Electronic Companion
Integrated Anesthesiologist and Room Scheduling for Surgeries: Methodology and Application

Sandeep Rath
Kenan-Flagler Business School
University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599
sandeep@unc.edu

Kumar Rajaram
UCLA Anderson School of Management, Los Angeles CA 90095
kumar.rajaram@anderson.ucla.edu

Aman Mahajan
Department of Anesthesiology and Perioperative Medicine
David Geffen School of Medicine at UCLA, Los Angeles CA 90095
amahajan@mednet.ucla.edu

EC.1. The Assignment and Scheduling Decision Making Environment

Assignment and scheduling decisions at operating services department in the UCLA RRMC is complicated by the large number of Operating Rooms (ORs) and anesthesiologists, variety in surgical procedures, variability in anesthesiologist workload, and unpredictability in surgical durations. This section provides more details on these aspects. Table EC.1 shows the average number of procedures performed per day and the range in the number of surgeries across the various specialties. Figure EC.1 shows the uncertainty in the total anesthesiologist workload per day. The two most apparent sources of this variability are in the number of surgeries and the mix of specialties. However, Figure EC.2 shows that even if we control for these factors, there is still considerable variability in total workload per day. This variability arises because of differences across procedures within a specialty and because different patients react differently even to the same procedure (Schaefer et al. 2004). As a direct consequence, the required number of resources (i.e., rooms with anesthesiologists) at any instant as shown in Figure EC.3 and surgical durations are unpredictable. Specifically, the coefficient of variation of surgery duration across these 2700 types procedures varies from 0.75% to 125%.

EC.2. Proof of propositions

Proof of Proposition 1

By definition, the *IARSP* is said to have relatively complete recourse if there exists a feasible second stage solution for any first stage feasible solution (Thiele et al. 2009). Thus, we need to show that there exist feasible $S_i, Over_a, Over_r \forall i \in I, a \in A, r \in R$ satisfying (28) through (32) $\forall \mathbf{d} \in \boldsymbol{\mathcal{D}}(\tau)$ $\forall (\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}) \in K_1$, where K_1 is the feasibility set of the first stage problem.

Let $\bar{S}_i = \max_{a \in A} t_a^{start} + \sum_{j \in I - \{i\}} \left(\bar{d}_j + \hat{d}_j \right)$ define an upper bound on S_i, the start time of surgery i. This follows as $\bar{S}_i$ would be the start time of surgery i if all surgeries start after the start time of the last shift ($\max_{a \in A} t_a^{start}$), i is the last surgery to be performed in the day after all the other surgeries have been performed and every other surgery j lasts for its maximum allowed duration $\left(\bar{d}_j + \hat{d}_j\right)$. Thus, S_i is bounded from above. We define $M_{seq} = \max_{i\in I} \bar{S}_i + \max_{i\in I}\left(\bar{d}_j + \hat{d}_j\right)$.

Let $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \in K_1$ be some first stage feasible solution. We next show that $S_i = \max_{h\in I}\{s_h\} + \sum_{h\in I,h\neq i} u_{hi}d_h \;\forall i \in I$ is a feasible solution to (28). First observe that by our definition of M_{seq}, this constraint is deactivated when $u_{ij} = 0$. To check the feasibility when $u_{ij} = 1$ note from (12) that $u_{ji} = 0$. Substituting the values of S_i and S_i in (28), we get:

$$\max_{h\in I}\{s_h\} + \sum_{\substack{h\in I\\ h\neq j}} u_{hj}d_h \geq \max_{h\in I}\{s_h\} + \sum_{\substack{h\in I\\ h\neq i}} u_{hi}d_h + d_i - M_{seq}\left(1-u_{ij}\right) \qquad \forall i,j\in I \tag{EC.1}$$

Separating the term $u_{ij}d_i$ from the summation on the LHS and the term $u_{ji}d_j$ from the summation on the RHS leads to:

$$\max_{h\in I}\{s_h\} + \sum_{\substack{h\in I\\ h\neq i,j}} u_{hj}d_h + u_{ij}d_i \geq \max_{h\in I}\{s_h\} + \sum_{\substack{h\in I\\ h\neq i,j}} u_{hi}d_h + d_i + u_{ji}d_j - M_{seq}\left(1-u_{ij}\right) \quad \forall i,j\in I \tag{EC.2}$$

Next, setting $u_{ij} = 1, u_{ji} = 0$ and simplifying the above expression we get,

$$\sum_{\substack{h\in I\\ h\neq i,j}} d_h\left(u_{hj} - u_{hi}\right) \geq 0 \tag{EC.3}$$

From (13) $u_{hj} \geq u_{hi} + u_{ij} - 1$. As $u_{ij} = 1$, this implies, $u_{hj} - u_{hi} \geq 0$. Therefore, $S_i = \max_{h\in I}\{s_h\} + \sum_{h\in I,h\neq i} u_{hi}d_h \;\forall i\in I$ is a feasible solution to (28) $\forall(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \in K_1$. Note that (29) is satisfied since $S_i \geq s_i$. Since there exists a feasible $S_i \;\forall i\in I$ in the recourse problem, then there will always exist $Over_a \geq 0 \;\forall a\in A$ and $Over_r \geq 0 \;\forall r \in R$ for any given $\mathbf{x},\mathbf{z}$ satisfying (30) through (32). Thus the *IARSP* has relatively complete recourse.

This result can be extended to the case where the integrality condition on $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ is relaxed. This extension will be used in the proof of Proposition 4. In preparation, let $IARSP^{LP}$ be the form of the *IARSP* with the integrality conditions on $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ relaxed and let the feasibility set of $IARSP^{LP}$ be denoted by K_1^{LP}. To prove that $ISRSP^{LP}$ has relatively complete recourse we show by mathematical induction that there exist feasible $S_i \geq 0, Over_a \geq 0, Over_r \geq 0 \;\forall i\in I, a\in A, r\in R \;\forall(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \in K_1^{LP}$ and $\forall \mathbf{d} \in \boldsymbol{\mathcal{D}}(\tau)$. The proof thus far implies that it is now sufficient to prove the existence of a feasible $S_i \forall i\in I$ that satisfies (28). We consider a base case and an induction step, to show this result.

