

The effect of decision fatigue on surgeons' clinical decision making

Emil Persson¹  | Kinga Barrafre¹  | Andreas Meunier² | Gustav Tinghög^{1,3} 

¹Department of Management and Engineering, Division of Economics, Linköping University, Linköping, Sweden

²Department of Clinical and Experimental Medicine, Orthopedics, Linköping University, Linköping, Sweden

³National Center for Health Care Priority Setting, Division of Health Care Analysis, Linköping University 581 83, Linköping, Sweden

Correspondence

Gustav Tinghög, Department of Management and Engineering, Division of Economics, Linköping University, Linköping 581 83, Sweden.
Email: gustav.tinghog@liu.se

Funding information

Östergötland County Council (ALF project); Ragnar Söderberg foundation; Marianne and Marcus Wallenberg foundation, Grant/Award Number: MMW2014.0187

Abstract

The depleting effect of repeated decision making is often referred to as decision fatigue. Understanding how decision fatigue affects medical decision making is important for achieving both efficiency and fairness in health care. In this study, we investigate the potential role of decision fatigue in orthopedic surgeons' decisions to operate, exploiting a natural experiment whereby patient allocation to time slots is plausibly randomized at the level of the patient. Our results show that patients who met a surgeon toward the end of his or her work shift were 33 percentage points less likely to be scheduled for an operation compared with those who were seen first. In a logistic regression with doctor-fixed effects and standard errors clustered at the level of the doctor, the odds of operation were estimated to decrease by 10.5% (odds ratio = 0.895, $p < .001$; 95% CI [0.842, 0.951]) for each additional patient appointment in the doctors' work shift. This pattern in surgeons' decision making is consistent with decision fatigue. Because long shifts are common in medicine, the effect of decision fatigue could be substantial and may have important implications for patient outcomes.

KEYWORDS

decision fatigue, ego depletion, medical decision making, time of day

1 | INTRODUCTION

Every day physicians make repeated decisions about patient care. Such decisions typically involve careful deliberation of the costs, risks, and benefits of a given course of action. An important insight from behavioral theory is that careful consideration of pros and cons in decision making is mentally taxing, leading people to shy away from engaging in cognitively demanding reasoning when tired (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Gabaix, Laibson, Moloche, & Weinberg, 2006; Payne, Bettman, & Johnson, 1993; Rubinstein, 1998; Simon, 1955; Vohs et al., 2008). The depleting effect of repeated decision making is often referred to as decision fatigue. Following repeated decision making, physicians may turn to decision-making heuristics, that is, mental shortcuts that allow us to make decisions on the basis of simple rules of thumb without engaging in cognitively demanding reasoning.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors Health Economics Published by John Wiley & Sons Ltd

One of the most well-documented heuristics in decision making is relying on a default option or preference for the status quo (Johnson & Goldstein, 2003; Samuelson & Zeckhauser, 1988; C. R. Sunstein, 2014). In a seminal study by Danziger, Levav, and Avnaim-Pesso (2011), more than 1,000 judges' parole decisions were reviewed. Danziger et al. found that the prisoners who came up for decision early in the morning received parole about 65% of the time, but this rate then steadily declined to only 15% just before a break. After morning and lunch breaks, the chance of parole was again high. The authors concluded that making decisions is depleting, and once depleted, judges start to look for the easiest and safest option, which is to stick with the status quo and leave the prisoners in jail. This interpretation has however been drawn into question in later work (Daljord, Urminsky, & Ureta, 2017; Glöckner, 2016; Weinshall-Margel & Shapard, 2011). Still, the influence of decision fatigue and default or status-quo effects have been recognized in other contexts, for example, automobile configurations and voter behavior (Augenblick & Nicholson, 2016; Levav, Heitmann, Herrmann, & Iyengar, 2010).

To date, little research has been conducted to explore if and how extraneous factors (such as patient ordering) that could be linked to decision fatigue affects medical decision making. The only study we are aware of found that primary care clinicians' likelihood of prescribing antibiotics increased throughout morning and afternoon sessions (Linder et al., 2014). Moreover, recent studies found that health care workers' compliance with professional standards for hand hygiene deteriorates during the course of their work shifts (Dai, Milkman, Hofmann, & Staats, 2014). In this study, we investigate whether medically extraneous factors that can be linked to decision fatigue affect clinical practice. To this end, we test whether orthopedic surgeons' decisions to operate depend on the sequence of patient appointments throughout the day.

Under ideal circumstances, patient ordering is an extraneous factor that should not influence the quality of medical decisions. However, a central component of decision fatigue is that making many choices depletes our mental resources, which in turn impairs subsequent reasoning and makes people more prone to choosing the status quo (Baumeister et al., 1998; Polman & Vohs, 2016; Vohs et al., 2008). Theoretically, decision fatigue should therefore make orthopedic surgeons more likely to postpone their decisions or making conservative recommendations without thoroughly considering the pros and cons of an operation. We therefore hypothesized that there would be fewer patients scheduled for operation by doctors who had already seen many patients during their work shift.

