Higher Quality and Lower Cost from Improving Hospital Discharge Decision Making*

By James C. Cox[a], Vjollca Sadiraj[a], Kurt E. Schnier[a] and John F. Sweeney[b]

Abstract: The added costs from excess, wasted resources in healthcare are estimated to exceed the size of the entire defense budget. Effectively addressing this problem requires decreasing the cost of healthcare while increasing its quality by improving healthcare decision making. This paper reports research on improving decisions about hospital discharges – decisions that are now made by physicians based on mainly subjective evaluations of patients' discharge status. The research has three components: (a) econometric identification of patient medical status and demographic characteristics that discriminate between the likelihood of successful discharge and unsuccessful discharge (resulting in unplanned readmission within 30 days); (b) development of decision support software that incorporates evidence-based discharge criteria; and (c) empirical evaluation of efficacy of the decision support software. Experimental data reported herein indicate that utilization of the decision support software with evidence-based selection of the default option reduces both average patient length of stay in the hospital and the likelihood of higher risk patients being readmitted.

Keywords: Healthcare, Experiment, Decision Support Software, Risk, Default Option

JEL Classification: C91, D81, I10

* The research was supported by a grant from the National Institutes of Health, National Institute on Aging (grant number 1RC4AG039071-01).

a. Experimental Economics Center (E*x*CEN) and Department of Economics, Andrew Young School of Policy Studies, Georgia State University

b. Department of Surgery, Emory University School of Medicine

Higher Quality and Lower Cost from Improving Hospital Discharge Decision Making

1. Introduction

In 2010 Americans spent 17.6 percent of GDP on healthcare, which was eight percentage points above the OECD average (Organization for Economic Cooperation and Development, 2012). [1] The United States is clearly an outlier when it comes to the growth in the share of our GDP spent on health care services (Chandra and Skinner 2012) and American healthcare is by far the most expensive in the world. Yet the quality patients receive from this investment is somewhere near the middle of the pack when compared to other economically developed nations (Fuchs and Millstein, 2011). Recent estimates suggest that the amount of wasted, excess costs in healthcare were $765 billion for 2009, which was $100 billion more than the entire Department of Defense budget for that year (Institute of Medicine, 2012). In order to brake the runaway train of expense that American healthcare has become we need to decrease the cost of care while increasing its quality. This paper addresses the issue through research on improving hospital discharge decision making.

The objective of decreasing medical costs, or at least reducing their outsized rate of increase, would seem to be well served by reducing hospital length of stay (LOS). If average LOS could be reduced by 10 percent, the savings would exceed $128 billion per year.[2] The research question we take up is how to assist physicians in making discharge decisions that decrease LOS and lower the likelihood of *unplanned* readmissions, an indicator of low quality and a cost inflator. Physicians have rapidly increasing access to large amounts of raw data on each patient they treat through electronic medical record systems. The problem for improving discharge decision making is not shortage of data on the patient but, rather, absence of evidence-based discharge criteria that can be effectively applied at the point of care where the discharge decision is made.

Our central activity is a collaboration between physicians who make discharge decisions and economists – with expertise in research on decisions under risk and information processing – aimed at improving hospital discharge decision making. The objectives are to design and

---

[1] Medicare, Medicaid, and CHIP spending alone made up 21 percent of the 2012 federal budget (Center on Budget and Policy Priorities, 2013). In addition, both Medicaid and CHIP (Children's Health Insurance Program) also require matching expenditures by the states.

[2] In 2010, 39 million patients spent on average 4.7 days in hospitals at a total cost in excess of $1.28 trillion (Agency for Healthcare Research and Quality, 2012).

implement decision support software that can be used to lower costs – by reducing average length of hospital stay – while increasing quality of medical care by decreasing the likelihood of unplanned readmissions.

An outline of current practice in hospital discharge decision making sheds light on the nature of the problem and a possible solution. Prior to deciding whether to discharge the patient, a physician examines the patient and reviews his or her electronic medical records. Criteria applied to making a discharge decision are derived from the physician's recall of his or her medical education and own previous practice and, perhaps, recommendations of one or more colleagues. The evidence base of these typical discharge criteria is extremely limited in comparison to the voluminous information that could be derived from the electronic medical records of the patient *population* of a hospital. A typical hospital will serve many thousands of patients per year. Each surviving patient will be discharged from the hospital and it will subsequently be revealed, in most cases, whether the discharge was successful or unsuccessful (i.e., led to unplanned readmission within 30 days). The central question addressed in our research is how to use this mass of data – from current and former patients' electronic medical records and outcomes from previous discharges of patients – in developing evidence-based discharge criteria that can be effectively applied at the point of care where the discharge decision is made.

We begin with the following question: Do the data profiles for patients who are successfully discharged differ in identifiable ways from the data profiles for patients who are unsuccessfully discharged? If the answer to this question is "yes" then that would open the possibility of building a decision support model that can inform discharge decisions for individual patients with the accumulated experience from discharging thousands of other patients. That is at the heart of our research agenda. We extract a large sample of de-identified data from the "data warehouse" of patient electronic medical records of a large southeastern teaching hospital. The data are used to build an econometric model, which provides the foundation for a decision support model that can be incorporated in software and applied at the point of care. The decision support software presents the physician with a recommended discharge decision and with estimated daily readmission probabilities (and 80% confidence intervals); in addition, it provides information on dynamically-selected key clinical variables for the individual patient in a user friendly format. Testing the software for efficacy in improving discharge decision making includes both a laboratory experiment and a field experiment, in the form of a hospital patient ward intervention. Ethical and practical considerations call for

laboratory evaluation of the efficacy of the software before its use in clinical intervention. The laboratory experiment is reported in this paper. We use a two-by-three experimental design defined over three information treatments and two time constraints on patient care. The experiments are conducted using resident physicians and fourth-year medical students at a university medical school as subjects.

The organization of the paper is as follows. The following section discusses related literature, section 3 describes the econometric model, and sections 4 and 5 report on the experimental design and the efficacy of the discharge support tool in the experiment. A summary of the main findings and conclusions in section 6 completes the paper.

2. Related Economic and Medical Journal Literature

The hospital is an exceptionally expensive care delivery environment and identification of the sources of the rampant increase in healthcare costs since the mid-1950s has been of considerable interest among health economists (Chandra et al. 2010; Chandra and Skinner 2012; Feldstein 1971, 1977; Finkelstein 2007; Pauly 1986; Weisbrod 1991) as well as whether or not these increased expenditures have been welfare enhancing (Feldstein 1973; Hall and Jones 2007; Newhouse 1992; Cutler et al. 1998). The objective of decreasing medical costs, or at least reducing their outsized rate of increase, would seem to be well served by reducing hospital length of stay (LOS). But simply decreasing LOS, without improving discharge decision making, would lead to higher hospital readmission rates, which is a current focus of concern of Medicare.

Indeed, hospital readmissions have recently become one of the critical healthcare quality metrics for American hospitals. In 2010, 19.2 percent of Medicare patients were readmitted within 30 days of discharge, resulting in additional hospital charges totaling $17.5 billion (Office of Information Products and Data Analytics, 2012). Hospitals and physicians are encountering increasing pressure to reduce hospital readmission rates, both from reputation effects of public disclosure of performance and from pay-for-performance reimbursement schemes that refuse payment for related readmissions.[3]

[3] Beginning in October 2012, the Centers for Medicare and Medicaid Services began publishing hospitals' readmission rates and penalizing those with "excess over expected" readmission rates for heart attack, heart failure and pneumonia patients. In 2012, a total of 2,217 hospitals were penalized; 307 of them were assessed the maximum penalty of 1 percent of their total regular Medicare reimbursements (*Kaiser Health News*, Oct 2, 2012). The scheduled penalties escalate in future years and apply to broader classes of treatment diagnosis codes.

The use of advanced information technology has been advocated as a method to increase health care quality and reduce costs (Cebul et al. 2008). Our research is part of a larger program in economics that aims at the creation of information technology for medical decision making and its application in clinical environments intended to improve quality and lower costs of healthcare. A seminal contribution by economists to improving healthcare is the mechanism design incorporated into information technology for kidney exchange by Roth, Sönmez, and Ünver (2004, 2007).[4] Their work provided a foundation for the New England Program for Kidney Exchange, and subsequent kidney exchange programs, which have led to increases in quality and length of life by matching patients with donors for transplant surgery while lowering the informational costs associated with organ matches. Support for improving medical decision making is needed in many additional areas. The present paper reports one such project. Our research lies at the intersection of healthcare cost and quality issues: it targets improving physician discharge decision making through development of clinical decision support systems and implementation of information mechanism design.

