary acute myelogenous leukemia. Cancer and Leukemia Group B. *N. Engl. J. Med.* 332, 1671–1677
19. Joss, R.A. *et al.* (1994) Combined-modality treatment of small-cell lung cancer: randomized comparison of three induction chemotherapies followed by maintenance chemotherapy with or without radiotherapy to the chest. Swiss Group for Clinical Cancer Research (SAKK). *Ann. Oncol.* 5, 921–928
20. Tummarello, D. *et al.* (1997) A randomized, controlled phase III study of cyclophosphamide, doxorubicin, and vincristine with etoposide (CAV-E) or teniposide (CAV-T), followed by recombinant interferon-alpha maintenance therapy or observation, in small cell lung carcinoma patients with complete responses. *Cancer* 80, 2222–2229
21. Stone, R.M. *et al.* (2001) Postremission therapy in older patients with de novo acute myeloid leukemia: a randomized trial comparing mitoxantrone and intermediate-dose cytarabine with standard-dose cytarabine. *Blood* 98, 548–553
22. Matthay, K.K. *et al.* (1999) Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid. Children's Cancer Group. *N. Engl. J. Med.* 341, 1165–1173
23. Matthay, K.K. *et al.* (2009) Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a children's oncology group study. *J. Clin. Oncol.* 27, 1007–1013
24. Habermann, T.M. *et al.* (2006) Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *J. Clin. Oncol.* 24, 3121–3127
25. van Oers, M.H. *et al.* (2006) Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood* 108, 3295–3301
26. Mateos, M.V. *et al.* (2010) Bortezomib, melphalan, and prednisone versus bortezomib, thalidomide, and prednisone as induction therapy followed by maintenance treatment with bortezomib and thalidomide versus bortezomib and prednisone in elderly patients with untreated multiple myeloma: a randomised trial. *Lancet Oncol.* 11, 934–941
27. Thall, P.F. *et al.* (2007) Adaptive therapy for androgen-independent prostate cancer: a randomized selection trial of four regimens. *J. Natl. Cancer Inst.* 99, 1613–1622
28. Kidwell, K.M. (2014) SMART designs in cancer research: past, present, and future. *Clin. Trials* 11, 445–456
29. Ruppert, A.S. *et al.* (2019) Application of a sequential multiple assignment randomized trial (SMART) design in older patients with chronic lymphocytic leukemia. *Ann. Oncol.* 30, 542–550
30. Altrock, P.M. *et al.* (2015) The mathematics of cancer: integrating quantitative models. *Nat. Rev. Cancer* 15, 730–745
31. Brady, R. and Enderling, H. (2019) Mathematical models of cancer: when to predict novel therapies, and when not to. *Bull. Math Biol.* 81, 3722–3731
32. Engelhardt, D. (2020) Dynamic control of stochastic evolution: a deep reinforcement learning approach to adaptively targeting emergent drug resistance. *J. Mach. Learn. Res.* 21, 1–30
33. Chmielecki, J. *et al.* (2011) Optimization of dosing for EGFR-mutant non-small cell lung cancer with evolutionary cancer modeling. *Sci. Transl. Med.* 3, 90ra59
34. Yu, H.A. *et al.* (2017) Phase 1 study of twice weekly pulse dose and daily low-dose erlotinib as initial treatment for patients with EGFR-mutant lung cancers. *Ann. Oncol.* 28, 278–284
35. McClatchy, D.M. *et al.* (2020) Modeling resistance and recurrence patterns of combined targeted-chemoradiotherapy predicts benefit of shorter induction period. *Cancer Res.* 80, 5121–5133
36. Zhang, J. *et al.* (2017) Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nat. Commun.* 8, 1816
37. Gatenby, R.A. *et al.* (2009) Adaptive therapy. *Cancer Res.* 69, 4894–4903
38. Brady-Nicholls, R. *et al.* (2020) Prostate-specific antigen dynamics predict individual responses to intermittent androgen deprivation. *Nat. Commun.* 11, 1750
39. Yamamoto, K.N. *et al.* (2019) Computational modeling of pancreatic cancer patients receiving FOLFIRINOX and gemcitabine-based therapies identifies optimum intervention strategies. *PLoS One* 14, e0215409
40. Khan, K.H. *et al.* (2018) Longitudinal liquid biopsy and mathematical modeling of clonal evolution forecast time to treatment failure in the PROSPECT-C phase II colorectal cancer clinical trial. *Cancer Discov.* 8, 1270–1285
41. Angelova, M. *et al.* (2018) Evolution of metastases in space and time under immune selection. *Cell* 175, 751–765 e16
42. Gutierrez, C. and Wu, C.J. (2019) Clonal dynamics in chronic lymphocytic leukemia. *Blood Adv.* 3, 3759–3769
43. Cescon, D.W. *et al.* (2020) Circulating tumor DNA and liquid biopsy in oncology. *Nat. Cancer* 1, 276–290
44. Keller, L. *et al.* (2020) Clinical relevance of blood-based ctDNA analysis: mutation detection and beyond. *Br. J. Cancer* 345–358

45. Parikh, A.R. *et al.* (2019) Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. *Nat. Med.* 25, 1415–1421
46. Sutton, R.S. and Barto, A.G. (2018) *Reinforcement Learning: an Introduction* (2nd edn), MIT Press
47. Francois-Lavet, V. *et al.* (2018) An introduction to deep reinforcement learning. *Found. Trends Mach. Learn.* 11, 219–354
48. Tseng, H.H. *et al.* (2017) Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med. Phys.* 44, 6690–6705
49. Liu, N. *et al.* (2019) Learning the dynamic treatment regimes from medical registry data through deep Q-network. *Sci. Rep.* 9, 1495
50. Yauney, G. and Shah, P. (2018) Reinforcement learning with action-derived rewards for chemotherapy and clinical trial dosing regimen selection. *Mach. Learn. Healthc. Conf.* 2018, 161–226
51. Gottesman, O. *et al.* (2019) Guidelines for reinforcement learning in healthcare. *Nat. Med.* 25, 16–18
52. Yu, C. *et al.* (2019) Reinforcement learning in healthcare: a survey. *arXiv* 1908.08796
53. Levine, S. *et al.* (2020) Offline reinforcement learning: tutorial, review, and perspectives on open problems. *arXiv* 2005.01643
54. Bertsimas, D. and Wiberg, H. (2020) Machine Learning in oncology: methods, applications, and challenges. *JCO Clin. Cancer Informa.* 4, 885–894
55. Rieke, N. *et al.* (2020) The future of digital health with federated learning. *NPJ Digit Med.* 3, 119
56. Xu, J. *et al.* (2020) Federated learning for healthcare informatics. *J. Healthc. Inform. Res.* 1–19
57. Stetson, P.D. *et al.* (2020) When predictive models collide. *Jco Clin. Cancer Informa.* 4, 547–550