

# Chapter 12

## Machine Learning in Personalized Anemia Treatment

**Adam E. Gaweda**  
*University of Louisville, USA*

### **ABSTRACT**

*This chapter presents application of reinforcement learning to drug dosing personalization in treatment of chronic conditions. Reinforcement learning is a machine learning paradigm that mimics the trial-and-error skill acquisition typical for humans and animals. In treatment of chronic illnesses, finding the optimal dose amount for an individual is also a process that is usually based on trial-and-error. In this chapter, the author focuses on the challenge of personalized anemia treatment with recombinant human erythropoietin. The author demonstrates the application of a standard reinforcement learning method, called *Q*-learning, to guide the physician in selecting the optimal erythropoietin dose. The author further addresses the issue of random exploration in *Q*-learning from the drug dosing perspective and proposes a “smart” exploration method. Finally, the author performs computer simulations to compare the outcomes from reinforcement learning-based anemia treatment to those achieved by a standard dosing protocol used at a dialysis unit.*

### **INTRODUCTION**

Pharmacological treatment of chronic illnesses often resembles a trial-and-error process. A physician usually initiates the treatment with a standard dose and observes the patient for a response and / or dangerous side effects. If the response is not sufficient, the dose is increased. On the other hand, if a side effect occurs, the dose is decreased. A standardized dosing strategy for a specific drug is usually established during Phase II of a clinical trial. However, such a standardized strategy is usually based on population dose-response characteristics and may not lead to optimal outcomes on individual basis.

DOI: 10.4018/978-1-60566-766-9.ch012

Optimal outcomes are more likely to be achieved if the dosing is adjusted over time based on the individual response. This is especially the case for chronic illnesses, such as anemia of End Stage Renal Disease (ESRD).

Anemia of ESRD is one of the common conditions in patients receiving hemodialysis (Eschbach and Adamson, 1985). It is caused by insufficient red blood cell production due to unavailability of a hormone called erythropoietin, produced primarily by the kidneys. Untreated, anemia leads to a number of conditions including cardiovascular disease (Harnett et al., 1995), as well as decreased quality of life (Wolcott et al., 1989) and increased mortality (Lowrie et al., 1994). External administration of recombinant human erythropoietin (rHuEPO) is a standard form of treatment for the anemia of ESRD. In the United States, the anemia treatment is governed by the National Kidney Foundation guidelines which recommend that the Hemoglobin (Hgb), a surrogate marker of red blood cell level, in patients receiving rHuEPO be maintained between 11 and 12 g/dL. Based on these recommendations, dialysis units develop their own sets of dose adjustment rules governing rHuEPO dosages in relation to Hgb levels. However, these rules are again often derived from population outcome measures and do not always provide the optimal outcome. For example, it is known that at a given point in time 2/3 of the ESRD patients in the U.S. are outside of the NKF recommended target range (Lacson et al., 2003).

To facilitate the personalization of drug administration, we have proposed the application of a Machine Learning paradigm called Reinforcement Learning to anemia treatment (Gaweda et al., 2005). Reinforcement Learning mimics the goal-oriented skill acquisition performed by humans (Dayan and Balleine, 2002). It represents an intelligent system, referred to as an “agent,” which discovers what actions to take through interaction with its environment, by trying them and receiving rewards, in order to drive the environment to a “goal” state. Agent’s choices may affect the environment not only immediately, but also over the long term. The first reported attempt at using Reinforcement Learning for drug dosing has been described by Sinzinger and Moore (2005). They show that a Reinforcement Learning based control of patient sedation performs better than conventional closed-loop controllers, such as PID.

In (Gaweda et al., 2005), we applied a Reinforcement Learning method called  $Q$ -learning, to anemia treatment in an on-line fashion. We concluded that in its simplest form this method’s performance is comparable to that of a rHuEPO dosing protocol used at our dialysis unit. However, a pure trial-and-error based learning was found insufficient to warrant any further improvement in the primary outcome, defined as maintaining the Hgb within the 11-12 g/dL range. To alleviate this drawback, we focused our attention on adding an element of heuristic supervision to guide the  $Q$ -learning (Gaweda et al., 2007). Others (Martin-Guerrero et al., 2006) have also applied the concept Reinforcement Learning to anemia treatment in an off-line fashion, using  $Q$ -learning to elicit optimal dosing strategy for rHuEPO from historical data records.

One of the key issues of Reinforcement Learning is the trade-off between exploration and exploitation (Sutton and Barto, 1998). To maximize the reward, the agent should take actions that are known to result in large rewards (exploitation). However, to discover such actions, the agent must try actions that it has not selected before (exploration). Our efforts so far, as well as those by Martin-Guerrero et al. (2006) focused on a purely exploiting form of Reinforcement Learning. An obvious drawback of the exploiting learning is the likelihood of achieving suboptimal results.

In this chapter, building upon the work presented in (Gaweda et al., 2007), we describe a Reinforcement Learning approach to anemia treatment which combines the guided  $Q$ -learning with the exploratory component. We first describe the issue of random exploration and its potential pitfalls in the case of drug dosing. To eliminate the risk involved in exploring undesired dose amounts, we propose the concept of

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/machine-learning-personalized-anemia-treatment/36989](http://www.igi-global.com/chapter/machine-learning-personalized-anemia-treatment/36989)

## Related Content

---

### User Consumption Behavior Recognition Based on SMOTE and Improved AdaBoost

Huijuan Hu, Dingju Zhu, Tao Wang, Chao He, Juel Sikderand Yangchun Jia (2022). *International Journal of Software Science and Computational Intelligence* (pp. 1-20).

[www.irma-international.org/article/user-consumption-behavior-recognition-based-on-smote-and-improved-adaboost/315302](http://www.irma-international.org/article/user-consumption-behavior-recognition-based-on-smote-and-improved-adaboost/315302)

### Music Content Analysis in MP3 Compressed Domain

Antonello D'Aguanno (2011). *Machine Learning Techniques for Adaptive Multimedia Retrieval: Technologies Applications and Perspectives* (pp. 301-321).

[www.irma-international.org/chapter/music-content-analysis-mp3-compressed/49114](http://www.irma-international.org/chapter/music-content-analysis-mp3-compressed/49114)

### Interval Mathematics as a Potential Weapon against Uncertainty

Hend Dawood (2014). *Mathematics of Uncertainty Modeling in the Analysis of Engineering and Science Problems* (pp. 1-38).

[www.irma-international.org/chapter/interval-mathematics-as-a-potential-weapon-against-uncertainty/94504](http://www.irma-international.org/chapter/interval-mathematics-as-a-potential-weapon-against-uncertainty/94504)

### A Particle Swarm Optimization Approach for Reuse Guided Case Retrieval

Nabila Nouaouria, Mounir Boukadoumand Robert Proulx (2014). *International Journal of Software Science and Computational Intelligence* (pp. 16-30).

[www.irma-international.org/article/a-particle-swarm-optimization-approach-for-reuse-guided-case-retrieval/127351](http://www.irma-international.org/article/a-particle-swarm-optimization-approach-for-reuse-guided-case-retrieval/127351)

### Symbiotic Aspects in e-Government Application Development

Claude Moulinand Marco Luca Sbodio (2010). *International Journal of Software Science and Computational Intelligence* (pp. 38-51).

[www.irma-international.org/article/symbiotic-aspects-government-application-development/39104](http://www.irma-international.org/article/symbiotic-aspects-government-application-development/39104)