The topic of healthcare cost has been of considerable interest since the rampant increase in healthcare costs following World War II as the rate of growth in healthcare expenditures has continually exceeded the annual growth in real GNP per capita (Newhouse 1992). Newhouse identified three demand-side and two supply-side factors that have influenced the cost of healthcare.[5] The supply side factors are physician-induced demand and low increase in productivity. To date little research has been conducted on the physician-induced demand causes of the increase in healthcare costs. However, it has been argued that the advent of the 468 Diagnosis Related Groups (DRG) payment mechanism implemented by Medicare in 1983 would help to mitigate the incentive problems that arise when physicians are compensated based on the level of healthcare provided, such as days in the hospital, and not the health problem being addressed, such as appendicitis (Pauly 1986). The recent actions by Medicare to reduce compensation when a hospital's rates of readmission are above reference levels is another effort to reduce the supplier-induced demand causes of the increase in healthcare costs. This action in essence lowers the effective DRG payment received by a hospital to treat a patient if readmission

[4] They developed a model of paired-kidney exchange for living donors and illustrated the benefits of two-way, three-way and higher level matches.

[5] The demand side factors influencing medical costs are the increased size of our aging population, the spread of insurance and the increase in consumer income.

rates are too high, by reducing the amount of additional compensation a hospital is eligible for if a patient is readmitted.

Our research begins with the premise that an integral component of achieving these cost reductions is the assimilation and dissemination of information to help physicians make better discharge decisions.[6] Our research is one of the first efforts to develop and test the efficacy of using information technology to address a supplier-induced increase in healthcare costs of primary interest to CMS, high rates of hospital readmission. In essence, our research endeavors to make improvements in the development of Comparative Effectiveness Research (CER) as well as testing the efficacy of its utilization.[7]

One of the earliest investigations of the determinants of hospital readmission in the medical literature was conducted by Anderson and Steinberg (1985) who found that a patient's disease history and diagnosis were important determinants of a patient's probability of readmission. More recent research has further illustrated the role that these patient-specific factors have on the probability of readmission (Demir 2012) and that the use of electronic medical record (EMR) data on a patient's vital signs and laboratory test results can be used to explain likelihood of readmissions (Amarasingham et al. 2010). Amarasignham et al.'s (2010) research is relevant to our research as: (1) it validates the use of electronic medical records data in recovering readmission probabilities; and (2) it highlights a fundamental flaw with the current Medicare regulations that generate expectations for readmission rates with models that do not contain clinical information. The importance of using clinical information to inform estimates of readmission rates is also supported by Lee et al. (2012), who study return visits to a pediatric emergency room within 72-hours. Interestingly, both Amarasignham et al. (2010) and Lee et al. (2012) conclude that their research suggests the importance of developing a clinical decision support tool to improve discharge decision making, which is something neither of those teams attempts but is the main objective of our research.

Given the complex relationship between a patient's length of stay (LOS) and their readmission probability a number of researchers have investigated ways that a patient's LOS can be reduced as well as actions that can be taken to reduce the readmission rate. Nabagiez et al.

---

[6] We are not the first to highlight the importance of information in lowering health care costs. Cebul et al. (2008) show that increasing information flows will lower healthcare organizational costs whereas Phelps (1992) argues that the dissemination of the information may be able to reduce regional variation in care.

[7] For a detailed discussion of Comparative Effectiveness Research (CER) in the economics literature see the discussion of Chandra et al. (2011).

(2013) claim that readmission rates could be reduced by 25% by using physician assistance home care (PAHC) developed to facilitate post-discharge care.[8] Providing physicians with information on the average LOS (of similar procedures) as well as outcomes of discharge audits across physicians effectively reduces a patient's LOS (Shea et al. 1995; Caminiti et al. 2013). Caminiti et al. (2013) conclude that LOS could be reduced by 16% without increasing the 30-day readmission rates.

The medical intervention we propose is referred to as a clinical decision support system (CDSS) in the medical literature. In a recent review of 148 studies, Bright et al. (2012), conclude that the current CDSSs (mostly *not* for discharge decisions) are effective at improving healthcare when assisting with physician decision making at the point of care. None of these studies, however, reports a test of efficacy of *discharge* decision support software that (a) applies at the point of care and (b) mandates physician attentiveness by replacing the default option of universal "opt in" to patient discharge with the alternative default option of "opt out" from the CDSS recommendations to discharge or not to discharge the patient on each day of hospital stay. We develop and test decision support software that applies at the point of care and includes a version which requires justification for overriding the software's discharge recommendations.

3. Estimation and Construction of the Decision Support Model

We used electronic medical records data for 3,202 patients who underwent complex gastrointestinal surgery at a large southeastern hospital between the dates of January 2007 and December 2009, stayed in the hospital for at least three days, and had a complete clinical profile of medical history, vital sign reports and laboratory test results during their hospital stay. Additional data were obtained from the electronic medical records on the medications administered, any diagnostic imaging that was conducted, the patient's diet status, and whether or not blood transfusions were provided. The patient's home address was also linked to their census tract to retrieve additional census tract level information (e.g., housing value and education level).[9] Clinical data were recorded in real time during the patient's stay in the hospital. Therefore, we observed the date and time for each of the observations. As a first step in

[8] Whether or not the costs incurred override the additional costs the hospital would have incurred by providing a longer LOS is yet an unanswered question.

[9] Once the information was abstracted from the clinical data warehouse, it was de-identified using the "Safe Harbor" Method as defined by the HIPAA Privacy Rule Section 164.514 (B)(2). Detailed description of the clinical data is contained in Appendix A.

our analysis we constructed a virtual patient chart for each patient that divided each day into eight-hour intervals. All of the variables were then assigned to one of the eight-hour intervals based on the date and time that observation was recorded. In the case that more than one observation was recorded in a particular time interval the average value was assigned to the time interval. The virtual patient charts are used in the econometric estimation and in the experiment software.

The primary outcome we are concerned with in our econometric model is the probability of readmission within 30 days of being discharged from the hospital (the Medicare criterion). We utilized a probit regression framework to estimate the probability of readmission. Since occurrence of the event of readmission is a binary variable, the information contained in a virtual patient chart on the day of discharge was used to estimate the probability of readmission (N=3,202). The explanatory variables included the average values of clinical variables during a patient's stay, the duration of time spent outside and within the normal range of values expected for a particular clinical variable, counts of medications, images and transfusions, as well as a full set of interaction terms between the laboratory test and vital sign variables.

Following construction of the data set we obtained over 900 variables that could affect the probability of readmission. However, some of these variables are highly correlated due to the biophysical nature of the data. To address the issue of multicollinearity, we eliminated some variables when a linear correlation greater than 0.90 with other variable(s) was detected. The elimination of variables was hierarchically determined in consultation with physicians in order to ensure that the variables retained would provide the most reliable profile of patient status.

Following estimation, the patient chart data were used to construct a data set that contained the daily values of each patient's health status and demographic variables over the course of their stay. This resulted in a data set consisting of 48,889 unique patient-day observations that corresponded to the observed value of each patient's data for each day during the hospital stay. This data set was then used to impute the expected probability of readmission for the patient if they were discharged from the hospital on that day using the probit estimates from the estimation algorithm discussed earlier. In addition to predicted probabilities, 80% confidence intervals were constructed using the estimated parameter distributions. An 80% confidence interval was selected because it captures a 10% one-sided error on the decision criterion (explained below) to discharge a patient on a given day.

An illustration of the estimated readmission probabilities is presented in Figure 1 for a sample patient used in the experiment. The kinks in the solid piecewise linear graph show the point

estimates of the readmission probabilities (vertical axis) if the patient were to be discharged on any one of several days during the hospital stay (horizontal axis). The dashed piecewise-linear graphs show the upper and lower bounds on the 80% confidence interval for the readmission probabilities. The horizontal (solid) line shows the target readmission probability for all patients with the same diagnosis code as the patient in this chart. The predicted readmission probabilities and confidence intervals provide primary constructed inputs for the experiment software designed to test the efficacy of the discharge decision support tool.

