EC.1. Table of Notation

Table EC.1 Table of Notation

Indices	
$i \in N$	Patients , where $N = \{1, ..., n\}$
$j \in \{1, ..., m\}$	Risk types
$k \in \{1, ..., p\}$	AFP observations
$t \in \{1, ..., T\}$	Time
Parameters	
P_{jj}, P_{jE}	Transition probability that a patient goes from state j to state j, or E, etc.
$\tau, \tilde{\tau}$	Transition matrices, for unscreened and screened patients, respectively
o_{jk}	Observation probability that a patient of type j gives observation k
Ω	Observation matrix
$\Omega_k, \Omega_E, \Omega_L$	The kth row of Ω diagonalized into a matrix
ρ	Reward vector
Variables	
$X_i^t(j)$	The belief that patient i is in state j at the beginning of time t
Y_i^t	The true health state of patient i at the beginning of time t unknown to the decision maker.
$[t]_s$	The patient chosen to be screened at time t
o_s^t	The observation received at time t
r_s^t	The reward received at time t

EC.2. Derivation of the Posterior Belief State X_i^{t+1} (Equation (2))

We wish to find X_i^{t+1}, given any previous belief state X_i^t, any action $[t]$ and any screening outcome o^t. For this, we need $X_i^{t+1}(j) = Pr(Y_i^{t+1} = j | [t], o^t, X_i^t)$, $\forall j$. We will divide this section into three cases, depending upon the action $[t]$ and the stochastic observation o^t. Because in our model, observations occur before transitions in each decision epoch, we denote t' to denote the belief immediately after an observation is received, but before transition occurs.

Case 1. $[t]=i$ **and** $k=1,...,p$ **(Screened patient and non-cancer reading)**

We wish to calculate

$$X_i^{t+1}(j)=Pr(Y_i^{t+1}=j|[t]=i,o^t=k,X_i^t) \quad \text{(EC.1)}$$

Expanding by conditioning on the belief at the intermediate time t', we get:

$$X_i^{t+1}(j)=\sum_{j'} Pr(Y_i^{t+1}=j|Y_i^{t'}=j',[t]=i,o^t=k,X_i^t)\cdot Pr(Y_i^{t'}=j'|[t]=i,o^t=k,X_i^t) \quad \text{(EC.2)}$$

Recognizing that the transition from t' to $t+1$ in EC.2 is independent of the action and observations, we get:

$$X_i^{t+1}(j)=\sum_{j'} Pr(Y_i^{t+1}=j|Y_i^{t'}=j')\cdot Pr(Y_i^{t'}=j'|[t]=i,o^t=k,X_i^t) \quad \text{(EC.3)}$$

Substituting for primitive data where applicable

$$X_i^{t+1}(j)=\sum_{j'} P_{j'j}\cdot Pr(Y_i^{t'}=j'|[t]=i,o^t=k,X_i^t) \quad \text{(EC.4)}$$

The remainder of this equation will be different, depending on 1 of 3 cases:

Case 1a. $j=1,...,m$ **(Probability of being cancer-free)**

Equation (EC.4) can be simplified because $P_{j'j}=0$ if $j\neq j'$

$$X_i^{t+1}(j)=P_{jj}\cdot Pr(Y_i^{t'}=j|[t]=i,o^t=k,X_i^t)$$

Applying Bayes' Law, we get:

$$X_i^{t+1}(j)=P_{jj}\cdot \frac{Pr(o^t=k|Y_i^{t'}=j,[t]=i,X_i^t)Pr(Y_i^{t'}=j|[t]=i,X_i^t)}{\sum_{\hat{j}} Pr(o^t=k|Y_i^{t'}=\hat{j},[t]=i,X_i^t)Pr(Y_i^{t'}=\hat{j}|[t]=i,X_i^t)}$$