Hypothesis 1. *The tendency to schedule a patient for operation is decreasing in the number of patients already seen by the doctor during the work shift.*

The strength of this effect largely depends on the attractiveness of the status quo. There are three main reasons for why the status quo (no operation) is an attractive option for orthopedic surgeons. First, the status quo option is reasonably safe because the risk of irreversible exacerbation due to “watchful waiting” is small. In contrast, surgery always poses some risks for the patient. For example, a systematic review that incorporated 16,424 surgical patients (from all surgical fields) found that an adverse event occurred in 14.4% of cases (Anderson, Davis, Hanna, & Vincent, 2013). Choosing the safer option is thus a plausible heuristic that people may turn to when they are depleted.¹ Second, additional effort is required for preoperative procedures and medical reporting in case patients are scheduled for operation, which makes the status quo option (no operation) costly to overturn (Kamenica, 2012; O'Donoghue & Rabin, 1999; Cass R. Sunstein & Thaler, 2003). The medical reporting to a separate journal in case of operation is always done by the surgeon, and the preoperative procedures (e.g., EKG and blood samples) are often planned by the surgeon but carried out by a supporting nurse, either right away or on a different day. Third, because this study is conducted within the context of a publicly funded health care system, there are no financial incentives associated with prescribing surgery.

If it is indeed the case that repeated decision making serves to deplete surgeons' mental resources, taking a break should logically allow them to replenish their energy. A pattern consistent with this effect was observed for the judges in the Danziger et al. (2011) study, where a food break was followed by a sharp increase in the proportion of prisoners who received parole, compared with the prisoners who were up for decision just before the break. Previous research has

¹Experimental studies using time-pressure or stress-induction protocols have found that people become less risk taking when prompted to make automated and effortless decisions (Cahliková & Cingl, 2017; Kirchler et al., 2017; Porcelli & Delgado, 2009). A more general effect in these studies is that people's choices become more in line with the reflection effect of prospect theory when they fall back on automated decision making, that is, they become more risk averse for gains but risk seeking for losses. However, no experimental study has tested the effect of decision fatigue on risk preference directly, and it is difficult to draw general conclusions based on studies that focused on related concepts, such as time pressure, ego depletion, or cognitive load. Ego depletion is an active research field at the moment and the most recent studies indicate that the correlation with risk preference is weak (Gerhardt, Schildberg-Hörisch, & Willrodt, 2017) or non-existing (Koppel, Andersson, Västfjäll, & Tinghög, 2019). We also know of one study that manipulated sleepiness in subjects, and they found that it increased risk taking (Castillo, Dickinson, & Petrie, 2017).

argued that a food break may help people to replenish their mental resources because of the resting period per se (Tyler & Burns, 2008) and also because eating helps to restore the amount of body glucose, which is an important energy source for the brain (Gailliot & Baumeister, 2007). We therefore hypothesized that there would be an upward jump in the proportion of patients scheduled for operation following the lunch break.

Hypothesis 2. *There is an upward discontinuity around lunch in the trend to schedule fewer patients to operation as the day passes.*

2 | MATERIALS AND METHODS

We collected data from the hospital registry of an orthopedic clinic in Sweden on patient appointments and decision to operate. Local administrators retrieved the data from clinical administrative registers. The researchers handled no personal data.

2.1 | Dataset description

Our data covers 848 patient appointments during the course of 133 different doctor work shifts. Eight orthopedic surgeons work at the hospital. All doctors are senior consultants. Our data covers all the scheduled appointments from the period between October and December 2015. Emergency cases were handled separately and are not included in the dataset. The period for the data collection was chosen to avoid summer and winter vacations, which could create a bias in scheduling patients. The dataset includes information about the appointments, such as date and time of the scheduled visit, location of the medical problem (knee, foot, hip, other), patient's age and gender, and information about doctor's work shift, whether it took place before noon (am) or in the afternoon (pm), as well as whether doctor had one (single shift) or two (double shift) shifts during the work day.

2.2 | Scheduling procedures

The scheduling of patients in this hospital is done consecutively. Referrals from primary care doctors arrive via a computer system consecutively as they are written. An orthopedic surgeon reviews the referrals and assigns the case to the hip, knee, or foot team. Thereafter, an administrator (secretary) enters the case on a consecutive list for a work shift with a fixed number of case slots. This means that cases are added in the same order as the referrals from the primary care unit arrive. Neither type nor extent of problem or the characteristics of the patient are taken into account. No patient can book a primary consultation without a referral from the primary care unit. The entire time from arrival of the referral to the visit is approximately 6 weeks, so there is no need to prioritize more urgent cases. A revisit (new appointment for the same problem) can be scheduled on request by the doctor or patient and is then added to the first available time slot. However, some revisits can be arranged on shorter notice, and these are added at the end of the list, thus making the shift longer than expected.