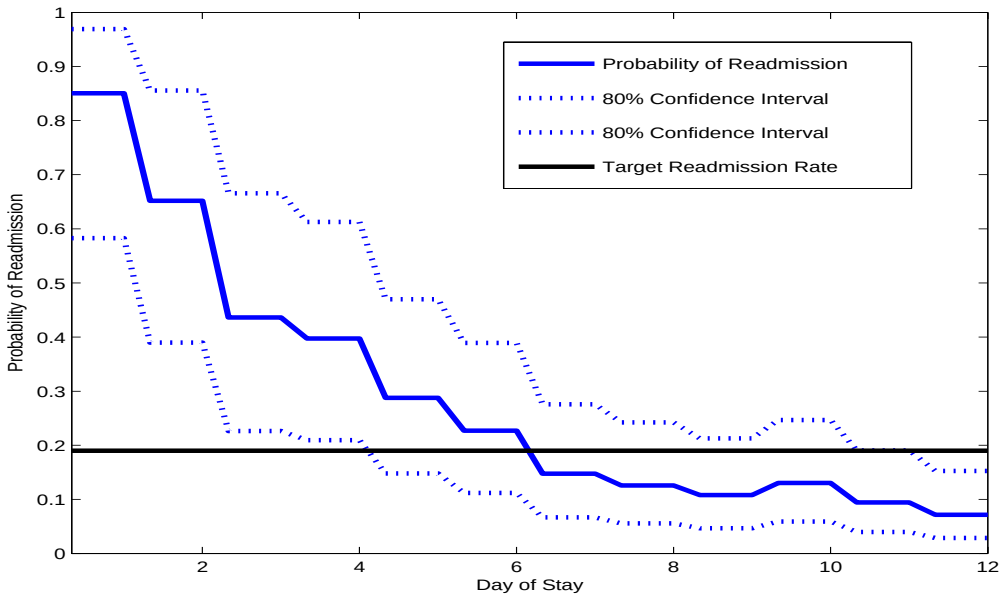

Figure 1: Estimated Daily Readmission Probabilities

The targeted readmission rates represent a uniform 10% reduction from historical readmission rates and are based on the targets stated by the Center for Medicare and Medicaid Services in 2010. The decision support software uses the estimated readmission probabilities and confidence intervals to make recommendations to a physician on daily basis. In the case that the point estimate of the probability of readmission (solid piece-linear graph in Figure 1) is above the targeted readmission rate (horizontal line in Figure 1) the software recommends that the patient *not* be discharged. If the point estimate lies below the target rate but the upper bound of the 80% confidence interval lies above the target readmission rate the software makes a "physician judgment" recommendation; that is, it makes no recommendation and leaves the decision up to the discretion of the physician but it provides much additional information (not contained in electronic medical records) on the readmission probability and six dynamically-selected key clinical variable charts, as explained below. Lastly, if the upper bound of the 80%

confidence interval lies below the targeted readmission rate the software recommends that the patient be discharged from the hospital.

In order to conduct our experiment we selected 30 patients from the sample of 3,202 patients used in the regression model. For selection of charts we partitioned the 3,202 patients into low, medium and high readmission risk categories and selected 10 patient charts from each of the three readmission-risk tiers.[10] We use patient charts that provide a clear test of the efficacy of the software by selecting patient charts for which there was enough variation in the daily recommendations of the decision support software. Figure 1 provides an example of a chart with these properties. For the sample patient shown in Figure 1, the decision support software's recommendations would be: (a) "do not discharge" on days 1 - 5; (b) "physician judgment" on days 6 - 10; and (c) "discharge" beginning on day 11.[11] Of course, in the experiment and eventually on the patient wards, a physician will get on day *t only* the recommendation for that day and the part of the time series of probabilities and clinical variables *from* day 1 *through* day *t*; some examples are presented in Figures 2 and 3 below.

In addition to discharge recommendations and readmission probabilities, the software dynamically displays six charts of selected patient information that the regression model indicates are most significant for the discharge status of the individual patient on that day of the hospital stay. The clinical variables displayed in the six charts for a patient would change during the course of the experiment to reflect the information that was most relevant to the physician's decision at that point in time. The six panels activated contain a temporal plot of the patient data as well as upper and lower bounds on the normal range of values for each dynamically-selected clinical variable (as in Figures 2 and 3 below).

4. Experimental Design and Protocol

The experimental design "crosses" the presence or absence of a 45 day constraint on the number of "experimental days" with three information and default conditions in a 2 by 3 design. Inclusion of the 45 day constraint (in three of the six treatment cells) increases the opportunity cost of keeping a patient longer in the hospital; this feature of the experiment is a stylized way of

[10] The partitioning was based on the historically observed readmission rates for the surgical procedures conducted with the patients. A "low risk" patient had a procedure with a historically observed readmission rate less than 10%, a "medium risk" patient was between 10% and 20%, and "high risk" patient was greater than 20%. These risks are not patient-specific but, instead, are associated with the complexity of the surgery and the procedure-specific potential for infection and other complications.

[11] The electronic medical record shows that this patient was actually discharged on day 12.

capturing the effect of a hospital's "capacity" on discharge decision making. The three information and default conditions (Baseline, Information, and Default) will be explained in detail below. We first explain features of the experiment that are present in all six treatment cells.

4.a Common Features of the Baseline, Information and Default Treatments

The information provided to subjects in all six treatment cells included clinical variables that were the same as they would get from a hospital's electronic medical records (EMR). Not only was the same information provided as in the hospital's EMR, we also used a graphical interface that was a facsimile of the EMR computer display screens.

The patient chart information used during the experiment was taken from the EMR for 30 patients included in our sample from the hospital clinical data warehouse. The patients' names and addresses had been removed from the charts and replaced with fictitious names and patient IDs. A subject in the experiment did not always have to review a patient's chart for the first few *calendar* days they were in the hospital if there was no realistic prospect for considering discharge during those days. The first *experimental* day on which a subject was asked to review a chart was randomly selected to be between one and four days before the estimated discharge model would first recommend that the patient be discharged; this one to four day period was independently selected for each of the 30 patient charts. During an experimental session, the 30 patient charts were presented in a random order that was independently drawn for each subject.

In order not to have the experimental subjects' discharge decisions be conditioned by the previous attending physicians' decisions, the dates of actual discharge were removed from the patient charts. Within the experiment it was, of course, possible that a patient could be retained longer than the observed length of stay in the EMR. Therefore, we constructed continuation charts for all 30 patients that imputed an extra five days of possible stay.[12]

In all treatments, the subjects were informed that they should assume that a patient was being managed at the appropriate standard of care while in the hospital and that the subjects were *not* being asked to speculate about additional tests or procedures that they might want to order. Instead, they were asked to make the hospital discharge decisions on the basis of the clinical information contained in the patient chart. The subjects were informed that they would be asked to make a series of choices between two options: "Discharge Patient" and "Do NOT Discharge

[12] The percentages of patients discharged in the experiment that occurred during the continuation chart periods are 7.76, 8.77 and 7.48 in the baseline, information and default treatments.

Patient." Any patient not discharged would return for consideration on the next "patient day" with updated chart information. Any patient who was discharged could turn out to be successfully discharged or, alternatively, could be readmitted. The likelihood of readmission was based on the estimated probit model.[13] After a patient was readmitted in the experiment, the patient remained in the hospital for at least two days. The subject had to view the readmitted patient's chart during the first two days following readmission but was not allowed to discharge the patient before the third day after readmission.

At the beginning of an experiment session, the subjects were welcomed to the decision laboratory by one of the researchers who self-identified as a medical doctor and explained that the research was supported by an NIH grant..[14] Each subject was informed that they would make a series of choices between the two options, "Discharge Patient" and "Do NOT Discharge Patient." They would be allowed to make at most a total of 30 choices of the Discharge Patient option. A choice of Discharge Patient would be "successful" if the patient was not readmitted. A choice of Discharge Patient would be "unsuccessful" if the patient was readmitted. A subject was paid $5 for each successful discharge and nothing for unsuccessful discharges. Each experiment session could last no more than 2 hours. In three treatment cells there was an additional constraint that the subject could not participate in more than 45 "experimental days". In contrast, there was no limit on the number of experimental days that a subject could use to make up to 30 discharge decisions in the other three treatment cells although a two-hour time limit applied. The two-hour time limit, however, was not a binding constraint for any subject. The purpose of the 45 experimental day constraint was to increase the opportunity cost of not discharging a patient. This 45 experimental day constraint was binding for some of the subjects in the three treatment cells in which it applied.

---

[13] In the case that a patient was readmitted after being discharged in the experiment the subject was presented with a readmission chart for the patient. The readmission chart was based on the observed complications following discharge within the population of patients served by the hospital. Subjects were informed of the complication that required readmission and the patient chart data were altered to be consistent with the presence of the complication as reflected in the empirical evidence reported in Kassin, et al. (2012). Each patient's chart was altered for only the first three to five days of their stay after readmission and the remaining chart days conformed to their observed data prior to being discharged.

[14] The subjects read and signed the IRB-approved consent form and subsequently began reading the subject instructions on their computer monitors. Subject instructions for the experiment can be found at (http://excen.gsu.edu/jccox/subjects.html).