Base case: Let there be two surgeries $I = \{1,2\}$ and also let there be some $(\mathbf{x},\mathbf{z},\mathbf{u},\mathbf{s}) \in K_1$ and $\mathbf{d} \in \boldsymbol{\mathcal{D}}(\tau)$. Further, let $M_{seq} = \sum_{i\in\{1,2\}}(\bar{d}_i + \hat{d}_i)$. Then the following is a feasible solution to (28) through (32),

$$S_1 = s_1 + s_2 + d_2 + M_{seq}(1-u_{12}) \tag{EC.4}$$

$$S_2 = s_1 + s_2 + M_{seq} \tag{EC.5}$$

Proof: Since $u_{12} \leq 1$, (29) is satisfied. For these two surgeries, (28) can be written as,

$$S_2 \geq S_1 + d_1 - M_{seq}(1 - u_{12}) \tag{EC.6}$$

$$S_1 \geq S_2 + d_2 - M_{seq}(1 - u_{21}) \tag{EC.7}$$

Substituting the value of S_1 and S_2 from (EC.4) and (EC.5) in (EC.6),

$$s_1 + s_2 + M_{seq} \geq s_1 + s_2 + d_1 + d_2 + M_{seq}(1 - u_{12}) - M_{seq}(1 - u_{12}) \tag{EC.8}$$

The above simplifies to $M_{seq} \geq d_i + d_2$, which is true by our choice of M_{seq}. Next, substituting the values of S_1 and S_2 in (EC.7),

$$s_1 + s_2 + d_2 + M_{seq}(1 - u_{12}) \geq s_1 + s_2 + M_{seq} + d_2 - M_{seq}(1 - u_{21}) \tag{EC.9}$$

This simplifies to $M_{seq} \geq M_{seq}(u_{12} + u_{21})$ which is true since $u_{12} + u_{21} \leq 1$ from (12). Therefore, the values of S_1 and S_2 satisfy (28) and $IARSP^{LP}$ has relatively complete recourse for $I = \{1, 2\}$.

Induction step: Let $IARSP^{LP}$ have relatively complete recourse for $I = \{1, 2, \ldots, k-1\}$ and for this I, $M_{seq} = M'$ and $S'_i \; \forall i \in I$ be a feasible solution. Thus,

$$S'_j \geq S'_i + d_i - M'(1 - u_{ij}) \qquad \forall i, j \in \{1, 2, \ldots, k-1\} \tag{EC.10}$$

We now prove that $IARSP^{LP}$ has relatively complete recourse for $I = \{1, 2, \ldots, k-1, k\}$.

Proof: Let $I = \{1, 2, \ldots, k\}$, $M_{seq} = 2M'$ and let there be some $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}) \in K_1^{LP}$ and $\mathbf{d} \in \boldsymbol{\mathcal{D}}(\tau)$.

We show that there exists feasible $S_i \; \forall i \in I$. To show this, we first provide a feasible solution for $S_i \; \forall i \in I - \{k\}$. Then we show that for any such feasible $S_i \; \forall i \in I - \{k\}$, there exists a feasible S_k. We define the following feasible solution for $i \in I - \{k\}$

$$S_i = S'_i + d_k + \max_{h \in I} \{s_h\} + M'(1 - u_{ik}) \qquad \forall i \in I - \{k\} \tag{EC.11}$$

All such S_i satisfy (29). To verify that they satisfy (28), we substitute the values of S_i and M_{seq} in (28).

$$S'_j + d_k + \max_{h \in I} \{s_h\} + M'(1 - u_{jk}) \geq S'_i + d_k + \max_{h \in I} \{s_h\} + M'(1 - u_{ik}) + d_i - 2M'(1 - u_{ij}) \qquad \forall i, j \in \{1, 2, \ldots, k-1\} \tag{EC.12}$$

Simplifying, we get,

$$S'_j \geq S'_i + d_i - M'(1 - u_{ij}) - M'(1 + u_{ik} - u_{jk} - u_{ij}) \tag{EC.13}$$

From (EC.10), $S'_j \geq S'_i + d_i - M'(1 - u_{ij})$. From (13), $1 + u_{ik} - u_{ij} - u_{jk} \geq 0$. Thus the above is always true. Next we show that for all such S_i given by (EC.11) there exists feasible S_k such that,

$$S_k \geq S_i + d_i - M_{seq}(1 - u_{ik}) \qquad \forall i \in I - \{k\} \tag{EC.14}$$

$$S_i \geq S_k + d_k - M_{seq}(1 - u_{ki}) \qquad \forall i \in I - \{k\} \tag{EC.15}$$

$$S_k \geq s_k \tag{EC.16}$$

Substituting the values of S_i from (EC.11) and $M_{seq} = 2M'$ we rewrite the above inequalities to get,

$$S_k \geq S_i' + d_i + d_k + \max_{h \in I}\{s_h\} - M'(1 - u_{ik}) \qquad \forall i \in I - \{k\} \qquad \text{(EC.17)}$$

$$S_k \leq S_i' + \max_{h \in I}\{s_h\} + M'(3 - 2u_{ki} - u_{ik}) \qquad \forall i \in I - \{k\} \qquad \text{(EC.18)}$$

For there to be a feasible S_k, the RHS of (EC.17) must be less than or equal to the RHS of (EC.18) i.e.,

$$S_i' + \max_{h \in I}\{s_h\} + M'(3 - 2u_{ki} - u_{ik}) \geq S_i' + d_k + \max_{h \in I}\{s_h\} + d_i - M'(1 - u_{ik}) \qquad \forall i \in I - \{k\} \qquad \text{(EC.19)}$$

Simplifying the above we get,

$$M'[4 - 2(u_{ki} + u_{ik})] \geq d_k + d_i \qquad \forall i \in I - \{k\} \qquad \text{(EC.20)}$$

As M' is large enough and $u_{ik} + u_{ki} \leq 1$, the above is true. We also need to check that the RHS of (EC.18) is greater the RHS of (EC.16), i.e.,

$$s_k \leq S_i' + \max_{h \in I}\{s_h\} + M'(3 - 2u_{ki} - u_{ik}) \qquad \forall i \in I - \{k\} \qquad \text{(EC.21)}$$

The above equation is always true as $u_{ik} + u_{ki} \leq 1$. Therefore there exists feasible S_i $\forall i \in I$ for $I = \{1, 2, \ldots, k\}$ and by principle of mathematical induction, $IARSP^{LP}$ has relatively complete recourse.