By virtue of the scheduling procedure at the hospital, each patient is essentially allocated to a specific time slot of a given doctor at the clinic by chance (assuming there is no systematic pattern in the timing of patient referrals sent from the primary care units). Therefore, neither type nor extent of problem or the patient's individual characteristics should influence the particular sequence of patient appointments for the doctor during the work shift. In this sense, we are exploiting a natural experiment, because a useful control group (first patient appointment during the work shift) arises naturally by the scheduling procedure at the hospital (Harrison & List, 2004). The only exception is that the proportion of revisiting patients could increase toward the end of the work shift, because some of these visits can be arranged on short notice and are then added to the existing list of patients for that particular day.

During the appointment, the surgeon decides whether to schedule an operation (for another date). In case the decision is to operate, the surgeon has to report the case in a separate journal, and the patient sometimes has to undergo a short preoperative examination right away (e.g., EKG and blood samples). These support functions are equally available all day during normal hours when surgeons are scheduled for patient appointments. Figure 1 illustrates the pathway for patients at the clinic. As can be seen in the figure, revisits include both post-surgery check-ups as well as follow-ups on patients for whom conservative treatment (no surgery) was chosen at the previous visit to the clinic.

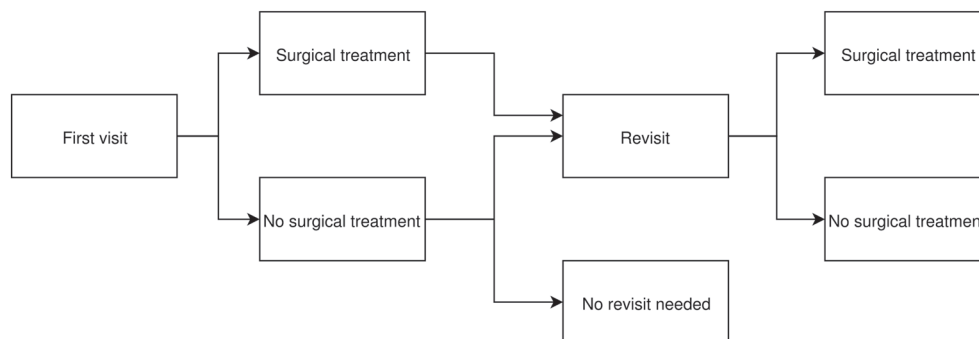


FIGURE 1 Flow chart of orthopedic patients. Note: Revisits are typically scheduled around 3 months from first appointment (and often 4–5 months for patients who have surgery), which means that the majority of revisits in our data comes from patients who had their first appointment before the beginning of the study period (3 months)

2.3 | Statistical methods

We use graphical analysis to present the trends in the proportion of scheduled surgeries depending on patient's ordinal position within a work shift. In our analyses, we first pooled morning and afternoon shifts together, meaning that for each doctor and workday, the first patient seen during a morning shift as well as the first patient seen during the afternoon shift is classified as *ordinal position* = 1. This enabled us to test for a downward trend in scheduled surgeries within doctors' work shifts (Section 3.1). To account for the panel structure of the dataset and for correlation between decisions of each doctor, we estimated fixed-effects logit model for panel data with decision to operate as dependent variable and the patient's ordinal position (within a work shift) as the main independent variable. The fixed-effects model in our study estimates the effect of a change in patient's ordinal position within work shift on the odds of scheduling an operation within each doctor and then averages it over all doctors. Furthermore, fixed-effects model partials out the effect of differences between doctors (e.g., doctor's skills, tendency to operate, age, and gender), so that the estimates do not suffer from time-invariant omitted-variable bias. In addition, to control for heteroscedasticity and serial correlation in the disturbances, we estimated robust standard errors clustered at the doctor's level. The use of robust standard errors in fixed-effects models has been suggested in previous literature (Wooldridge, 2010). We used a dummy for revisit as one of the regressors in order to control the fact that the number of revisiting patients increased toward the end of the shift. We also ran the regression separately in the two subsamples that included only first visits and only revisits, respectively.

In Section 3.2, we illustrate the discontinuity in scheduled operations around lunch by plotting the proportion of scheduled operations against the clock hour of appointment. In our formal analysis, we used the same estimation strategy as above (fixed-effects logit model for panel data) except that we used patients' ordinal position within the whole day rather than within a shift (i.e., we did not reset the count for patients' ordinal position after lunch) and we restricted the sample to double shifts (both morning and afternoon for a given doctor on a given date). The discontinuity around lunch is captured by a dummy for patients seen in the afternoon and we allowed for potentially different trends (slopes) before vs after lunch by including an additional variable that only counts patients' ordinal position in the afternoon.