A subject begins each experimental day by selecting patients from a list on a screen that displays fictitious patient IDs and names and real summary information from each of three patient charts that includes patient age, sex, and length of stay in the hospital (up to the current experimental day) taken from electronic medical records. After selecting a patient, a subject in the experiment gets access to that patient's chart information. The information is presented in several charts that are facsimiles of the patient's actual charts in the EMR (but, of course, de-identified). The charts for one patient, taken from the subject instructions, are reproduced in Appendix A: Figures A1 to A4 are facsimiles of "pages" from the EMR for the patient.

4.b Idiosyncratic Features of the Baseline, Information and Default Treatments

In the Baseline Treatment, a subject makes the discharge decisions using only the information in the EMR (contained in appendix Figures A1 to A4). The default option in the Baseline Treatment is the same as in current medical practice: the patient remains in the hospital unless the physician with authority actively decides to discharge the patient. The Information Treatment presents all of the EMR-facsimile screens used in the Baseline Treatment plus additional screens with selected patient information and a recommendation about the discharge decision. The default option in an Information Treatment is the same as in the Baseline treatment. The Default Treatment presents all of the same information as in the Information Treatment, including a discharge recommendation, but uses a different default option. In the Default Treatment, a recommendation by the decision support software to discharge or not to discharge the patient would be automatically implemented unless the decision maker overrules the recommendation and enters an explanation.

The subjects enter their decisions on screens that differ across the three information and default treatments. The decision screen for the Baseline Treatment (shown in Figure A5) includes only the patient's ID, name, age, and sex and two buttons to be clicked in order to record a decision whether to discharge the patient on the experimental day recorded at the top of the screen. These buttons are labeled "Discharge Patient" and "Do NOT Discharge Patient." If the subject clicks on the Do NOT Discharge Patient button, the patient remains "in the hospital" and reappears in the subject's list of patients on the following experimental day. If the subject clicks on the Discharge Patient button, the patient is discharged. In the event of a successful discharge, the subject is paid five dollars. In the event of an unsuccessful discharge, the subject receives no payment and the patient is readmitted and reappears in the subject's list of patients.

There are three alternative decision screens for the Information Treatment. Which decision screen a decision maker encounters depends on the recommendation of the decision support software for the patient on that day. Figure 2 shows three screens for the same patient ("Linda Doe") for three days on which the decision support software provides different recommendations. The recommendation about discharging the patient appears at the bottom left of a screen. The left side of a decision screen shows probabilities of readmission if the patient were to be discharged on any experimental day up to the present decision day, as estimated by the decision support software. The dots at kinks in the piecewise linear graph show point estimates of the probability of readmission. The vertical dashed lines that pass through the dots (at kinks) correspond to the 80% confidence intervals of the readmission probability. The horizontal line shows the target readmission probability for patients with the diagnosis code of this patient. The six charts on the right two-thirds of a screen show the probabilistically most important clinical variables for the discharge decision for this specific patient on the present experimental day (the last day shown on a screen).

In case of a negative recommendation the decision maker encounters a decision screen like the one shown in the *top panel* of Figure 2 that reports the recommendation Do Not Discharge Patient at the bottom left of the screen. For the selected patient, the left part of the figure shows point estimates and 80% confidence intervals that lie entirely above the horizontal line showing the target readmission rate (in this case at 17%); that is why the decision support software makes the negative recommendation. For this patient the WBC (white blood cell count) chart on the right makes highly salient a clinical variable that is indicative of a negative discharge recommendation; this indicator of infection is highly elevated at the middle of day 2, at the boundary of normal in the middle of day 3, and still somewhat elevated in the middle of day 4. The subject enters her decision by clicking on one of the two buttons at the lower right of the screen.

The middle panel of Figure 2 shows an example where the decision support software does not make any recommendation whether to discharge the patient; instead, it exhibits the "recommendation" for days 8, 9 and 10 as Physician Judgment.[15] Recall that a Physician Judgment "recommendation" occurs when the target readmission rate falls between the point estimate and the upper bound on the 80% confidence interval for the readmission probability.

[15] Note that the clinical variables exhibited in the top and middle panels of Figure 2 are not all the same variables.

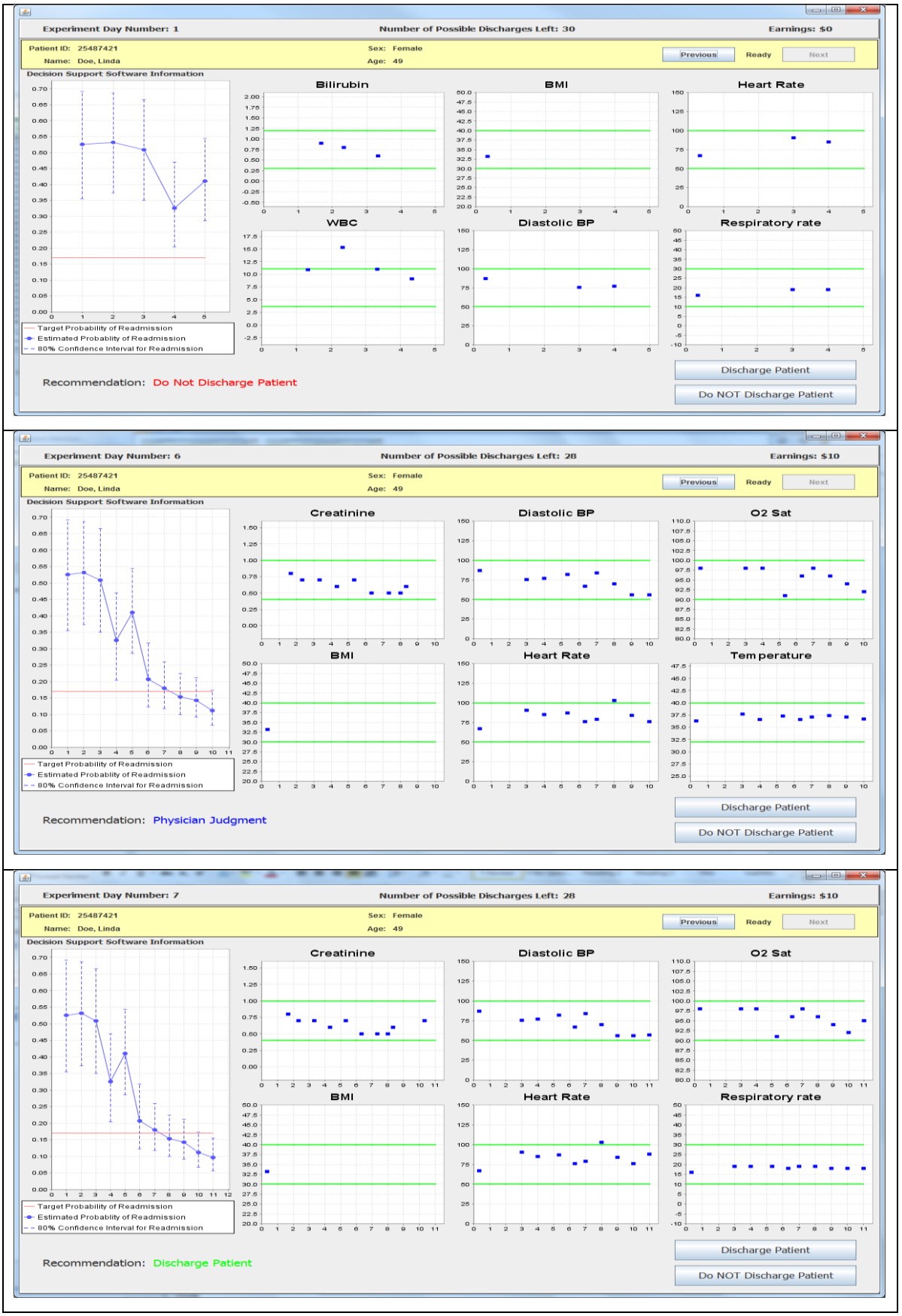

Figure 2. Information Treatment Decision Screens with Alternative Recommendations

The software first recommends that the patient be discharged on the experimental day in which the top of the 80% error bar dips below the target readmission rate. This criterion reflects choice of a 10% type 1 error. In the example shown in the bottom panel of Figure 2, the first day on which the top of the error bar drops below the target readmission rate is experimental day 11. The software's recommendation on that day is Discharge Patient.