Proof of Proposition 2

For a given first stage solution $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s})$ and an arbitrary $\mathbf{d} \in \boldsymbol{\mathcal{D}}(\tau)$, the dual of the recourse function $\boldsymbol{\mathcal{R}}^{\mathbf{D}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}, \mathbf{d})$ is given by:

$$\boldsymbol{\mathcal{R}}^{\mathbf{D}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}, \mathbf{d}) = \max \Bigg\{ \sum_{i \in I} d_i \Bigg(\sum_{j \in I - \{i\}} \lambda_{ij} + \sum_{a \in A} \mu_{ia} + \sum_{r \in R} \theta_{ir} \Bigg) + \sum_{i \in I} s_i \phi_i - M_{seq} \sum_{i,j \in I i \neq j} \lambda_{ij}(1 - u_{ij})$$
$$- M_{anesth} \sum_{\substack{i \in I \\ a \in A}} \mu_{ia}(1 - x_{ia} + y_a) - M_{room} \sum_{\substack{i \in I \\ r \in R}} \theta_{ir}(1 - z_{ir}) - \sum_{\substack{i \in I \\ r \in R}} \theta_{ir} T^{end} - \sum_{\substack{i \in I \\ a \in A}} \mu_{ia} t_a^{end} \Bigg\}$$

subject to,

$$\sum_{i \in I} \mu_{ia} \leq c_{oa} \qquad \forall a \in A \qquad \text{(EC.22)}$$

$$\sum_{i \in I} \theta_{ir} \leq c_{or} \qquad \forall r \in R \qquad \text{(EC.23)}$$

$$\sum_{\substack{j \in I \\ j \neq i}} \lambda_{ij} - \sum_{\substack{j \in I \\ j \neq i}} \lambda_{ji} + \sum_{a \in A} \mu_{ia} + \sum_{r \in R} \theta_{ir} - \phi_i \geq 0 \qquad \forall i \in I \qquad \text{(EC.24)}$$

$$\lambda_{ij}, \mu_{ia}, \theta_{ir}, \phi_i \geq 0 \qquad \forall i, j \in I, a \in A, r \in R \qquad \text{(EC.25)}$$

Here, $\lambda_{ij}, \phi_i,\ \mu_{ia},\ \theta_{ir},\ \forall i, j \in I, a \in A, r \in R$ are dual variables corresponding to constraints (28)-(32) respectively. $\boldsymbol{\mathcal{Q}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s})$ can now be written as

$$\boldsymbol{\mathcal{Q}}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) = \max \left\{ \sum_{i\in I} \left(\bar{d}_i + f_i\hat{d}_i\right) \left(\sum_{j\in I-\{i\}} \lambda_{ij} + \sum_{a\in A}\mu_{ia} + \sum_{r\in R}\theta_{ir} \right) + \sum_{i\in I} s_i\phi_i - M_{seq} \sum_{i,j\in I i\neq j} \lambda_{ij}\left(1-u_{ij}\right) \right.$$
$$\left. - M_{anesth} \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}\left(1-x_{ia}+y_a\right) - M_{room}\sum_{\substack{i\in I\\ r\in R}} \theta_{ir}\left(1-z_{ir}\right) - \sum_{\substack{i\in I\\ r\in R}} \theta_{ir}T^{end} - \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}t_a^{end} \right\}$$

subject to,

$$(EC.22)-(EC.25)$$
$$\sum_{i\in I}|f_i| \le \tau \tag{EC.26}$$
$$-1\le f_i \le 1 \qquad \forall i\in I \tag{EC.27}$$

Simplifying the objective function,

$$\boldsymbol{\mathcal{Q}}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) = \max \left\{ \sum_{i\in I} \left(\bar{d}_i\pi_i + f_i\pi_i\hat{d}_i\right) + \sum_{i\in I} s_i\phi_i - M_{seq} \sum_{i,j\in I i\neq j} \lambda_{ij}\left(1-u_{ij}\right) \right.$$
$$\left. - M_{anesth} \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}\left(1-x_{ia}+y_a\right) - M_{room}\sum_{\substack{i\in I\\ r\in R}} \theta_{ir}\left(1-z_{ir}\right) - \sum_{\substack{i\in I\\ r\in R}} \theta_{ir}T^{end} - \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}t_a^{end} \right\}$$

subject to, $(EC.22)-(EC.27)$

$$\pi_i = \sum_{j\in I-\{i\}} \lambda_{ij} + \sum_{a\in A}\mu_{ia} + \sum_{r\in R}\theta_{ir} \qquad \forall i\in I \tag{EC.28}$$
$$\pi_i \ge 0 \qquad \forall i\in I \tag{EC.29}$$

$\boldsymbol{\mathcal{Q}}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ is a bilinear optimization problem because of the presence of bilinear terms $f_i\pi_i$ in the objective function. Further, at optimality $f_i^* \ge 0\ \forall i\in I$. This is true as $\pi_i \ge 0$ and the feasibility set of f_i is independent of π_i. If τ is a positive integer, constraints (EC.26) and (EC.27) constrain the feasibility set of $f_i\ \forall i \in I$ to be the set $\{0,1\}$. This implies that:

$$f_i\pi_i = \begin{cases} \pi_i, & \text{if } f_i = 1 \\ 0, & \text{if } f_i = 0 \end{cases} \qquad \forall i\in I$$

We introduce an additional variable $\xi_i = f_i\pi_i$ and rewrite $\boldsymbol{\mathcal{Q}}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ as follows:

$$\boldsymbol{\mathcal{Q}}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) = \max \left\{ \sum_{i\in I} \left(\bar{d}_i\pi_i + \xi_i\hat{d}_i\right) - M_{seq} \sum_{i,j\in I i\neq j} \lambda_{ij}\left(1-u_{ij}\right) + \sum_{i\in I} s_i\phi_i \right.$$
$$\left. - M_{anesth} \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}\left(1-x_{ia}+y_a\right) - M_{room}\sum_{\substack{i\in I\\ r\in R}} \theta_{ir}\left(1-z_{ir}\right) - \sum_{\substack{i\in I\\ r\in R}} \theta_{ir}T^{end} - \sum_{\substack{i\in I\\ a\in A}} \mu_{ia}t_a^{end} \right\}$$

subject to,

$$(EC.23)-(EC.29)$$

$$\xi_i \leq M_f f_i \qquad \forall i \in I \tag{EC.30}$$

$$\xi_i \leq \pi_i \qquad \forall i \in I \tag{EC.31}$$

$$\xi_i \geq 0 \qquad \forall i \in I \tag{EC.32}$$

Where, M_f is a sufficiently large positive number. Thus, $\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ can be written as a mixed-integer program.