3 | RESULTS

Table 1 provides summary statistics for our sample. During the time period covered, there were 848 patient visits. Surgery was prescribed in 32% of cases. The median number of patients per shift was six (mean = 6.37) and 89.5% of the shifts contained between five and nine patients (see online Supplementary information Table S1 for more details). Revisiting patients constituted 61% of the recorded cases. The average age of the patients was 60 years, and 61% were females. There were three types of workdays for surgeons: single shift in the morning (before lunch), single shift in the afternoon (after lunch), and double shift (first shift before lunch and second shift after lunch). We can see in Table 1 that the shifts were evenly divided between morning (53%) and afternoon (47%) and that double shifts were somewhat more common compared with single shifts (64% vs. 36%). Five observations have missing data on decision to operate and we dropped them from the analysis. We found no systematic relationship between the sequential time point of appointment and case characteristics or individual patient characteristics (see Figure S3 and the regression results in Tables S4–S5 for details). This supports our assumption that patient assignment to time slots is unbiased with

TABLE 1 Sample characteristics of patients and surgeons. Values are numbers (percentages) unless otherwise stated

Variable	Full sample
<i>Surgeon level</i>	
Participating surgeons	8
Total number of patients seen by the surgeons	848
Recommended treatment:	
Surgery	270 (32)
No surgery	573 (68)
<i>Surgeons' work shifts</i>	
Total number of work shifts	133
Work shift time:	
Before lunch	71 (53)
After lunch	62 (47)
Workday type:	
Single shift	29 (36)
Double shift (am + pm)	52 (64)
Mean (SD) number of patients within a work shift	6.37 (1.80)
<i>Patient level</i>	
Type of patient visit:	
First visit	330 (39)
Revisit	518 (61)
Type of condition:	
Hip	160 (19)
Knee	324 (38)
Foot	330 (39)
Other	33 (4)
Mean (SD) age of patients	60.05 (15.52)
Patient's gender:	
Female	516 (61)
Male	332 (39)

respect to patient age, gender, location of the medical problem, time of the shift, and type of a work day. The only exception was that revisits, as expected, were more common toward the end of the shift. To account for this, we undertake additional robustness checks, testing our main hypothesis separately for first-visiting and revisiting patients.

3.1 | Downward trend in scheduling surgery within doctors' work shifts

We found that patients who met a surgeon toward the end of his or her work shift were less likely to be scheduled for an operation compared with those who were seen first. This negative pattern is visible in Figure 2, which shows how the proportion of scheduled surgeries depended on patients' ordinal position within a surgeon's work shift. In the figure, we see a negative relationship between the proportion of operations and the ordinal position of patients. For example, whereas 40.2% of patients allocated the first timeslot (either in the morning or after lunch, see Section 2 for details) were scheduled for surgery, the proportion decreased to 21.7% for patients who were allocated the sixth timeslot (Chi-squared test, $p = .003$, $n = 229$), and to only 6.7% at the ninth timeslot (Chi-squared test, $p = .011$, $n = 147$).² The negative relationship between the patient's ordinal position and the doctor's decision to operate can be observed for both first-visiting and revisiting patients (Figure S4).

Next, we estimated the effect of an additional patient appointment on the odds of the decision to schedule a patient for operation. Results from the logistic regression analysis are presented in Table 2. A one-unit increase in the patient's ordinal position in a work shift is associated with a decrease in odds of being scheduled for an operation by 10.5% (odds ratio [OR] = 0.895, $p < .001$; 95% CI [0.842, 0.951]). This finding provides support for Hypothesis 1. We can also see that revisiting patients have lower odds of being scheduled for a surgery than patients who visit the doctor for the first time with a given medical issue. As an additional robustness check, we also performed the regression separately for each of the two subsamples, that is, first visits and revisits, respectively. In that case, the odds of being scheduled for an operation decreased by 11.9% for a one-unit increase in the patient's ordinal position (OR = 0.881, $p = .019$; 95% CI [0.792, 0.979]) for patients visiting for the first time and by 10.3% (OR = 0.897, $p = .061$; 95% CI [0.801, 1.005]) for revisiting

FIGURE 2 The patients' ordinal position and doctors' decision to operate ($n = 843$). Note: Morning and afternoon shifts are pooled together, that is, for each doctor and workday, *patient position* = 1 represents a patient who is seen first in a morning shift as well as a patient who is seen first in an afternoon shift (see Section 2 for more information and Supporting information text for additional analyses). The figure does not show average values for the 10th and 11th patients within working shift due to very low sample size ($n = 3$ and $n = 1$, respectively), which are not considered representative. Dashed lines represent 95% confidence intervals that account for binary nature of the variable and are constructed based on clustered standard errors at the doctor's level