As explained above, the Default Treatment presents subjects with the same information as the Information Treatment. The only difference between the two treatments is change in the default option. For example, in case of a positive recommendation the decision screen for the Default Treatment, shown in the top panel of Figure 3, presents the decision maker with two buttons labeled "Overrule and Enter Reasons" and "Discharge Patient." This contrasts with the decision buttons for the Information Treatment shown in Figure 2, which are "Discharge Patient" and "Do NOT Discharge Patient." In either treatment, if the decision maker clicks on Discharge Patient then that occurs without any further action. If, in the Information Treatment, the decision maker clicks on "Do NOT Discharge Patient" then the patient remains in the hospital and the experiment immediately continues. In contrast, if in the Default Treatment the decision maker clicks on Overrule and Enter Reasons then he is presented with the response screen shown in the bottom panel of Figure 3 that requires entry of the reasons for overruling the positive recommendation. In case that the decision support software produces a negative (do not discharge) recommendation, the decision screen has two buttons labeled Do NOT Discharge and Overrule and Enter Reasons. Clicking on the button to overrule opens a screen requiring entry of the reasons for overruling the negative recommendation. These two response screens are shown in Figures A6 and A7 in Appendix A. In case that the decision support software produces the recommendation Physician Judgment, the subject's decision screen looks exactly the same as it does for this recommendation in the Information Treatment, as shown in the middle panel of Figure 2.

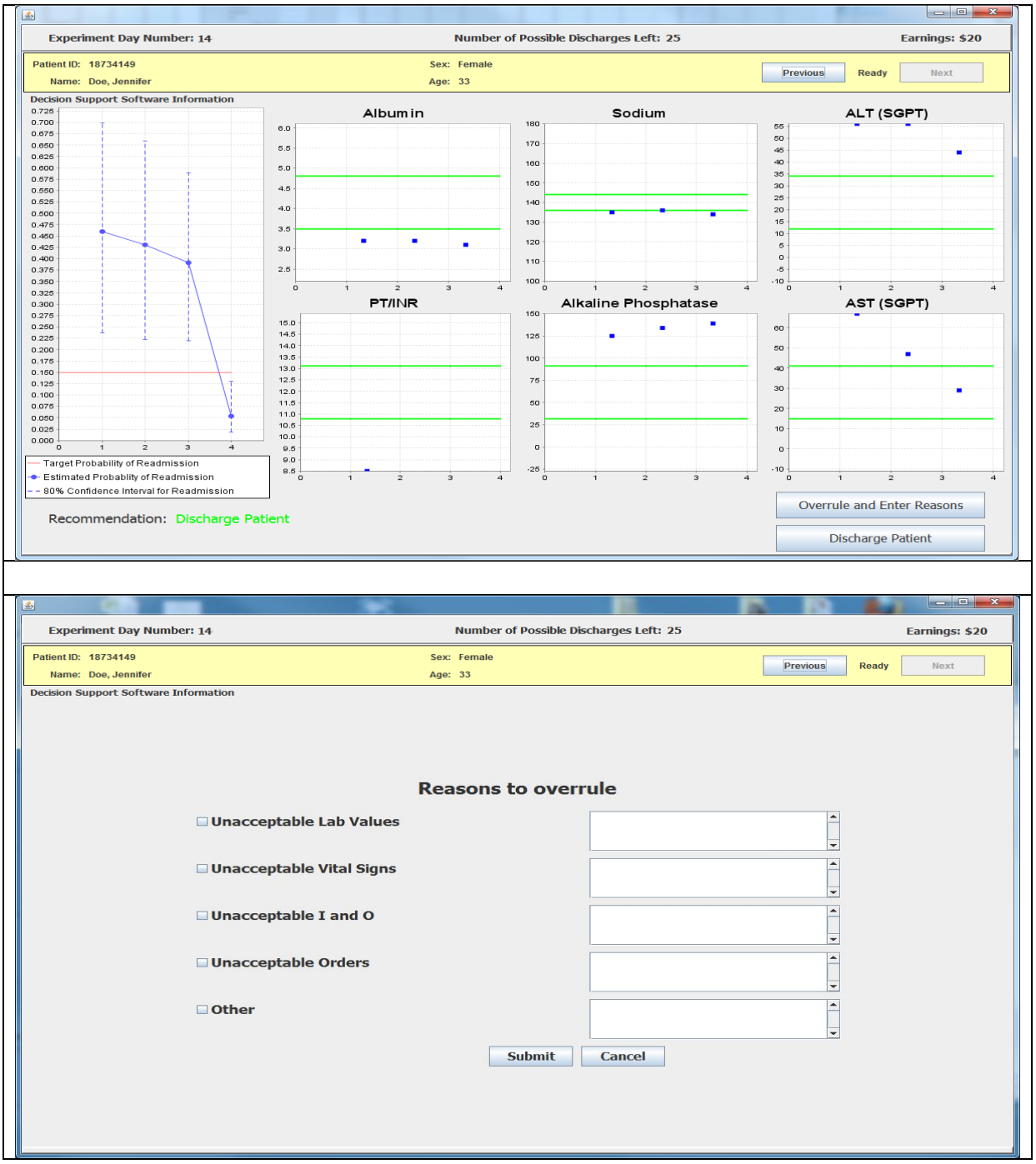

Figure 3: Default Treatment Decision Screens for Positive Recommendation

5. Results from the Experiment

One hundred and twenty-five subjects participated in the experiment; twenty of these subjects were resident physicians and the rest were fourth-year medical students. Subjects were distributed almost equally across the baseline (43 subjects), information (42 subjects), and default (40 subjects) treatments. The overall number (64) of female participants was similar to

the number (61) of male subjects, as was the gender composition across treatments ((21F, 22M); (20F, 22M) and (23F, 17M); Pearson chi2(2)=0.95, p-value=0.62). Academic performance of subjects who participated in different treatments was at comparable levels.[16]

After making their discharge decisions in one of the six treatment cells, subjects completed an online questionnaire that was embedded in the experiment software. The questionnaire elicited demographic information and also included hypothetical response questions about risk attitudes.[17] After completing the questionnaire, subjects exited the lab one at a time to be paid in cash in private.

We ran two designs that differ from each other only with respect to whether there was a constraint (of 45 days) on the number of experimental days. Fifty-four subjects (out of 125) participated in the design with the 45 experimental days constraint and 71 participated in the design with no constraint on the number of experimental days.

Data from our experiment provide support for efficacy of the decision support software with respect to four measures of performance: subject earnings, quality of service (readmission rate), hospital length of stay, and time efficiency (number of experimental days utilized to make a certain number of discharges). We report several ways of describing the data and statistical analysis for significance of treatment effects.

5.a Decision Time Efficiency and Daily Experiment Earnings

In the treatment cell without the 45 day constraint, subjects took on average 54 experimental days to finish the experiment (i.e. to make 30 discharges) in the baseline treatment but in the information and default treatments they were able to complete the task of making 30 discharges within 47 and 42 experimental days, respectively, an improvement in time efficiency of 7 and 12 experimental days. The null hypothesis of equal time efficiency across treatments is rejected (chi2=6.42; p-value=0.04 according to Kruskal-Wallis test).[18] Data from treatments in which the

[16] Reported average grades for previous undergraduate studies are 3.68, 3.64 and 3.74 in the baseline, information and default treatments (Kruskal-Wallis test: chi2 (corrected for ties)=4.84, p-value=0.09); for medical graduate studies we see 3.59, 3.57 and 3.46 (Kruskal-Wallis test: chi2(corrected for ties)=3.32, p-value=0.19).

[17] The questionnaire can be found at (http://excen.gsu.edu/jccox/subjects.html).

[18] Since in the 45 constrain design subjects couldn't go above 45 days we are excluding these data in the analysis of time efficiency in the main text because of potential bias. If we include data from the design with 45 day constraint, we still get higher time efficiency in completing the task of 30 discharges: 49, 44, and 40 experimental days, respectively, for the baseline, information, and default treatments. According

discharge support software is used are significantly more efficient than the baseline but the effect is stronger in the default treatment (one sided p-values reported by the t-test are 0.043 and 0.004 when the baseline is compared with the information and default treatment, respectively). Figure 4 shows cumulative distributions of experimental days in the baseline and default treatment.