Substituting for any known quantities:

$$X_i^{t+1}(j)=P_{jj}\cdot \frac{o_{jk}X_i^t(j)}{\sum_{\hat{j}} o_{\hat{j}k}X_i^t(\hat{j})} \quad \text{(EC.5)}$$

Case 1b. $j=E$ **(Probability of being in early-stage cancer)**

Containing from Equation (EC.4):

$$X_i^{t+1}(E) = \sum_{j'} P_{j'E} \cdot Pr(Y_i^{t'} = j' | [t] = i, o^t = k, X_i^t) \tag{EC.6}$$

The summands $j' = E, L$ can be removed from EC.6 because the state at t' cannot be E or L if an observation k was received

$$X_i^{t+1}(E) = \sum_{j'=1}^{m} P_{j'E} \cdot Pr(Y_i^{t'} = j' | [t] = i, o^t = k, X_i^t) \tag{EC.7}$$

Apply Bayes' Law to get:

$$X_i^{t+1}(E) = \sum_{j'=1}^{m} P_{j'E} \frac{Pr(o^t = k | X_i^{t'}(j) = j', [t] = i, X_i^t) Pr(Y_i^{t'} = j' | [t] = i, X_i^t)}{\sum_{\hat{j}} Pr(o^t = k | X_i^{t'}(j) = \hat{j}, [t] = i, X_i^t) Pr(Y_i^{t'} = \hat{j} | [t] = i, X_i^t)} \tag{EC.8}$$

Substituting for any known quantities:

$$X_i^{t+1}(E) = \frac{\sum_{j'}^{m} P_{j'E} o_{j'k} X_i^t(j)}{\sum_{\hat{j}} o_{\hat{j}k} X_i^t(\hat{j})} \tag{EC.9}$$

Case 1c. $j = L$ **(Probability of being in late-stage cancer)**

Continuing from Equation (EC.4):

$$X_i^{t+1}(L) = \sum_{j'} P_{j'L} \cdot Pr(Y_i^{t'} = j' | [t] = i, o^t = k, X_i^t) \tag{EC.10}$$

The summands $j' = E, L$ can be removed from EC.10 because the state at t' cannot be E or L if an observation k was received

$$X_i^{t+1}(L) = \sum_{j'=1}^{m} P_{j'L} \cdot Pr(Y_i^{t'} = j' | [t] = i, o^t = k, X_i^t) \tag{EC.11}$$

Applying Bayes' Law, we get:

$$X_i^{t+1}(L) = \sum_{j'}^{m} P_{j'L} \frac{Pr(o^t = k | X_i^{t'}(j) = j', [t] = i, X_i^t) Pr(Y_i^{t'} = j' | [t] = i, X_i^t)}{\sum_{\hat{j}} Pr(o^t = k | X_i^{t'}(j) = \hat{j}, [t] = i, X_i^t) Pr(Y_i^{t'} = \hat{j} | [t] = i, X_i^t)} \tag{EC.12}$$

Substituting for any known quantities:

$$X_i^{t+1}(L) = \frac{\sum_{j'}^{m} \delta o_{j'k} X_i^t(j)}{\sum_{\hat{j}} o_{\hat{j}k} X_i^t(\hat{j})} \tag{EC.13}$$

Therefore, Equations (EC.5), (EC.9), and (EC.13) can be summarized in vector form as follows. When $[t]=i$ observation $o^t=1,2,...,p$:

$$(X_i^{t+1}|[t]=i,o^t=k)=\left[\begin{array}{c|c} X_i^{t+1}(1) & \\ \vdots & \\ X_i^{t+1}(j) & \\ \vdots & [t]=i,o^t=k \\ X_i^{t+1}(m) & \\ X_i^{t+1}(E) & \\ X_i^{t+1}(L) & \end{array}\right]^T = \begin{bmatrix} \frac{P_{11}o_{1k}X_{i1t}}{\sum_{j=1}^m o_{jk}X_i^t(j)} \\ \vdots \\ \frac{P_{jj}o_{jk}X_i^t(j)}{\sum_{j=1}^m o_{jk}X_i^t(j)} \\ \vdots \\ \frac{P_{mm}o_{mk}X_i^t(m)}{\sum_{j=1}^m o_{jk}X_i^t(j)} \\ \frac{\sum_{j=1}^m P_{jE}o_{jk}X_i^t(j)}{\sum_{j=1}^m o_{jk}X_i^t(j)} \\ \frac{\sum_{j=1}^m \delta o_{jk}X_i^t(j)}{\sum_{j=1}^m o_{jk}X_i^t(j)} \end{bmatrix}^T = \overline{X_i^t\Omega_k}\tau_S$$

$$\forall k=1,...,p \tag{EC.14}$$

Case 2. $[t]=i$ **and** $k=\bar{E},\bar{L}$ **(Screened patient and cancer reading)**

The belief state update when $[t]=i$ and the observation $o^t=\bar{E}$ or $o^t=\bar{L}$ is far simpler:

$$X_i^{t+1}(j)=Pr(Y_i^{t+1}=j|[t]=i,o^t=\bar{E},X_i^t) \tag{EC.15}$$

Expanding by conditioning on the intermediate state at time t' we get:

$$X_i^{t+1}(j)=\sum_{j'}Pr(Y_i^{t+1}=j|Y_i^{t'}=j',[t]=i,o^t=\bar{E},X_i^t)\cdot Pr(Y_i^{t'}=j'|[t]=i,o^t=\bar{E},X_i^t) \tag{EC.16}$$

But when we see an observation $\bar{E}$, we know that the state at time t' is E with probability 1, so all but 1 of the summands are 0

$$X_i^{t+1}(j)=Pr(X_i^{t+1}=j|Y_i^{t'}=\bar{E},[t]=i,o^t=\bar{E},X_i^t) \tag{EC.17}$$

Lastly, we know from $\tilde{\tau}$ that a patient who is found to be in state E will replaced by new patient with probabilities q_j

$$X_i^{t+1}(j)=q_j \tag{EC.18}$$

Equation (EC.18) can be summarized in matrix form in Equation (EC.19):

$$(X_i^{t+1}|[t]=i,o^t=k)=\left[\begin{array}{c|c} X_i^{t+1}(1) & \\ \vdots & \\ X_i^{t+1}(j) & \\ \vdots & [t]=i,o^t=k \\ X_i^{t+1}(m) & \\ X_i^{t+1}(E) & \\ X_i^{t+1}(L) & \end{array}\right]^T = \begin{bmatrix} q_1 \\ \vdots \\ q_j \\ \vdots \\ q_m \\ 0 \\ 0 \end{bmatrix}^T = \begin{cases} \overline{X_i^t\Omega_E}\tau_S & \text{if } k=\bar{E} \\ \overline{X_i^t\Omega_L}\tau_S & \text{if } k=\bar{L} \end{cases} \tag{EC.19}$$

Case 3. $[t]\neq i$ **(Non-screened patient)**

On the other hand, the subsequent belief state of the unscreened patient, $[t]\neq i$ evolves regardless of the observation o^t, therefore:

$$(X_i^{t+1}|[t]\neq i,o^t=k)=\left[\begin{array}{c|c} X_i^{t+1}(1) & \\ \vdots & \\ X_i^{t+1}(j) & \\ \vdots & [t]\neq i,o^t=k \\ X_i^{t+1}(m) & \\ X_i^{t+1}(E) & \\ X_i^{t+1}(L) & \end{array}\right]^T = \begin{bmatrix} P_{11}X_i^t(1) \\ \vdots \\ P_{jj}X_i^t(j) \\ \vdots \\ P_{mm}X_i^t(m) \\ \sum_{j=1}^m P_{jE}X_i^t(j)+P_{EE}X_i^t(E) \\ \sum_{j=1}^m \delta X_i^t(j)+P_{EL}X_i^t(E)+1X_i^t(L) \end{bmatrix}^T = X_i^t\tau \tag{EC.20}$$