Proof of Proposition 3

This proof is adapted from Thiele et al. (2009). For conciseness of notation we represent the set defined by (EC.22)-(EC.25) as $\Lambda(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$. Also let $\boldsymbol{\lambda}=(\lambda_{ij})\forall i,j\in I$, $\boldsymbol{\mu}=\mu_{ia}\forall i\in I, a\in A$, $\boldsymbol{\theta}=\theta_{ir}\forall i\in I, r\in R$, $\boldsymbol{\phi}=\phi_i\forall i\in I$. From the proof of Proposition 2 and strong duality, we get:

$$\begin{aligned}\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}) = \max_{\substack{(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\phi})\\ \in \\ \Lambda(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})}} \Bigg\{ &\sum_{i\in I} d_i \left(\sum_{j\in I-\{i\}} \lambda_{ij} + \sum_{a\in A}\mu_{ia} + \sum_{r\in R}\theta_{ir}\right) + \sum_{i\in I} s_i\phi_i - M_{seq}\sum_{i,j\in I i\neq j}\lambda_{ij}(1-u_{ij}) \\ &- M_{anesth}\sum_{\substack{i\in I\\ a\in A}}\mu_{ia}(1-x_{ia}+y_a) - M_{room}\sum_{\substack{i\in I\\ r\in R}}\theta_{ir}(1-z_{ir}) - \sum_{\substack{i\in I\\ r\in R}}\theta_{ir}T^{end} - \sum_{\substack{i\in I\\ a\in A}}\mu_{ia}t_a^{end}\Bigg\}\end{aligned} \tag{EC.33}$$

Let $\boldsymbol{d}^l \in \arg\max_{\boldsymbol{d}\in\mathcal{D}(\tau)} \mathcal{R}(\mathbf{x}^l,\mathbf{y}^l,\mathbf{z}^l,\mathbf{u}^l,\mathbf{s}^l,\mathbf{d})$. From equation (24), $\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) = \max_{\mathbf{d}\in\mathcal{D}(\tau)} \mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d})$. Then for an arbitrary $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$,

$$\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \geq \mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}^l) \tag{EC.34}$$

Furthermore, let $(\boldsymbol{\lambda}^l,\boldsymbol{\mu}^l,\boldsymbol{\theta}^l,\boldsymbol{\phi}^l)$ be an optimal solution of (EC.33) with $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}) = (\mathbf{x}^l,\mathbf{y}^l,\mathbf{z}^l,\mathbf{u}^l,\mathbf{s}^l,\mathbf{d}^l)$. From (EC.33) we get:

$$\begin{aligned}\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}^l) \geq &\sum_{i\in I} d_i^l \left(\sum_{j\in I-\{i\}} \lambda_{ij}^l + \sum_{a\in A}\mu_{ia}^l + \sum_{r\in R}\theta_{ir}^l\right) + \sum_{i\in I} s_i\phi_i^l - M_{seq}\sum_{i,j\in I i\neq j}\lambda_{ij}^l(1-u_{ij}) \\ &- M_{anesth}\sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l(1-x_{ia}+y_a) - M_{room}\sum_{\substack{i\in I\\ r\in R}}\theta_{ir}^l(1-z_{ir}) - \sum_{\substack{i\in I\\ r\in R}}\theta_{ir}^l T^{end} - \sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l t_a^{end}\end{aligned} \tag{EC.35}$$

From, equation (EC.34) and (EC.35) we get,

$$\begin{aligned}\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \geq &\sum_{i\in I} d_i^l \left(\sum_{j\in I-\{i\}} \lambda_{ij}^l + \sum_{a\in A}\mu_{ia}^l + \sum_{r\in R}\theta_{ir}^l\right) + \sum_{i\in I} s_i\phi_i^l - M_{seq}\sum_{i,j\in I i\neq j}\lambda_{ij}^l(1-u_{ij}) \\ &- M_{anesth}\sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l(1-x_{ia}+y_a) - M_{room}\sum_{\substack{i\in I\\ r\in R}}\theta_{ir}^l(1-z_{ir}) - \sum_{\substack{i\in I\\ r\in R}}\theta_{ir}^l T^{end} - \sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l t_a^{end}\end{aligned} \tag{EC.36}$$

From (EC.28) and (EC.36) we get:

$$\begin{aligned}\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) \geq \sum_{i\in I}\left(\bar{d}_i\pi_i^l + \xi_i^l\hat{d}_i\right) + \sum_{i\in I} s_i\phi_i^l - M_{seq}\sum_{i,j\in I i\neq j}\lambda_{ij}^l\left(1-u_{ij}\right)\\ -M_{anesth}\sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l\left(1-x_{ia}+y_a\right) - M_{room}\sum_{\substack{i\in I\\ r\in R}}\theta_{ir}^l\left(1-z_{ir}\right) - \sum_{\substack{i\in I\\ r\in R}}\theta_{ir}T^{end} - \sum_{\substack{i\in I\\ a\in A}}\mu_{ia}^l t_a^{end}\end{aligned} \tag{EC.37}$$

Therefore, the right hand side of (EC.37) is a lower bound on $\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$.

Proof of Proposition 4

The value function of a linear program $z(b) = \min\{c^T x | Ax \geq b\}$ is a piecewise linear convex function over the domain for which the linear program is feasible (Martin 1999)[Corollary 2.49, pp. 75].