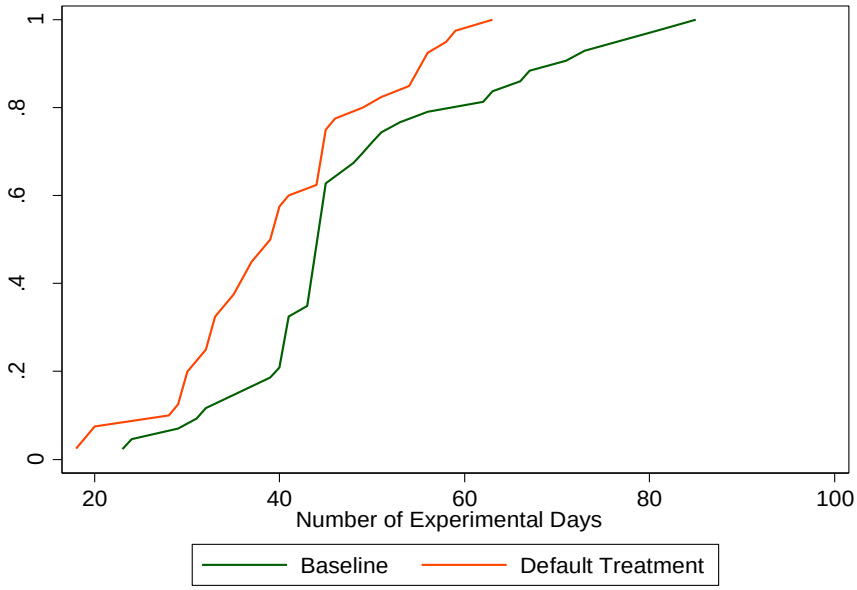

Figure 4. Cumulative Distributions of Observed Experimental Days

Average subject earnings per experimental day were $2.83, $3.15 and $3.62 in the baseline, information and default treatments, as reported in the top panel of Table 1 (with standard deviations in braces). We see that subjects' daily earnings are highest in the Default treatment and lowest in the Baseline treatment. We ran one-way analysis of variance by ranks (Kruskal-Wallis test) to see whether daily earnings across different treatments come from the same distribution. This null hypothesis is rejected by this test (chi-squared statistic is 12.33, two-sided p-value is 0.002).

Next, we ask which treatments are responsible for this rejection. The means of the ranks of daily earnings of three treatments are shown in the lower panel of Table 1. The p-values for each pairwise comparison are shown in the bottom two rows. If we use (Bonferroni) adjusted p-values for multiple comparisons, we conclude that baseline and default data on earnings per day

---

to the Kruskal-Wallis test, the null hypothesis of equal distributions of experimental days across treatments is rejected (chi2 (correcting for ties) =8.32, p-value=0.016).

are coming from different distributions; subjects in the default treatment are earning more per experimental day than subjects in the baseline.

Table 1. Comparisons of Daily Earnings across Treatments

	Baseline	Information	Default
Number of Subjects	43	42	40
Mean {st.dev.}	\$2.83 {0.94}	\$3.15 {0.94}	\$3.62 {1.16}
Kruskal-Wallis Test			
RankMean	49.22	63.64	77.14
Information	0.033	--	
Default	0.0002***	0.046	--

(Rank Means and p-values correspond to Kruskal-Wallis test; the adjusted p-value for multiple comparisons is 0.008 in case of all pairwise comparisons and 0.0125 in case the baseline is treated as a control group, i.e. comparing the baseline to the other two.)

So far, the analysis of data at the aggregate level has been focused on the treatment effects. We are also interested in other features of the experimental design and individuals' characteristics that are correlated with higher daily earnings. So we ran linear regressions (with robust standard errors) of daily earnings as a dependent variable and treatment dummies, a dummy for the 45-day constraint, and subject demographic variables as right-hand variables. The parameter estimates are as follows (standard errors in braces):[19]

$$\begin{aligned} Daily\ Earnings = {} & \underset{\{1.424\}}{3.63^{**}} - \underset{\{0.164\}}{0.28^{*}} \times D_{45day} - \underset{\{0.231\}}{0.46^{*}} \times D_{resident} - \underset{\{0.366\}}{0.15} \times GPA_{undergrad} + \underset{\{0.142\}}{0.15} \times GPA_{med} \\ & - \underset{\{0.152\}}{0.09} \times D_{music} - \underset{\{0.172\}}{0.06} \times D_{athlete} - \underset{\{0.153\}}{0.47^{***}} \times D_{female} + \underset{\{0.043\}}{0.02} \times risk + \underset{\{0.186\}}{0.72^{***}} \times D_{default} + \underset{\{0.180\}}{0.31^{*}} \times D_{info} \end{aligned}$$

Daily earnings are lower for residents (by 46 cents) and female physicians (by 47 cents); they are also lower (by 28 cents) in the treatment with the 45-day constraint on the maximum number of experiment days. There are insignificant effects of musical training and record of competitive athletics (although such characteristics are selected for in admissions to the surgical specialty). With respect to treatment effects, consistent with findings at the aggregate level of data analysis, daily earnings of subjects in the default and information treatments are 20% and 9% higher than

[19] F(10,108)=2.94 (p-value=0.003), 119 observations. Six observations were dropped during the regression because of incomplete demographics responses. The estimates of the robust regression without demographics for the information and default treatments are 0.31 (p-value = 0.10) and 0.65 (p-value = 0.001) and F(4,120)= 4.08 (p-value=0.004).

in the baseline (p-values are 0.000 and 0.087).[20] Subjects are also making more money per day in the default treatment than in the information treatment.[21] We conclude that:

Result 1. *Use of the decision support software, with or without making the software's recommendation the default option:*

(i) increases decision time efficiency; and

(ii) increases daily earnings.

One may wonder whether the increased decision time efficiency that we observe in the treatments has a negative effect on the quality of care. Given the design of our experiment, lower quality would be manifested in higher readmissions. Readmission rate is one of the factors that affect the ranking of a hospital and it is also one that has attracted increasing attention from Medicare, including fines for excess readmissions beginning in October 2012. In the following section we look closely at the interaction between different treatments and readmission rates in our experiment.

5.b Readmissions as an indicator of the quality of care

An earlier discharge is not an indicator of better discharge decision making if it decreases the quality of care. An indicator of the quality of care is the readmission rate since a premature discharge increases the likelihood of an unplanned but necessary readmission. Averages of readmission rates of all regular patients[22] observed across treatments are 10.21%, 10.40% and 9.84%, respectively, for the baseline, information and default treatments.[23] For regular patients with high levels of *targeted* readmission probabilities (at least 17%) the mean readmission rates

[20] Note that $0.72/3.63 \cong 0.20$ and $0.31/3.63 \cong 0.09$ using the coefficient estimates.

[21] Estimates for the information and default treatments are different from each other ($F(1,108)=4.80$, p-value=0.031.)

[22] A patient is called "regular" if he has not previously been readmitted.

[23] If we include data from the patients after they are readmitted we get: averages of readmission rates per patient observed across treatments are 9.6%, 9.7% and 9.2%. For patients with high levels of targeted readmission probabilities (at least 17%) the mean readmission rates are 12.4%, 11.7% and 10% for the baseline, information and default treatments. Recall that the discharge software was developed to reduce the readmission rates by 10% of the current level. The "current" levels in our experiment correspond to the readmission rate of 9.6% and 12.4% for high risk patients. The effectiveness of the discharge software requires that the mean of the readmissions be 8.6%; for high risk patients the target figure is 11% .

are 13.49%, 12.70% and 10.80% for the baseline, information and default treatments.[24] We observe that the discharge decision support software combined with change in the default option helps reduce the overall mean of readmissions by 3.6% and the rate for high risk patients by 20%. In contrast, there is no significant difference between the baseline and information treatments; hence providing only daily likelihood of readmissions at the patient level seems to be

Table 2. Probit Regressions of Readmissions for Regular Patients

VARIABLES	High Risk Patients		All Patients	
Rec. LOS	…	0.005	…	-0.009***
	…	(0.185)	…	(0.000)
Understay	0.004	0.000	0.016***	0.025***
	(0.617)	(0.976)	(0.000)	(0.000)
Overstay	-0.005	-0.003	-0.004	-0.004
	(0.498)	(0.660)	(0.311)	(0.252)
Information	-0.020	-0.019	-0.000	0.001
	(0.386)	(0.397)	(0.972)	(0.958)
Default	-0.047**	-0.045**	-0.015	-0.015
	(0.022)	(0.027)	(0.211)	(0.190)
Female	-0.013	-0.015	-0.006	-0.004
	(0.492)	(0.439)	(0.570)	(0.723)
Athlete	0.019	0.019	0.006	0.007
	(0.384)	(0.396)	(0.592)	(0.563)
Musical	-0.034*	-0.035*	-0.023**	-0.021**
	(0.075)	(0.066)	(0.027)	(0.040)
Medical GPA	0.013	0.013	0.006	0.007
	(0.297)	(0.314)	(0.537)	(0.508)
Undergrad GPA	0.003	0.004	-0.017	-0.021
	(0.949)	(0.929)	(0.511)	(0.431)
Risk Avers. Index	-0.017***	-0.017***	-0.003	-0.003
	(0.003)	(0.002)	(0.200)	(0.262)
Resident	-0.013	-0.016	0.010	0.014
	(0.629)	(0.563)	(0.597)	(0.446)
Target Probability	4.207***	4.480***	0.762***	0.954***
	(0.000)	(0.000)	(0.000)	(0.000)
45 Day Constraint	0.051**	0.050**	0.044***	0.043***
	(0.030)	(0.031)	(0.000)	(0.001)
Nobs	1,063		3,197	

[24] One third of "patients" in our experiments had a targeted probability of readmission higher than 0.17; we call such patients "high risk" patients.

insufficient for driving readmissions down. To test for statistical significance we ran probit regressions with binary dependent variable that takes value 1 if a regular patient is readmitted. Covariates include subjects' demographics, a risk aversion index,[25] patients' targeted probabilities, whether the patient was discharged before the first recommended discharge day (Understay = 1) or after that day (Overstay = 1), and the recommended length of stay until first discharge recommendation (Rec. LOS). Table 2 reports the estimated coefficients (and p-values of the probit regressions) with clusters at subject level for the high risk patients (columns (1) and (3)) and all patients (columns (2) and (4)). The probit regressions reported in columns (3) and (4) include the Rec. LOS variable, reported in the first row of the table, while regressions reported in the other columns exclude it.