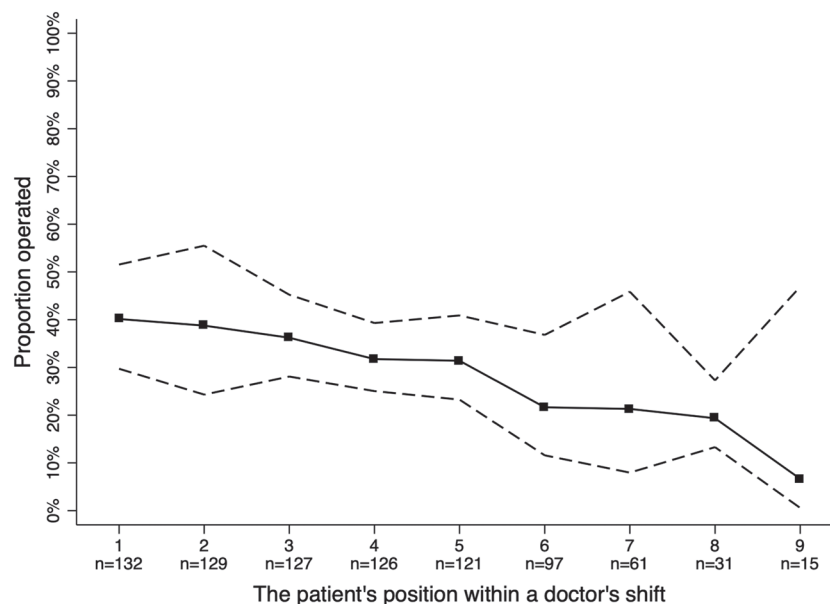


TABLE 2 The effect of patient's ordinal position on the odds of scheduling an operation (logistic regression with fixed-effects)

Dependent variable: Operate (1 = yes 0 = no)	All patients	First visit	Revisit
Patient's ordinal position	0.895*** (0.028)	0.881** (0.048)	0.897* (0.052)
First visit	REF		
Revisit	0.344*** (0.064)		
Morning shift	REF	REF	REF
Afternoon shift	0.997 (0.232)	0.737 (0.174)	1.29 (0.476)
Single shift	REF	REF	REF
Double shift (am + pm)	0.854 (0.119)	0.851 (0.231)	0.787 (0.127)
Location: hip	REF	REF	REF
Location: knee	0.923 (0.571)	0.749 (0.433)	1.282 (1.057)
Location: foot	0.626 (0.218)	1.390 (0.462)	0.180*** (0.115)
Location: other	0.441 (0.382)	0.442 (0.479)	0.452 (0.420)
Patient's age	1.002 (0.006)	1.013 (0.010)	0.995 (0.007)
Patient's gender: male	REF	REF	REF
Patient's gender: female	1.193 (0.145)	1.212 (0.209)	1.113 (0.255)
N	842	330	512

Note. Odds ratios with standard errors (in parentheses) clustered at the doctor's level.

* $p < .1$; ** $p < .05$; *** $p < .01$.

patients. These analyses indicate that the relationship between patient order and the odds of being scheduled for operation is similar for first visits and revisits, and this is further supported by a regression that includes an interaction between patient orders and revisits (Table S6). The interaction term is close to unity and not significantly different from one (OR = 0.907, $p = .789$; 95% CI [0.817, 1.166]). Our main results are robust when the data is analysed separately for morning shift, afternoon shift, single shifts, and double shifts (see Supporting information text and

²The sixth and ninth patients correspond to the 50th and 95th percentiles of the number of patients within a work shift, respectively.

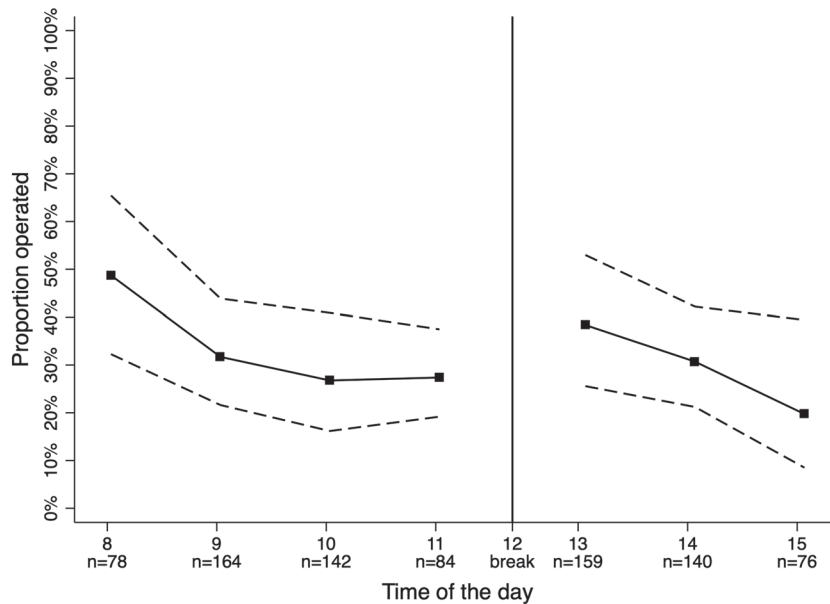


FIGURE 3 Time of the day of the appointment and doctors' decision to operate ($n = 843$). Note: We categorize patient appointments by the time of the day. All appointments that started between 08:00 a.m. and 08:59 a.m. are categorized as 8:00 a.m., all appointments between 09:00 a.m. and 09:59 a.m. are categorized as 9:00 a.m., and so on. Dashed lines represent 95% confidence intervals that account for binary nature of the variable and are constructed on the basis of clustered standard errors at the doctor's level

Tables S2–S3 and Figure S1–S2). They are also robust to removing the last patient for each shift and surgeon ($OR = 0.887$, $p = .001$; 95% CI [0.828, 0.95]), and to additionally removing the second-to-last ($OR = 0.909$, $p = .007$; 95% CI [0.848, 0.975]) and third-to-last ($OR = 0.824$, $p = .003$; 95% CI [0.726, 0.936]) patient appointments (see Table S7 for more information). This alleviates the concern that our results are driven by time constraints at the end of the shift rather than decision fatigue.