From the proof of Proposition 1, for any $\mathbf{d}\in\mathcal{D}(\tau)$, if K_1^{LP} is the feasibility set of the first stage problem with the integrality condition on $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ relaxed and $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})\in K_1^{LP}$, then $\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d})$ is a feasible linear program over its domain as $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ appear in the right hand side of its constraints. Thus, from above, $\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d})$ is a piecewise linear convex function in $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ for any given $\mathbf{d}\in\mathcal{D}(\tau)$. From equation (24), $\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s}) = \max_{\mathbf{d}\in\mathcal{D}(\tau)}\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d})$. Therefore, $\mathcal{Q}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$ is also piecewise-linear and convex in $(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s})$.

EC.3. Sample Average Approximation Solution for the Integrated Anesthesiologist and Room Scheduling Problem

In this section we provide the formulation and a short description of sample average approximation procedure based solution to solve the integrated anesthesiologist and room scheduling problem.

The variable and parameter description is as described in §3 in the paper. The difference between the sample average approximation formulation and the robust optimization formulation is that the surgery duration $d_i\ \forall i\in I$, instead of belonging to an uncertainty set are now modeled as random variables and are denoted by $d_i(\omega)$ under scenario $\omega\in\Omega$. For $|I|$ surgeries the random vector $\mathbf{d}(\omega) = \{d_1(\omega), d_2(\omega),\ldots,d_{|I|}(\omega)\}$ is the vector of surgery durations under scenario $\omega\in\Omega$. The support of $\mathbf{d}$ is $\mathbb{R}_+^{|I|}$. Following the evidence found in clinical literature (Strum et al. 2012) and standard assumptions in surgery scheduling literature (Batun et al. 2011, Denton et al. 2010) we assumed that these d_i are independent and have log-normal distribution. The standard formulation of the resulting two-stage stochastic program with recourse is:

$$\min\left\{\sum_{r\in R} c_r v_r + \sum_{a\in A} c_q y_a + \mathbb{E}\left[\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}(\omega))\right]\right\} \tag{EC.38}$$

subject to,

$$(2) - (23)$$

and,

$$\mathcal{R}(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{s},\mathbf{d}(\omega)) = \min\left\{\sum_{a\in A} c_{oa} Over_a(\omega) + \sum_{r\in R} c_{or} Over_r(\omega)\right\} \tag{EC.39}$$

subject to,

$$S_j(\omega) \geq S_i(\omega) + d_i(\omega) - M_{seq}(1 - u_{ij}) \qquad \forall i, j \in I \quad \text{(EC.40)}$$

$$S_i(\omega) \geq s_i \qquad \forall i \in I \quad \text{(EC.41)}$$

$$Over_a(\omega) \geq S_i(\omega) + d_i(\omega) - t_a^{end} - M_{anesth}(1 - x_{ia} + y_a) \qquad \forall i \in I, a \in A \quad \text{(EC.42)}$$

$$Over_r(\omega) \geq S_i(\omega) + d_i(\omega) - T^{end} - M_{room}(1 - z_{ir}) \qquad \forall i \in I, r \in R \quad \text{(EC.43)}$$

$$S_i(\omega), Over_a(\omega), Over_r(\omega) \geq 0 \qquad \forall i \in I, a \in A, r \in R \quad \text{(EC.44)}$$

Assuming a log-normal distribution, we estimate the conditional mean and standard deviation for each cluster of CPT codes. From this estimated distribution, we draw $N_s = 1000$ samples and formulate the sample average approximation of the two-stage stochastic program as:

$$\min \left\{ \sum_{r \in R} c_r + \sum_{a \in A} c_q y_a + \frac{1}{N_s} \sum_{n=1}^{N_s} \boldsymbol{\mathcal{R}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}, \mathbf{d}^n) \right\} \quad \text{(EC.45)}$$

subject to,

$$(2) - (23)$$

and,

$$\boldsymbol{\mathcal{R}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{s}, \mathbf{d}^n) = \min \left\{ \sum_{a \in A} c_{oa} Over_a^n + \sum_{r \in R} c_{or} Over_r^n \right\} \quad \text{(EC.46)}$$

subject to,

$$S_j^n \geq S_i^n + d_i^n - M_{seq}(1 - u_{ij}) \qquad \forall i, j \in I \quad \text{(EC.47)}$$

$$S_i^n \geq s_i \qquad \forall i \in I \quad \text{(EC.48)}$$

$$Over_a^n \geq S_i^n + d_i^n - t_a^{end} - M_{anesth}(1 - x_{ia} + y_a) \qquad \forall i \in I, a \in A \quad \text{(EC.49)}$$

$$Over_r^n \geq S_i^n + d_i^n - T^{end} - M_{room}(1 - z_{ir}) \qquad \forall i \in I, r \in R \quad \text{(EC.50)}$$

$$S_i^n, Over_a^n, Over_r^n \geq 0 \qquad \forall i \in I, a \in A, r \in R \quad \text{(EC.51)}$$

We solve the above program by the integer L-shaped decomposition framework (Birge & Louveaux 1997). The results of the solution procedure are provided in Table 6 of the main paper.

EC.4. Quantile Regression Procedure for Estimating Uncertainty Set $\boldsymbol{\mathcal{D}}(\tau)$

We develop a quantile regression model for predicting $g_L(\tilde{\mathbf{b}}; \rho)$ and $g_U(\tilde{\mathbf{b}}; \rho)$, which are used in (45) to estimate the uncertainty set $\boldsymbol{\mathcal{D}}(\tau)$. As the objective of this model was to predict quantiles, we prefer a parsimonious model with predictive power rather than an over-specified model. For evaluation of the predictive power of this model we use an out of sample Mean Square Prediction Error (MSPE) with respect to the conditional median. To avoid over-specification of the model, we use an Akaike Information Criterion (AIC). The dataset Δ^E had 25700 observations. Δ^E was divided into two disjoint datasets $\Delta^{E-Train}$ (19300 observations) and Δ^{E-Test} (6400 observations). The quantile regression was built on $\Delta^{E-Train}$ and the out-of-sample MSPE was performed on Δ^{E-Test}. The regression model was built using the following steps:

Step 1: Dimension Reduction. Due to the large number of CPT codes and surgeons compared to the number of observations, we performed clustering to reduce the number of factor variables corresponding to CPT codes and surgeons. For surgeries with multiple CPT codes, we concatenated the CPT codes to create a composite code. For example, if a procedure consisted of CPT codes *A, B* and *C*, we created the composite code *A-B-C*. After this procedure the total number of unique surgery types grew from 2706 to 4061 as there are now more unique procedures with each combination being a different code. Subsequently we performed a *k-means* clustering on the median values of observed surgery duration for each unique procedure code. This clustering procedure is similar to that performed in He et al. (2012). We plot the number of clusters against the % variance explained and choose the elbow point for selecting the total number of clusters. We perform similar clustering procedure for the surgeons. From Figure EC.4 and EC.5 we choose 6 clusters for both CPT codes and surgeons. This choice of clusters explained 95% of the variance in median surgical durations. We use these clustered variables in the subsequent quantile regression and name them $CPTCLUST$ for the CPT clusters and $PROVCLUST$ for the surgeons.

