

## 公募助成「腎不全病態研究助成」研究サマリー

研究名	強化学習による血液透析患者の赤血球造血刺激因子製剤処方の最適化
所属機関	新潟県立吉田病院
氏名	笠井 昭男

### 【背景】

腎性貧血は血液透析患者で最も多くみられる合併症であり、治療に赤血球造血刺激因子製剤(ESAs)が用いられる。しかし、患者間でその反応性に違いがあり、患者毎にもヘモグロビン(Hb)変動の原因となる。近年、強化学習の手法が様々な分野で応用されている。強化学習はエージェントが環境の中で、価値を最大化するような行動を学習するものである。ESAsの処方を最適化するにあたり、適した方法であると考えられる。

### 【方法】

#### 1. 強化学習の概要

強化学習は「環境」における価値を最大化するようにエージェントに学習させる。ここで、環境として赤血球产生モデルと鉄代謝モデルを統合したものを与える。

状態  $s$  は、環境の現在の状態を表し、ここでは、Hb 値、血清鉄、TSAT、フェリチン値になる。

行動  $a$  は、エージェントの取り得る行動を表し、ここでは、ESAs の投与量、経口鉄あるいは静注鉄を使用することになる。

報酬  $r$  は、エージェントがある状態で行動をおこした時に得られる報酬を表し、ここでは、 $10 \text{ g/dL} \leq \text{Hb} \leq 12 \text{ g/dL}$  で  $+1 \text{ g/dL}$  を最大とする釣鐘型曲線の報酬とし、 $\text{Hb} < 10 \text{ g/dL}$  では  $-1 \text{ g/dL}$  を最大とする釣鐘型曲線の報酬として、 $\text{Hb} > 12 \text{ g/dL}$  では  $-1 \text{ g/dL}$  を最大とする釣鐘型曲線の報酬とする。

行動状態価値  $Q(s, a)$  の学習には Neural fitted Q アルゴリズムを用いる。

#### 2. 数理モデル

強化学習の環境を設定するために赤血球生成と鉄代謝を統合した数理モデルが必要となる。

##### a) 赤血球产生モデル

赤血球は造血幹細胞(HSC)から CFU-GEMM、burst forming unit erythroid(BFU-E)、colony forming unit erythroid(CFU-E)、前赤芽球、赤芽球、脱核を経て網赤血球、成熟赤血球と分化する。これを機能的に BFU-E、CFU-E、赤芽球、骨髄網赤血球、成熟赤血球の5クラスに分類する。赤血球产生モデルの各クラスは年齢構造化モデルを使って、以下のような偏微分方程式モデルであらわされる。

$$\frac{\partial}{\partial t} u(t, \mu) + v(E(t)) \frac{\partial}{\partial \mu} u(t, \mu) = (\beta(E(t)) - \alpha(E(t), z(t), \mu)) u(t, \mu) \quad (1)$$

ここで、 $u(t, \mu)$  は時刻  $t$ 、成熟度  $\mu$  の人口密度関数、 $\beta$  : 増殖率、 $\alpha$  : アポトーシス率、 $E(t)$  : 時刻  $t$  でのエリスロポイエチン濃度、 $z(t)$  : 時刻  $t$  での鉄濃度、 $v(\cdot)$  : 成熟速度である。各クラスにあわせて、 $\beta$ 、 $\alpha$  などのパラメータを設定した。数値解は差分法により計算する。

##### b) 鉄代謝モデル

鉄代謝制御系として、ヘプシジンの量を決定するモデルを作る。ヘプシジン濃度は血清鉄、貯蔵鉄、酸素飽和度に依存し、炎症によって増加すると考えられる。ヘプシジンはフェロポルチンの働きを制御する。

鉄動態モデルとしては、経口鉄による腸管からの鉄吸収と静注鉄による鉄補充、血清鉄濃度、貯蔵鉄、赤血球系細胞の鉄、鉄リサイクル系、透析や採血による鉄喪失をモデル化していく。

現在、鉄代謝を常微分方程式を用いて、モデル化している。

赤血球产生モデルと鉄代謝モデルを統合して、患者データからパラメータ推定をおこない強化学習の環境とする。