3.2 | Discontinuity in the trend of scheduled surgeries around lunch

We investigated the temporal pattern of surgeons' decision making and found a declining trend in scheduled surgeries both before and after the lunch break (Figure 3). As can be seen in the figure, the proportion of patients who were scheduled for surgery was high in the early morning and decreased until lunch break, after which, the proportion was again high and decreased until the end of the afternoon shift. The difference between the two data points just before and just after lunch is 11.0 percentages and it is weakly significant in a Chi-squared test ($p = .087$, $n = 243$). The lunch effect is stronger and significant if we restrict attention to the doctors who worked full-day shifts on a given day. In that case, the proportion of patients who were scheduled for operation increased with 14.8 percentages, from 24.2% during the final hour of the morning shift to 39.0% during the first hour of the afternoon shift (Chi-squared test, $p = .042$, $n = 198$).

We follow-up with a fixed-effects logistic regression to formally test for an upward discontinuity around lunch in the trend to schedule fewer patients to operation as the day passes. As expected, we found a declining trend in scheduled surgeries throughout the day such that each additional patient appointment was associated with a decrease in the odds of being scheduled for an operation by 11.1% ($OR = 0.889$, $p = .048$; 95% CI [0.791, 0.999]). In line with Hypothesis 2, there is a clear discontinuity around lunch; the odds of operation are on average 1.96 times higher for patients who are seen in the afternoon compared with patients seen in the morning ($OR = 1.955$, $p = .023$; 95% CI [1.110, 3.486]). This means, for example, that the predicted probability of operation for a patient who has the sixth appointment during the day is 25.5% if the appointment occurs just before lunch, and increases to 38.8% if the appointment instead occurs just after lunch (still as the sixth patient for the day). The full specification and regression results can be found in Table S8.

4 | DISCUSSION

We found a strong link between the surgeons' decisions to operate and the sequence of patient appointments they face throughout the day: much fewer patients are scheduled for operation by surgeons who are nearing the end of their work shift. We attribute this effect to decision fatigue; that surgeons become more inclined to rely on heuristics and go with the status quo option when tired.

Our interpretation of the data is supported by the fact that the sequence of patient appointments was associated with the decision to operate, in the expected direction, but not with case characteristics or individual patient characteristics. The only exception was that revisits were more common toward the end of the shift, because unscheduled revisits were sometimes added at the end. This was compensated for in the analysis by investigating first visits and revisits separately. At a general level, there are other behavioral phenomena beside decision fatigue that could give rise to sequencing effects in randomly ordered data. In the context of court decisions, a previous literature has argued for the possibility of both positive and negative autocorrelation of decisions, for example, due to judges (or juries) striving to be internally consistent or using self-imposed quotas, such as a maximum number of affirmative decisions on a single day (Bindler & Hjalmarsson, 2018; Chen, Shue, & Moskowitz, 2016). Sequencing effects could also arise due to a belief bias called the gambler's fallacy, whereby individuals underestimate the likelihood of long streaks occurring by chance and thus take the frequency of previous cases with similar characteristics into account when making their current decision (Chen et al., 2016). For example, a doctor who saw several consecutive cases with a clear indication of operation in all of them would, all else equal, be slightly biased toward non-operative treatment in the upcoming appointment. However, none of these explanations should, arguably, yield the type of downward trend that we see in our data, and the upward jump around lunch is difficult to explain without making further assumptions. In contrast, both the downward trend and the lunch effect are consistent with decision fatigue.

Doctors' decisions might also be influenced by patients' demands. Saying no to patients can be mentally burdensome and doctors may experience discomfort if they feel they are denying care to the patient (Persson et al., 2018; Tinghög, 2011; Ubel, 2001; Ubel & Goold, 1997). For Linder et al. (2014), who showed that the likelihood of primary care clinicians' prescribing antibiotics increased throughout the day, patient demand is likely an important factor. Orthopedic patients with complex problems, on the contrary, are less likely to strongly push for a surgical intervention, which could be both inconvenient and risky. This arguably tips the scale for our study compared with Linder et al., meaning that for decisions that are close calls between surgery and no surgery, tired orthopedists are pulled toward the immediate situation, which in this case is not to operate.