Step 2: Quantile Regression From the dataset Δ^{E}, there are 8 possible explanatory variables that could be included for modeling surgical durations. However, using the CPT cluster variable ($CPTCLUST$) and surgery service variable (*SERVICE*) at the same time led to an ill-defined model matrix. This was due to the collinearity between these two variables as measured by the Variance Inflation Factor (VIF) as described in Hair et al. (2006). The VIF for *SERVICE* was 40.19. We removed this variable and found that none of the other variables demonstrated a $VIF > 10$, thus meeting the criteria for not exhibiting significant collinearity (Table EC.2).

Subsequently we performed quantile regression incorporating the above 7 variables. We then sequentially remove variables in increasing order of significance and compare the out-of-sample MSPE and the AIC for the resulting models. The results are shown in Table EC.3. Incorporating surgeon information (i.e., $PROVCLUST$), CPT information (i.e., $CPTCLUST$), age and ASA score improves the prediction and leads to a lower AIC value. Therefore, these variables were incorporated in the model. The number of CPT codes (NUMCPT) was excluded from the final model specification as adding it did not lead to any improvement in MSPE or AIC. The final specification of the quantile regression model was thus,

$$ACTUALHRS = BOOKEDHRS + ASA + AGE + PATCLASS + CPTCLUST + PROVCLUST \tag{EC.52}$$

Next, in Table EC.4 we provide the coefficients of the above quantile regression model at several quantile values (10^{th}, 30^{th}, 60^{th} and 90^{th} quantile). This table demonstrates an advantage of using quantile regression instead of OLS regression. In particular, it allows for the effect of variable to be different at different quantiles. The coefficient of BOOKEDHRS, representing the surgeons' estimate of surgery duration, increases with increasing quantiles of surgery duration. This implies that more weight is provided to surgeons estimate for longer surgeries. This could imply that surgeons provide a more accurate estimate for longer surgeries than for shorter surgeries. One reason for this could be that surgeons often round up or down to the nearest quarter of an hour while providing estimates. This rounding off would lead to more significant differences for shorter surgeries than for longer surgeries. Another variable that demonstrates changing coefficients with quantiles is ASA. The coefficient of the ASA code increases with increasing quantiles. This could imply that for longer surgeries each increase in ASA score contributes more to the surgery duration than for shorter surgeries. This is also intuitive as the fitness of the patient before surgery would be more significant for longer and more complex procedures.

EC.5. The Decision Support System at the UCLA RRMC

The decision support system for integrated anesthesiologist and operating room scheduling at the UCLA RRMC was built using the Python programming language and a schematic of this system is shown in Figure EC.6. This system consists of an uncertainty set estimation module and an optimization module. The surgery characteristics were provided using the CareConnect database at the hospital and inputted to the uncertainty set estimation module. The output from this module was sent along with anesthesiologist availability, and the surgery resource specialty information provided by the Qgenda database to the optimization module. This solved the IARSP using the model based heuristic and generated an optimized schedule for the following day. This specified the assignment decisions of anesthesiologists and operating room to surgeries along with their scheduled start times. The output of this module was provided to the planner who made adjustments as needed to accommodate special requests by surgeons to change start times or for rooms with additional specialized equipment or for specific anesthesiologists.

There were several challenges in implementing this system. First, the CareConnect database and the Qgenda database had to be accessed daily, and their output had to be reformatted to be compatible with the uncertainty set estimation and the optimization modules. This necessitated the development of a specialized automated interface which required regular maintenance. Second, the planners were initially skeptical about the ability of our system to consider all specialties and constraints. Further, they were unsure if the prediction of surgical durations was better than estimates made by the surgeons, and if the model had adequately captured uncertainty in surgical durations. They felt that if these aspects were not effectively incorporated, this could lead to schedule disruptions, unsatisfied patients and extensive overtime costs. To ensure that the planners were confident with this system, we ran the model in parallel with their approach as described in §5.2 of the paper, so that they could understand the solution of the model and compare this with their own rules. They were reassured that the model solution corresponded to their solution when the number of daily surgeries was small. However, they also appreciated how the model solution outperformed their rules when the number of daily surgeries was high and the resulting scheduling complexity was larger. In this scenario, the model was more effective in utilizing the flexibility of the multiple parallel resources and considering uncertainty in surgical durations. A third challenge in implementing the system was that surgical characteristics had to be updated and new procedures added to the system. Since these required specialized clinical input, a formalized procedure had to be instituted, where the feedback of the surgeons and anesthesiologists had to be solicited and manually updated in the CareConnect database. This was time consuming and had to be done in monthly basis. However, this was essential to ensure the continued efficacy of our system.

EC.6. Evaluation of the Prediction and Scheduling Benefits of the Model Based Heuristic

In this section, we analyze what proportion of the gains were due to a better prediction method for surgical durations and how much was due to the scheduling policy. To conduct this analysis, we used the estimates of surgical durations made by the surgeons and ran our model for the live and historical validation. Note that the resulting gains would now be entirely from the scheduling policy. We then subtracted these gains from the original benefits which considered prediction and scheduling to calculate the gains of the prediction method. This analysis showed that on average 41% of the benefits were due to better prediction and 59% was due a better scheduling policy. These results demonstrate that due to the complexity of the problem, the most benefits can be got by combining prediction and optimization, an aspect for which robust optimization is particularly well suited.

References

Batun S, Denton BT, Huschka TR, Schaefer AJ. 2011. Operating room pooling and parallel surgery processing under uncertainty. *INFORMS J. Comput.* 23(2) 220-237.