We can now combine Equations (EC.14), (EC.19) and (EC.20) into Equation (5) to tell us the subsequent belief state $X_i^{t+1}(j)$, given any previous belief state X_i^t, any action $[t]$ and any screening outcome o^t

$$X_i^{t+1} = \begin{cases} \overline{X_i^t \Omega_k \tilde{\tau}} & \text{if } [t] = i \text{ and } o^t = 1, ..., p \\ \overline{X_i^t \Omega_E \tilde{\tau}} & \text{if } [t] = i \text{ and } o^t = \bar{E} \\ \overline{X_i^t \Omega_L \tilde{\tau}} & \text{if } [t] = i \text{ and } o^t = \bar{L} \\ X_i^t \tau & \text{if } [t] \neq i \end{cases} \tag{EC.21}$$

EC.3. Deriving the Probability of Observations (Equation (3))

We wish to know the probability of any observation, given the current belief state $\boldsymbol{X^t}$ and a patient choice $[t] = i$:

$$Pr(o^t = k | [t] = i) \tag{EC.22}$$

Conditioning on the underlying state of the patient i, we get:

$$Pr(o^t = k | [t] = i) = \sum_j Pr(o^t = k | [t] = i, X_i^t = j) Pr(X_i^t = j | [t] = i) \tag{EC.23}$$

Substituting for any known quantities:

$$Pr(o^t = k | [t] = i) = \begin{cases} \sum_{j=1}^{m} o_{jk} X_i^t(j) & \text{for } k = 1, ..., p \\ X_i^t(E) & \text{for } k = \bar{E} \\ X_i^t(L) & \text{for } k = \bar{L} \end{cases} = \begin{cases} X_{[t]} \Omega_k \vec{1} & \text{for } k = 1, ..., p \\ X_{[t]} \Omega_E \vec{1} & \text{for } k = \bar{E} \\ X_{[t]} \Omega_L \vec{1} & \text{for } k = \bar{L} \end{cases} \tag{EC.24}$$

EC.4. Proof of Propositions 2 and 3

Recall Propositions 2 and 3:

Suppose $C^*(r)$ holds at time t. Then $C(r)$ also holds at time t.

Suppose $C^*(r)$ holds at time t. Then $C^*(r-1)$ holds at time t.

Proof of Proposition 2 Suppose $C^*(r)$ holds at time t. Then by definition, $\forall b \in \{1, ..., n\}, \exists a \in \{1, ..., n\}, b \neq a$ such that

$$X_a \alpha^* \geq \overline{X_b \Omega_k} \beta^* \tag{EC.25}$$

But we know by construction that $\alpha(r) \geq \alpha^*(r)$, $\forall r$ and $\beta^*(r) \geq \beta(r)$, $\forall r$ Therefore, it is a direct consequence that

$$X_a\alpha(r) \geq X_a\alpha^*(r) \geq \overline{X_b\Omega_k}\beta^*(r) \geq \overline{X_b\Omega_k}\beta(r) \tag{EC.26}$$

Therefore $C(r)$ holds.

Proof of Proposition 3 Suppose $C^*(r)$ holds at time t. Then by definition, $\forall b \in \{1, ..., n\}, \exists a \in \{1, ..., n\}, j \neq i$ such that

$$X_b\alpha^* \geq \overline{X_a\Omega_k}\beta^* \tag{EC.27}$$

We know that α^* is non-increasing, and β^* is non-decreasing. Therefore $\alpha^*(r-1) \geq \alpha^*(r)$ and $\beta^*(r) \geq \beta^*(r-1)$. Therefore, we can write the following chain of inequalities:

$$X_a\alpha^*(r-1) \geq X_a\alpha^*(r) \geq \overline{X_b\Omega_k}\beta^*(r) \geq \overline{X_b\Omega_k}\beta^*(r-1) \tag{EC.28}$$

Therefore $C^*(r-1)$ holds.