Some limitations related to our study should be noted. First, without a physiological measure of fatigue at the point of decision making, we cannot completely rule out competing explanations for the pattern that we see in our data. One alternative explanation could be that surgeons become reluctant to consider operation as the end of the shift approaches simply because they do not want to work overtime. However, the fact that our results were robust to excluding the final patient(s) for each doctor and work shift indicates that this is likely not the main explanation of our result. Second, the unique identifiers in our data are at the level of each patient appointment for a specific surgeon at a specific date-time. This means that we cannot track individual patients who visit more than once during the study period. However, because our study covered a relatively short time period, a second appointment for the same patient would under most circumstances have occurred after the end of the study period. Finally, even though patients were added to the consultation list as the referrals came in, it cannot be ascertained that the sequence of patients was completely random. We can only test this assumption based on a limited number of observable patient and case characteristics. In addition, our data does not allow us to assess whether an appointment has been rescheduled by patient request. In principle, if some types of patients are more likely to reschedule to certain time points during the day, and if these patients differ systematically from the rest of the sample, then this could bias the results. Still, the general routines for scheduling patients to time slots should prevent any systematic ordering for the majority of cases (except the revisits), and the administrators reported that they took no such measures.

Our study is restricted to the decision to operate, and to orthopedic surgeons in one single unit. A general interpretation of the results should therefore be made with caution, keeping in mind that both theory and previous results from behavioral experiments stress the importance of the decision environment for the onset of decision fatigue. Danziger et al. (2011) is the only other study we are aware of that investigates decision fatigue in expert decision making using quasi-randomized data, in their case concerning the effect of the ordinal position of Israeli prisoners coming up for decisions by a parole board. Although their sample is limited, like ours, the combined results of the two studies speak toward the more general conclusion that decision fatigue is an important factor in real-life contexts where decisions are taken on behalf of other individuals.

Taken together, our results point to a few key aspects in the decision environment that can be targeted by policy interventions. First, time spent on the task seems to be strongly related to decision fatigue. Short breaks could potentially alleviate this effect. Second, apart from the effort required in weighing the pros and cons when making the actual decision, the doctor faces an additional effort requirement if deciding to operate because extra paperwork and patient communication must be undertaken right away. This makes the status quo option (no operation) costly to overturn,

and policy makers should therefore try to minimize this additional effort required by the doctor in case he or she were to decide on operation. On a more general level, decision fatigue could be one aspect to consider when designing physician payment systems in the future, because it is an example of how pushing doctors to work too hard may create inefficiencies and affect patient outcomes. Our results show that the tendency to operate is higher when a doctor is relatively well-rested and decreases as doctors make more and more decisions. This is bad news for patients because it suggests that their treatment is affected by incidental factors and that there is some arbitrariness in decisions taken by the doctors. From a societal point of view, this is an inefficient and arguably unfair use of medical resources. Further research is needed to conclude whether the declining trend in the decision to operate is specific to orthopedics or holds also in areas where lack of surgery might have more detrimental effects.

ACKNOWLEDGEMENTS

We are thankful to Per Aspenberg who played an important role in instigating this project but unfortunately did not live to see it in print. Moreover, we are grateful to Magnus Johannesson, Daniel Västfjäll, Gaute Torsvik, Manja Gärtner, Peter Martinsson, and two anonymous reviewers for valuable comments and suggestions and Margareta Carlsson and Mathilda Cassel for data collection. This study was supported by grants from Marianne and Marcus Wallenberg foundation, Ragnar Söderberg foundation, and Östergötland County Council (ALF project). Funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. All authors declare no competing interests that could have appeared to influence the submitted work.

ORCID

Emil Persson  <https://orcid.org/0000-0003-2994-0541>

Kinga Barrafreem  <https://orcid.org/0000-0003-3171-8640>

Gustav Tinghög  <https://orcid.org/0000-0002-8159-1249>

REFERENCES

- Anderson, O., Davis, R., Hanna, G. B., & Vincent, C. A. (2013). Surgical adverse events: A systematic review. *The American Journal of Surgery*, 206(2), 253–262. <https://doi.org/10.1016/j.amjsurg.2012.11.009>
- Augenblick, N., & Nicholson, S. (2016). Ballot position, choice fatigue, and voter behaviour. *The Review of Economic Studies*, 83(2), 460–480. <https://doi.org/10.1093/restud/rdv047>
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252–1265. <https://doi.org/10.1037/0022-3514.74.5.1252>
- Bindler, A., & Hjalmarsson, R. (2018). Path dependency in jury decision making. *Journal of the European Economic Association*. <https://doi.org/10.1093/jeea/jvy046>
- Cahlíková, J., & Cingl, L. (2017). Risk preferences under acute stress. *Experimental Economics*, 20(1), 209–236. <https://doi.org/10.1007/s10683-016-9482-3>
- Castillo, M., Dickinson, D. L., & Petrie, R. (2017). Sleepiness, choice consistency, and risk preferences. *Theory and Decision*, 82(1), 41–73. <https://doi.org/10.1007/s11238-016-9559-7>
- Chen, D. L., Shue, K., & Moskowitz, T. J. (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics*, 131(3), 1181–1242. <https://doi.org/10.1093/qje/qjw017>
- Dai, H., Milkman, K. L., Hofmann, D. A., & Staats, B. R. (2014). The impact of time at work and time off from work on rule compliance: The case of hand hygiene in health care. *Journal of Applied Psychology*, 100, 846–862. <https://doi.org/10.1037/a0038067>
- Daljord, O., Urminsky, O., & Ureta, J.-M. (2017). The status quo theory of depletion does not explain the Israeli parole decisions. <http://home.uchicago.edu/ourminsky/Status%20Quo%20Theory%20Revisited.pdf>.
- Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108(17), 6889–6892. <https://doi.org/10.1073/pnas.1018033108>
- Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043–1068. <https://doi.org/10.1257/aer.96.4.1043>
- Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, 11(4), 303–327. <https://doi.org/10.1177/1088868307303030>
- Gerhardt, H., Schildberg-Hörisch, H., & Willrodt, J. (2017). Does self-control depletion affect risk attitudes? *European Economic Review*, 100, 463–487. <https://doi.org/10.1016/j.euroecorev.2017.09.004>