Referring to the estimates of Understay, we find that for data from all patients (right two columns) the patient Understay variable has a positive effect on readmission while the estimate of the Overstay variable is insignificant. This gives us the second result:

Result 2. *Discharging a patient earlier than recommended significantly increases the likelihood of unplanned readmission while later discharge does not significantly decrease it.*

Readmissions are significantly higher in the design with a constraint (of 45 days) on the maximum number of experimental days. Subjects with musical training or a higher risk aversion index have lower readmissions with high risk patients. Treatment effects are consistent with the aggregated data figures reported in the preceding paragraph; the default treatment induces a significant reduction in the readmissions of high risk patients. Probabilities of readmissions in the default treatment (holding all other covariates at the means) are 4.5% - 4.7% lower than in the baseline for the high risk patients.[26] We conclude that:

Result 3. *Use of the decision support software and making the discharge software's recommendation the default option reduces readmissions of high risk patients.*

---

[25] In the hypothetical ten ordered tasks in our study, a risk neutral subject switches from the safer option to the riskier option in task 5. The risk index variable is the difference between the number of the task that a subject switches (for the first time) from choosing the safer option to choosing the riskier one and task five. Hence, the risk index is negative for a risk lover and positive for a risk averse subject (the later the switch the more risk averse the subject).

[26] Probabilities of readmissions of regular patients in the information treatment (holding all other covariates at the means) are 1.7% and 0.57% lower than in the baseline, but these figures are not statistically different from 0.

The recommended hospital length of stay (Rec. LOS) also has a significantly negative effect on readmissions[27] but the effect disappears when only the high risk patients are considered. We next turn our attention to study of hospital length of stay (LOS) across treatments.

5.c Hospital Length of Stay

The average figures for LOS of regular patients were 7.72, 7.15 and 6.64 for the baseline, information and default treatments.[28] The OLS estimates (and robust standard errors in braces) of the hospital length of stay for regular patients are[29]

$$\underset{}{\text{LOS}} = \underset{\{2.053\}}{7.75^{**}} - \underset{\{0.248\}}{0.32} \times D_{45day} + \underset{\{0.301\}}{0.84^{***}} \times D_{resident} - \underset{\{0.548\}}{0.29} \times GPA_{undergrad} - \underset{\{0.126\}}{0.13} \times GPA_{med}$$
$$+ \underset{\{0.220\}}{0.40^{*}} \times D_{music} - \underset{\{0.285\}}{0.08} \times D_{athlete} + \underset{\{0.217\}}{0.71^{***}} \times D_{female} + \underset{\{0.072\}}{0.06} \times risk - \underset{\{0.305\}}{0.92^{***}} \times D_{default} - \underset{\{0.275\}}{0.50^{*}} \times D_{info}$$

Female physicians and residents kept their patients in the hospital longer but their readmissions were not lower than others (as shown in Table 2). The hospital length of stay of regular patients is lower in both information (by a half day) and default (by almost one day) treatments while there are no higher readmission rates in these treatments (see Table 2 estimates for the Information and Default dummy variable parameters). We conclude that:

Result 4. *Use of the decision support software, with or without making the discharge software's recommendation the default option, reduces hospital length of stay without increasing readmissions.*

5.d Value of Decision Support Information

After they finished making all discharge decisions, subjects who participated in the default and information treatments were asked (by the experimental software) to report their ranking on a five-point scale (where higher is better) of the usefulness of being provided the estimated readmission probabilities and the 80% confidence intervals. Half of the subjects reported a score 3 or higher for both the point estimate and the 80% confidence interval.

---

[27] The estimated marginal effect is - 0.01 (std. error is 0.002) for model (4) in Table 2.

[28] If we include days in the hospital after a patient is readmitted in LOS then we get the following average hospital length of stay: 8.60, 8.11 and 7.42.

[29] F(10,118)=6.35 (p-value=0.00), nobs=3274, 119 clusters. If we include in LOS days in the hospital as a readmitted patient we still get similar results; the estimates for the information and the default treatments are -0.48 and -1.05.

The difference between daily earnings of the subjects who gave a score of three or higher for the usefulness of the point estimate (call it group H) and the average daily earnings of those who reported a lower score than three (call it group L) can be used as an economic measure of the effects of subjects' acceptance of the value of the readmission probability information. The mean daily earnings are \$3.15 and \$3.53 for groups L and H; the median figures are \$2.87 and \$3.42 for groups L and H. The Kolmogorov-Smirnov test rejects at 1% significance level (p-value is 0.001) the null hypothesis of daily earnings of the two groups being drawn from the same distribution. We conclude that:

Result 5. *Subjects who report they place relatively high value on usefulness of reported readmission probabilities earn significantly higher payoffs.*

We also looked at the economic significance of following a recommendation. If we exclude rounds when the software reports Physician Judgment, we find that the (average) compliance rates are 80.78% and 81.44% in the information and in the default treatments. In those treatments, subjects who discharged a patient when the recommendation was *not* to discharge earned \$4.06 whereas those subjects who discharged a patient when the recommendation was to discharge earned \$4.75, an increase of 17%. On average, total payoffs earned by subjects in the information and default treatment were \$130 and \$134.[30] The Tobit estimates of total earnings are[31]

$$\begin{aligned} Earnings = & \underset{\{23.73\}}{49.57}^{**} + \underset{\{11.26\}}{30.70}^{***} \times complyRate - \underset{\{2.71\}}{18.17} \times D_{45day}^{***} - \underset{\{4.00\}}{1.01} \times D_{resident} + \underset{\{6.01\}}{18.59}^{***} \times GPA_{undergrad} \\ & - \underset{\{1.98\}}{0.469} \times GPA_{med} + \underset{\{2.44\}}{3.04} \times D_{music} - \underset{\{2.73\}}{3.56} \times D_{athlete} - \underset{\{2.62\}}{3.91} \times D_{female} + \underset{\{0.68\}}{0.49} \times risk + \underset{\{0.20\}}{3.31} \times D_{default} \end{aligned}$$

where the "complyRate" variable is the rate of compliance, defined as the number of times that a subject's decisions were the same as the decision recommended by the decision support model

[30] In the baseline subjects earned on average \$127. On average in 60% of the cases subjects' decisions in the baseline are the same as the recommendations by the decision model (though this information was not given to the subjects).

[31] Log-likelihood is -282, LR chi2(10) is 58.04, number of observations is 79. There are five right censored observations, that is five out of 79 subjects earned the maximum amount of money, \$150. Three observations were dropped because of incomplete demographic responses. Tobit estimate of the complyRate variable when demographics are not included (all 82 observations included) is 35.18 (p-value=0.002).

divided by the total number of subject's decisions.[32] The coefficient estimate for the complyRate variable is positive and significantly different from 0 (p-value is 0.008).

Similarly, the difference between days a patient is kept in the hospital by subjects of groups L and H can be used as an indicator of the value of the information. The means are 3.43 and 2.82 days for groups L and H. Hence, in our experiment, the value of the information is a reduction of length of stay of 0.61 days per patient, a decrease of 18%. The Kolmogorov-Smirnov test rejects at 1% significance level (p-value is 0.000) the null hypothesis that the distributions of observed length of stay (from the first day the patient is seen by subjects) are the same for the L and H groups of subjects. We conclude that:

Result 6. *Subjects who comply with the software's recommendations have:*

(i) significantly higher earnings and

(ii) significantly lower average patient length of stay.