EC.5. Proof of Proposition 4

Recall Proposition 4:

Let $d_1, d_2, ...d_r \in \mathbb{R}^{(1\times n)}$ and $X_1, X_2, ..., X_n \in \mathbb{R}^{(n\times 1)}$ Then

$$\begin{aligned} &\max_{[1]\in N}\left\{d_1X_{[1]} + \max_{[2]\in N\setminus[1]}\left\{d_2X_{[2]} + ... \max_{\substack{[r]\in N\setminus \\ [1],[2],...,[r-1]}}\left\{d_rX_{[r]}\right\}\right\}\right\} \\ &= \max_{\{[1],...,[r]\}\subset N}\left\{d_1X_{[1]} + ... + d_rX_{[r]}\right\} \end{aligned} \tag{EC.29}$$

Furthermore, the same arguments maximize the left-hand side and right-hand side expression.

We prove this statement for any fixed n, and by induction on the number of summands r. Clearly this statement is true for the case of a single summand, because they equate the exactly the same expression.

Now suppose that this statement holds true for r summands.

$$\begin{aligned} &\max_{[r]\in N}\left\{d_rX_{[r]} + \max_{[r-1]\in N\setminus[r]}\left\{d_{r-1}X_{[r-1]} + ... \max_{\substack{[1]\in N\setminus \\ [2],...,[r]}}\left\{d_1X_{[1]}\right\}\right\}\right\} \\ &= \max_{\{[1],...,[r]\}\subset N}\left\{d_rX_{[r]} + ... + d_1X_{[1]}\right\} \end{aligned} \tag{EC.30}$$

We can use this inductive hypothesis to simplify the claim for $r+1$ summands

$$\max_{[r+1]\in N}\left\{d_{r+1}X_{[r+1]}+\max_{[r]\in N\setminus[r+1]}\left\{d_rX_{[r]}+\max_{\substack{[r-1]\in N\setminus\\ [r],[r+1]}}\left\{d_{r-1}X_{[r-1]}+\max_{\substack{[1]\in N\setminus\\ [2],...,[r],[r+1]}} d_1X_{[1]}\right\}\right\}\right\}$$
$$=\max_{[r+1]\in N}\left\{d_{r+1}X_{[r+1]}+\left[\max_{\{[1],...,[r]\}\subset N\setminus[r+1]}\{d_rX_{[r]}+...+d_1X_{[1]}\}\right]\right\} \tag{EC.31}$$

We can move the term $d_{r+1}X_{[r+1]}$ inside the inner max because it is not a function of the arguments being maximized.

$$=\max_{[r+1]\in N}\left\{\max_{\{[1],...,[r]\}\subset N\setminus[r+1]}\left[d_{r+1}X_{[r+1]}+d_rX_{[r]}+...+d_1X_{[1]}\right]\right\} \tag{EC.32}$$

We are now maximizing a single expression first over two sets of arguments. However the arguments have no intersection, so it is possible to evaluate both simultaneously.

$$=\max_{\substack{[r+1]\in N\\ \{[1],...,[r]\}\subset N\setminus[r+1]}}\left\{d_{r+1}X_{[r+1]}+d_rX_{[r]}+...+d_1X_{[1]}\right\} \tag{EC.33}$$
$$=\max_{\{[1],...,[r],[r+1]\}\subset N}\left\{d_{r+1}X_{[r+1]}+d_rX_{[r]}+...+d_1X_{[1]}\right\} \tag{EC.34}$$

EC.6. Proof of Proposition 5

Recall Proposition 5:

Let $d_1, d_2, ...d_r \in \mathbb{R}^{(1\times n)}$ *and* $X_1, X_2, ..., X_n \in \mathbb{R}^{(n\times 1)}$ *and Suppose* $\exists\lambda$ *such that every vector* $d_1, ..., d_{\lambda-1}$ *strictly dominates every vector* $d_\lambda, ..., d_r$ *in every component. Then the first* $[1], ..., [\lambda-1]$ *arguments maximize both of the following two expressions.*

$$\max_{\{[1],...,[r]\}\subset N}\left\{d_1X_{[1]}+...+d_{\lambda-1}X_{[\lambda-1]}+d_\lambda X_{[\lambda]}+...+d_rX_{[r]}\right\}$$
$$\max_{\{[1],...,[\lambda-1]\}\subset N}\left\{d_1X_{[1]}+...+d_{\lambda-1}X_{[\lambda-1]}\right\}$$

We prove this Proposition by considering the former problem:

$$\max_{\{[1],...,[r]\}\subset N}\left\{d_1X_{[1]}+...+d_{\lambda-1}X_{[\lambda-1]}+d_\lambda X_\lambda+...+d_rX_{[r]}\right\} \tag{EC.35}$$

Notice that if every vector $d_1, ..., d_{\lambda-1}$ strictly dominates every vector $d_\lambda, ..., d_r$ in every component, then this problem can be solved in two discrete stages without any loss of optimality. That is the first $1, ..., \lambda-1$ terms can be maximized without regard to the final $\lambda, ..., r$ terms.

EC.7. Proof of Superiority of Patient-by-Patient Screening

First, we prove that under certain conditions, it is advantageous to delay a reward by 1 period.

LEMMA EC.1. *If* $\sum_j P_{jE}X_i^t(j) + P_{EE}X_i^t(E) \geq X_i^t(E)$, *then* $\mathbb{E}\left[(r^{t+1}|[t+1]=i)\right] \geq \mathbb{E}\left[(r_t|[t]=i)\right]$

Proof of lemma

$$\mathbb{E}\left[(r^{t+1}|[t+1]=i)\right] = X_i^{t+1}\rho \quad \text{by definition} \tag{EC.36}$$

$$= X_i^t \tau \rho \quad \text{derived in (EC.21)} \tag{EC.37}$$

$$= \sum_j P_{jE}X_i^t(j) + P_{EE}X_i^t(E) \tag{EC.38}$$

$$\geq X_i^t(E) \quad \text{by supposition} \tag{EC.39}$$

$$= X_i^t \rho \tag{EC.40}$$

$$= \mathbb{E}\left[(r^t|[t]=i)\right] \quad \square \quad \text{by definition} \tag{EC.41}$$

Now we will prove that the value function for screening 1 patient per period is greater than or equal to screening 2 patients every 2 periods, provided that $\sum_j P_{jE}X_i^t(j) + P_{EE}X_i^t(E) \geq X_i^t(E)$. The proof for the suboptimality of screening S patients every S periods follows similarly.

Recall the value function under the policy of screening 1 patient per period from Equation (5).

$$V^t(\boldsymbol{X^t}) = \max_{i \in N}\left\{\mathbb{E}\left[r^t + V^{t+1}(\boldsymbol{X^{t+1}})\Big|[t]=i\right]\right\} \tag{EC.42}$$

The same value function will be true at time $t+1$:

$$V^{t+1}(\boldsymbol{X^{t+1}}) = \max_{j \in N}\left\{\mathbb{E}\left[r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}})\Big|[t+1]=j\right]\right\} \tag{EC.43}$$

Next, we substitute (EC.43) into (EC.42):

$$V^t(\boldsymbol{X^t}) = \max_{i \in N}\left\{\mathbb{E}\left[r^t + \max_{j \in N}\left\{\mathbb{E}\left[r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}})\big|[t+1]=j\right]\right\}\Big|[t]=i\right]\right\} \tag{EC.44}$$

This is still the recursive value function for screening 1 patient per period, but written in terms of every 2 periods (patient i at time t then patient j at time $t+1$).