Birge, JR, Louveaux F. 1997. *Introduction to Stochastic Programming*. Springer, New York.

Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR. 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Oper. Res.* 58(4-part-1) 802816.

He B, Dexter F, Macario A, Zenios S. 2012. The timing of staffing decisions in hospital operating rooms: incorporating workload heterogeneity into the newsvendor problem. *Manufacturing Service Oper. Management* **14**(1) $99-114$.

Schaefer, AJ, Bailey MD, Shechter SM, and Roberts MS. "Modeling medical treatment using Markov decision processes." *Operations research and health care*, pp. 593-612. Springer US, 2005.

Strum, DP, May JH, Vargas LG. 2000. Modeling the uncertainty of surgical procedure times: Comparison of log-normal and normal models. *Anesthesiology* **92**(4) 1160-1167.

Table EC.1 Number of surgeries across specialties

Surgery Specialty	Average Number of Surgeries per day	Range of surgeries per day
Vascular	1.75	0-10
Neuro	5.97	0-18
Plastics	6.01	0-15
ENT	1.53	0-6
Urology	3.23	0-15
Liver	5.15	0-10
Thoracic	1.66	0-12
Cardiac	7.07	1-12
Trauma	1.02	0-7
Pediatric	1.34	0-6
Eye Surgery	1.44	0-5
General	26.34	0-62

Table EC.2 Variance Inflation Factors

Variable	VIF
BOOKEDHRS	3.22
ASA	1.23
AGE	1.15
PATCLASS	1.76
CPTCLUST	4.96
PROVCLUST	3.19
NUMCPT	1.04

Table EC.3 Results of independent variable selection for quantile regression

	Dependent variable:				
	ACTUALHRS				
	(1)	(2)	(3)	(4)	(5)
BOOKEDHRS	0.206***	0.206***	0.204***	0.206***	0.210***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
ASA	0.044***	0.044***	0.036***		
	(0.009)	(0.009)	(0.009)		
AGE	−0.001***	−0.001***			
	(0.0002)	(0.0003)			
NUMCPT	−0.001*				
	(0.0003)				
CPTCLUST2	−4.966***	−4.964***	−4.954***	−5.001***	−4.955***
	(0.145)	(0.147)	(0.157)	(0.143)	(0.133)
CPTCLUST3	−5.872***	−5.869***	−5.869***	−5.908***	−5.888***
	(0.147)	(0.149)	(0.159)	(0.145)	(0.135)
CPTCLUST4	−6.622***	−6.619***	−6.616***	−6.662***	−6.652***
	(0.148)	(0.150)	(0.160)	(0.146)	(0.136)
CPTCLUST5	−3.726***	−3.725***	−3.720***	−3.768***	−3.724***
	(0.147)	(0.149)	(0.159)	(0.145)	(0.134)
CPTCLUST6	−2.350***	−2.353***	−2.342***	−2.373***	−2.336***
	(0.147)	(0.150)	(0.159)	(0.146)	(0.137)
PROVCLUST2	0.363***	0.367***	0.354***	0.374***	0.333***
	(0.066)	(0.067)	(0.067)	(0.068)	(0.065)
PROVCLUST3	0.967***	0.970***	0.962***	0.953***	0.969***
	(0.094)	(0.096)	(0.094)	(0.092)	(0.095)
PROVCLUST4	0.484***	0.483***	0.478***	0.475***	0.464***
	(0.024)	(0.025)	(0.024)	(0.025)	(0.025)
PROVCLUST5	0.331***	0.330***	0.332***	0.331***	0.329***
	(0.015)	(0.016)	(0.016)	(0.017)	(0.018)
PROVCLUST6	0.417***	0.416***	0.400***	0.399***	0.398***
	(0.019)	(0.020)	(0.019)	(0.020)	(0.020)
PATCLASS-INPATIENT	0.077**	0.074**	0.078**	0.078***	
	(0.032)	(0.032)	(0.033)	(0.029)	
PATCLASS-SAME DAY ADMIT	0.159***	0.159***	0.151***	0.141***	
	(0.031)	(0.031)	(0.032)	(0.029)	
Constant	7.245***	7.241***	7.218***	7.359***	7.421***
	(0.156)	(0.158)	(0.168)	(0.153)	(0.141)
Observations	19,335	19,335	19,335	19,335	19,335
AIC	58529	58528	58549	58563	58589
Out-of-sample MSPE	1.312	1.312	1.313	1.313	1.314

*p<0.1; **p<0.05; ***p<0.01

Table EC.4 Coefficients of quantile regressions at various quantile levels

	Dependent variable:			
	ACTUALHRS			
	$q = 0.1$	$q = 0.3$	$q = 0.6$	$q = 0.9$
BOOKEDHRS	0.096***	0.151***	0.243***	0.469***
	(0.017)	(0.009)	(0.010)	(0.020)
ASA	−0.023***	0.011	0.066***	0.122***
	(0.009)	(0.009)	(0.010)	(0.023)
AGE	−0.002***	−0.002***	−0.001***	−0.0004
	(0.0003)	(0.0002)	(0.0003)	(0.001)
CPTCLUST2	−4.337***	−5.027***	−4.895***	−4.563***
	(0.684)	(0.099)	(0.121)	(0.214)
CPTCLUST3	−4.843***	−5.919***	−5.803***	−5.143***
	(0.683)	(0.101)	(0.122)	(0.220)
CPTCLUST4	−5.254***	−6.573***	−6.591***	−6.219***
	(0.684)	(0.103)	(0.124)	(0.225)
CPTCLUST5	−3.420***	−3.783***	−3.688***	−3.433***
	(0.693)	(0.107)	(0.125)	(0.216)
CPTCLUST6	−2.821***	−2.495***	−2.328***	−2.420***
	(0.701)	(0.115)	(0.122)	(0.216)
PROVCLUST2	0.318***	0.297***	0.473***	0.924***
	(0.064)	(0.079)	(0.077)	(0.123)
PROVCLUST3	0.391***	0.742***	1.087***	1.812***
	(0.134)	(0.096)	(0.083)	(0.193)
PROVCLUST4	0.228***	0.414***	0.520***	0.764***
	(0.039)	(0.025)	(0.029)	(0.066)
PROVCLUST5	0.207***	0.295***	0.325***	0.380***
	(0.026)	(0.018)	(0.017)	(0.047)
PROVCLUST6	0.181***	0.352***	0.417***	0.538***
	(0.032)	(0.021)	(0.020)	(0.050)
PATCLASS-INPATIENT	0.185***	0.075	0.080	0.176
	(0.071)	(0.078)	(0.066)	(0.142)
PATCLASS-SAME DAY ADMIT	0.589***	0.274***	0.093	−0.141
	(0.075)	(0.078)	(0.065)	(0.141)
Constant	5.543***	7.130***	7.237***	7.005***
	(0.691)	(0.136)	(0.146)	(0.282)
Observations	19,335	19,335	19,335	19,335
Note:		*p<0.1; **p<0.05; ***p<0.01		