- Glöckner, A. (2016). The irrational hungry judge effect revisited: Simulations reveal that the magnitude of the effect is overestimated. *Judgment and Decision Making*, 11(6), 601–610.
- Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42(4), 1009–1055. <https://doi.org/10.1257/0022051043004577>
- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338–1339. <https://doi.org/10.1126/science.1091721>
- Kamenica, E. (2012). Behavioral economics and psychology of incentives. *Annual Review of Economics*, 4(1), 427–452. <https://doi.org/10.1146/annurev-economics-080511-110909>
- Kirchler, M., Andersson, D., Bonn, C., Johannesson, M., Sørensen, E. Ø., Stefan, M., ... Västfjäll, D. (2017). The effect of fast and slow decisions on risk taking. *Journal of Risk and Uncertainty*, 54(1), 37–59. <https://doi.org/10.1007/s11166-017-9252-4>
- Koppel, L., Andersson, D., Västfjäll, D., & Tinghög, G. (2019). No effect of ego depletion on risk taking. *Scientific Reports*, 9, 9724. <https://doi.org/10.1038/s41598-019-46103-0>
- Levav, J., Heitmann, M., Herrmann, A., & Iyengar, S. S. (2010). Order in product customization decisions: Evidence from field experiments. *Journal of Political Economy*, 118(2), 274–299. <https://doi.org/10.1086/652463>
- Linder, J. A., Doctor, J. N., Friedberg, M. W., Reyes Nieva, H., Birks, C., Meeker, D., & Fox, C. R. (2014). Time of day and the decision to prescribe antibiotics. *JAMA Internal Medicine*, 174(12), 2029–2031. <https://doi.org/10.1001/jamainternmed.2014.5225>
- O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1), 103–124. <https://doi.org/10.1257/aer.89.1.103>
- Payne, J., Bettman, J., & Johnson, E. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139173933>
- Persson, E., Andersson, D., Back, L., Davidson, T., Johannesson, E., & Tinghög, G. (2018). Discrepancy between health care rationing at the bedside and policy level. *Medical Decision Making*, 38(7), 881–887. <https://doi.org/10.1177/0272989x18793637>
- Polman, E., & Vohs, K. D. (2016). Decision fatigue, choosing for others, and self-construal. *Social Psychological and Personality Science*, 7(5), 471–478. <https://doi.org/10.1177/1948550616639648>
- Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3), 278–283. <https://doi.org/10.1111/j.1467-9280.2009.02288.x>
- Rubinstein, A. (1998). *Modeling bounded rationality*. Cambridge, Massachusetts: MIT Press.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59. <https://doi.org/10.1007/BF00055564>
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99–118. <https://doi.org/10.2307/1884852>
- Sunstein, C. R. (2014). Choosing not to choose. *Duke Law Journal*, 64(1), 1–52.
- Sunstein, C. R., & Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 70, 1159–1202. <https://doi.org/10.2307/1600573>
- Tinghög, G. (2011). *The art of saying no: The economics and ethics of healthcare rationing (thesis)*. Linköping: Linköping University Press.
- Tyler, J. M., & Burns, K. C. (2008). After depletion: the replenishment of the self's regulatory resources. *Self and Identity*, 7(3), 305–321. <https://doi.org/10.1080/15298860701799997>
- Ubel, P. A. (2001). Physicians, thou shalt ration: the necessary role of bedside rationing in controlling healthcare costs. *HealthcarePapers*, 2(2), 10–21. <https://doi.org/10.12927/hcpap.2002.17453>
- Ubel, P. A., & Goold, S. (1997). Recognizing bedside rationing: clear cases and tough calls. *Annals of Internal Medicine*, 126(1), 74–80. <https://doi.org/10.7326/0003-4819-126-1-199701010-00010>
- Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making Choices Impairs Subsequent Self-Control: A Limited-Resource Account of Decision Making, Self-Regulation, and Active Initiative. *Journal of Personality and Social Psychology*, 94(5), 883–898. <https://doi.org/10.1037/0022-3514.94.5.883>
- Weinshall-Margel, K., & Shapard, J. (2011). Overlooked factors in the analysis of parole decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(42), E833. <https://doi.org/10.1073/pnas.1110910108>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Persson E, Barrafreem K, Meunier A, Tinghög G. The effect of decision fatigue on surgeons' clinical decision making. *Health Economics*. 2019;28:1194–1203. <https://doi.org/10.1002/hec.3933>