We can move the term r^t inside of the maximization over j because it does not depend on j.

$$= \max_{i \in N} \left\{ \mathbb{E}\left[\max_{j \in N} \left\{ r^t + \mathbb{E}\left[r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j \right] \right\} \middle| [t] = i \right] \right\} \tag{EC.45}$$

The same term r^t can be moved inside the inner expectation operator $\mathbb{E}$, because it is an expectation over stochastic outcomes at time $t+1$. Therefore the term is deterministic with respect to the expectation.

$$= \max_{i \in N} \left\{ \mathbb{E}\left[\max_{j \in N} \left\{ \mathbb{E}\left[r^t + r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j \right] \right\} \middle| [t] = i \right] \right\} \tag{EC.46}$$

Notice that we have a maximization over $j \in N$ inside of an expectation operator $\mathbb{E}[\cdot]$. For any two independent random variables A, B, $\mathbb{E}\left[\max\{A, B\} \right] \geq \max\left\{ \mathbb{E}[A], \mathbb{E}[B] \right\}$, therefore:

$$\geq \max_{i \in N} \left\{ \max_{j \in N} \left\{ \mathbb{E}\left[\mathbb{E}\left[r^t + r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j \right] \middle| [t] = i \right] \right\} \right\} \tag{EC.47}$$

We can re-write the two maximizations into a single maximization because they are over the same shared expression.

$$= \max_{i,j \in N} \left\{ \mathbb{E}\left[\mathbb{E}\left[r^t + r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j \right] \middle| [t] = i \right] \right\} \tag{EC.48}$$

The two nested expectations $\mathbb{E}\left[\mathbb{E}[\cdot] \right]$ over stochastic outcomes at times t and $t+1$, respectively, can be re-defined into the expected value over two periods:

$$= \max_{i,j \in N} \left\{ \mathbb{E}\left[r^t + r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j, [t] = i \right] \right\} \tag{EC.49}$$

In this maximization, it is currently possible that $i = j$. If we restrict the decisions such that $i \neq j$, then the value function over a more constrained action set can only decrease. Therefore:

$$\geq \max_{\substack{i,j \in N \\ i \neq j}} \left\{ \mathbb{E}\left[r^t + r^{t+1} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t+1] = j, [t] = i \right] \right\} \tag{EC.50}$$

Lastly, the lemma above tells us that accruing the reward from patient j a period earlier results in a lower value.

$$\geq \max_{\substack{i,j \in N \\ i \neq j}} \left\{ \mathbb{E}\left[r^t + r^{t'} + V^{t+2}(\boldsymbol{X^{t+2}}) \middle| [t'] = j, [t] = i \right] \right\} \tag{EC.51}$$

We have used t' to indicate the second additional screening performed at time t. This final line is precisely the value function under the policy of screening two patients every two periods, and only updating beliefs every two periods. This chain of inequalities proves that screening 1 patient per period is superior to batch screening. □

We note that it is possible that the lemma which upholds the final inequality could be violated. However, consider that in our case study, a new patient always has $\Pr(Early) = 0$, and must be unscreened for 8 consecutive periods, or 240 days, before the patient first transitions into a belief state which violates the lemma's conditions. Any patient who has been screened within 8 periods, or 240 days, will satisfy the lemma's conditions.

Furthermore, the lemma is applied to an expression within a maximization over all patients j. Although the lemma's conditions may be violated for some patients j, the inequality is invalid only if both (1) the condition is violated for the patient j achieving the maximum on the left-hand side, as well as (2) the patient j' achieving the maximum on the right-hand side attains a strictly higher value. It would require an unusual sequence of outcomes for the belief state of the population to reach such a state.

Acknowledgments

The authors would like to acknowledge the National Science Foundation Graduate Research Fellowship Program, and the Seth Bonder Foundation, for their support of this work. We would like to thank Gabriel Zayas-Caban for his comments and review of this paper.