Figure EC.1 Histogram of average daily anesthesia hours

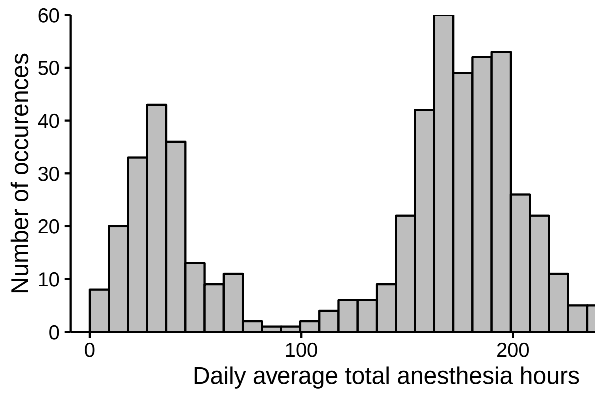

Figure EC.2 Histogram of average daily anesthesia hours controlling for number and mix of surgeries

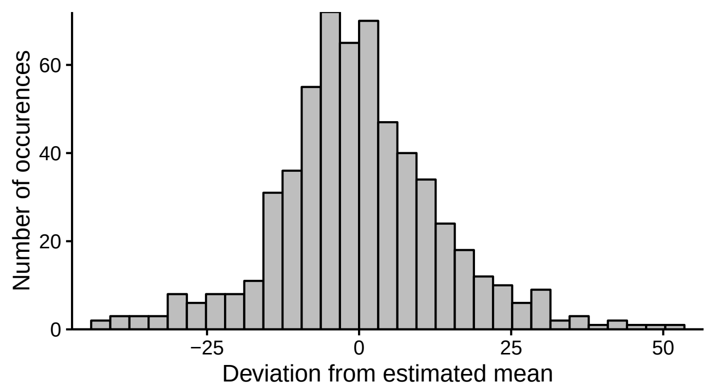

Figure EC.3 Number of rooms with assigned anesthesiologists by hour of day

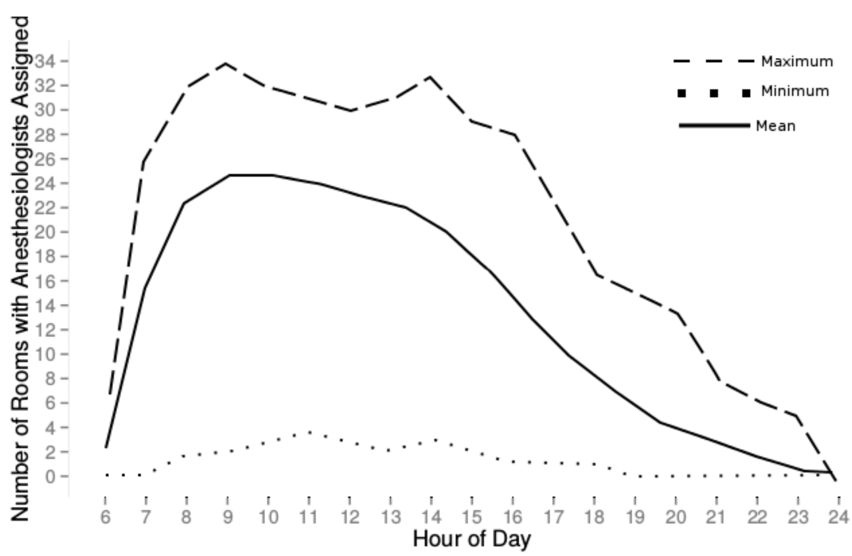

Figure EC.4 Number of CPT code clusters and % explained variance

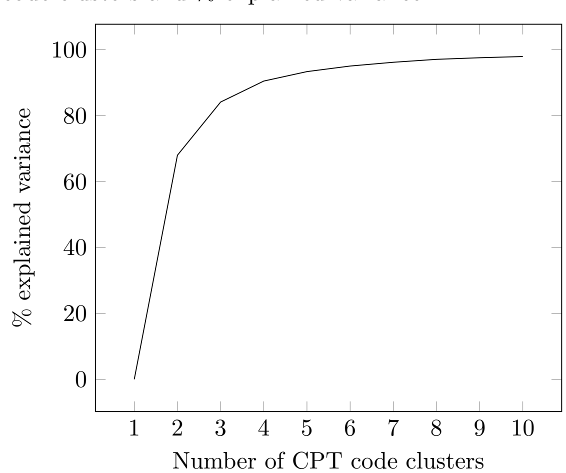

Figure EC.5 Number of surgeon clusters and % explained variance

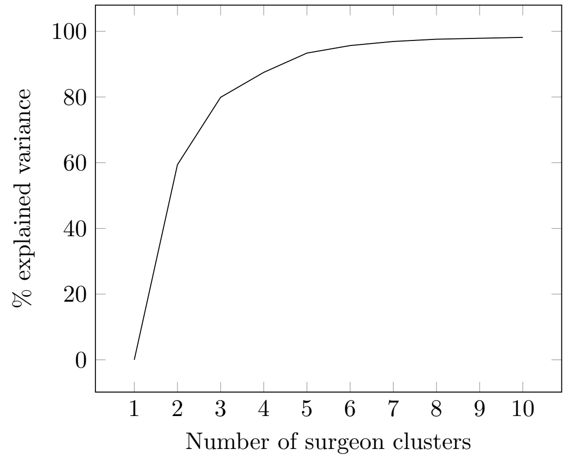

Figure EC.6 Software schematic of decision support system