Together, Results 5 and 6 inform us that subjects obtain higher earnings in the experiment when they: (a) report that they place relatively high value on usefulness of the decision support software's readmission probability estimates; or (b) comply with the software's recommendations to discharge or not to discharge patients. These conclusions are especially clear in the default treatment. These results suggest another question concerning the reasons why decision support software does lead to better discharge decisions in the experiment. Does this occur because a high proportion of subjects are reluctant to overrule the recommendation of the software, most especially when they have to enter reasons for doing so? Some insight into this question is provided by analyzing only that subset of the data in which the software provides probability information and six dynamically selected patient charts but does not provide a recommendation to discharge or not to discharge the patient; these are the experimental days in patient charts in which the software's "recommendation" is Physician Judgment. We compare the mean outcomes from decisions in the baseline and default treatments using only the data from experiment days in which the "recommendation" is Physician Judgment and find that: mean total payoff is higher in the default treatment; mean experimental days is lower in the default treatment; and mean readmission rate for high risk patients is lower in the default

[32] In the construction of the variable complyRate, physicians' decisions when the recommendation is "physician judgment" are not included.

treatment. Furthermore, payoff per experimental day is significantly higher in the default treatment and the total number of experimental days is significantly lower in the default treatment. In this way we find that the information provided by the decision support software, *per se*, improves discharge decision making even in the absence of a definitive recommendation about discharge, at least in the context of the default treatment. This suggests that the default treatment may better focus decision makers' attention on the new information (shown in the decision screen charts) provided by the decision support software even when the software does not state a recommended decision.

5.e Effects of Capacity Constraint

The probit regressions in Table 2 show higher readmission rates in the presence of the 45-day constraint. The OLS regression shows, however, that LOS of regular patients is not affected by the 45-day constraint. But these figures are confounded with differences in numbers of patients discharged between the no-constraint and 45-day constraint treatments. We seek to isolate the effects on readmissions per patient discharged. For each subject we constructed a new variable, ReadmissionRate, which is the number of patients readmitted divided by the total number of discharges. Figure 5 shows kernel densities of this variable for both designs, with and without

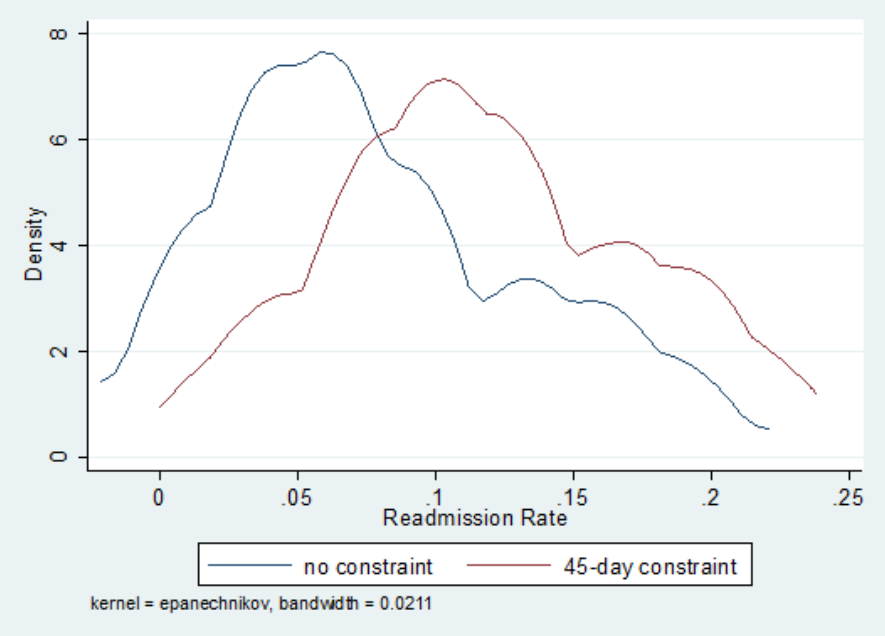

Figure 5. Kernel Densities for Readmissions/Discharges

the 45-day constraint. It can be easily seen that the readmission rates are higher for subjects who were making discharge decisions under the 45-day constraint. The mean readmission rates are 7.5% and 11.9%, respectively, in the no-constraint and 45-day constraint treatments; the differences are statistically significant at 1% (p-value reported by t-test is 0.0001; nobs are 51 and 54).[33]

5.f Reduction in Length of Stay

According to their electronic medical records, the 30 patients whose de-identified charts are used in our experiment were kept in the hospital an average of 9.37 days. In our baseline experimental treatment, the average length of stay (LOS) was 8.6 days. This reduction may have resulted from the exclusive focus on the discharge decision created by the experimental environment. The average LOS in the default treatment was 7.42 days, which is 14% lower than the baseline number of 8.6 days. This reduction in LOS did not produce higher readmission rates since the rates for the baseline and default treatments were, respectively, 10.2% and 9.8%

6. Concluding Remarks

The research reported in this paper is part of an ongoing collaboration between economists affiliated with Georgia State University's Experimental Economics Center and surgeons affiliated with Emory University's School of Medicine. We began our collaboration by researching the causes of surgical patient readmissions at a large southeastern teaching hospital (Kassin, et al. 2012). We subsequently elicited the discharge criteria reported by physicians (Leeds, et al., 2013) and compared the responses to econometrically-identified predictors of successful and unsuccessful discharges (Cox, et al., 2013). Although many stated criteria coincide with significant predictors, as one would expect, various inconsistencies were identified which suggested the importance of research on improving discharge decision making. The present paper reports development and experimental testing of decision support software that incorporates evidence-based criteria for making hospital discharge decisions.

[33] If we include residents (who participated only in the no-constraints design), we still find that the mean readmission rates are 7.8% and 11.9% respectively in the no-constraint and 45-day constraint treatments. The t-test rejects the null hypothesis in favor of the alternative hypothesis of higher readmissions in the presence of the 45-day constraint at 1% (p-value is .0001). Numbers of observations are 71 and 54 in the no-constraint and 45-day constraint treatments.

The hospital discharge decision plays a central role in the increasingly important interplay between the quality of healthcare delivery and medical costs. Premature discharge can lead to unplanned readmission with higher costs and questionable quality of care. Needlessly delayed discharge wastes increasingly expensive healthcare resources.

The project includes four component parts. First, we estimate an econometric model of the clinical and demographic variables that are significant predictors of successful or unsuccessful discharge. Secondly, we develop a decision support model that is incorporated into user-friendly software. Thirdly, we conduct laboratory experiments on the efficacy of this software. These first three stages of the project are reported in this paper. Completion of the laboratory experiment is ethically and practically necessary before proceeding to the fourth stage of the project, which is using the discharge decision support software in an intervention on hospital patient wards (if it has been found to be efficacious in the laboratory experiment).

The laboratory experiment uses a 2 by 3 design that crosses presence or absence of a 45 experimental day constraint with baseline, information, and default treatments. Introduction of the 45-day constraint is intended to put the subjects under time pressure in making their discharge decisions. The baseline treatment presents subjects only with the kind of information that they receive from currently-used electronic medical records (EMR); indeed, the subject screens used in the baseline treatment are facsimiles of EMR screens. The information treatment uses these same EMR-facsimile screens plus a new screen that reports information provided by the decision support model. The new screen shows point estimates of quasi-marginal readmission probabilities and their 80 percent error bounds for experimental days prior to and including the relevant experimental decision day. The new screen displays six charts of dynamically-selected clinical variables that the regression model indicates have the highest quasi-marginal effects for predicting outcomes from discharge of that patient on that day during their hospital stay. Finally, the new screen shows one of three recommendations, discharge, physician judgment or do not discharge, that are based on the relationship between the target readmission probability for patients with the relevant diagnosis code and the readmission probability point estimate and 80 percent error bounds for the patient whose data are under consideration. The default treatment differs from the information treatment by changing the default option for patient discharge. In the information treatment, the default option is that the patient is *not* discharged unless the decision maker makes the decision to discharge regardless of the recommendation of the decision support model. In contrast, in the default treatment the patient is discharged or *not* discharged according to the recommendation of the decision support

software unless the decision maker overrides that recommendation and provides reasons for rejecting it.

Data from our experiment provide support for efficacy of the decision support software with respect to the two central performance measures for hospital discharge decision making, readmission rate and length of stay. The data also provide support for effectiveness of the decision support software in promoting time efficiency in making discharge decisions and for the traditional experimental economics measure of subject earnings in the experiment. The decision support software is more effective in the default treatment than in the information treatment; in other words, combining the information provided by the decision support software with making the software's recommendation the default option is more effective in promoting better discharge decisions than simply providing the information. Subjects perform generally better in the absence than presence of a ("45 experiment day") constraint that puts them under time pressure.

Further research collaboration is in progress. The next stage of research on the hospital discharge decision involves patient ward intervention. This requires development of a version of the discharge decision support software that can interact with electronic medical records systems in real time and development of the protocol for the intervention.

References

Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, "HCUP Nationwide Inpatient Sample". http://hcupnet.ahrq.gov/HCUPnet.jsp Downloaded December 14, 2012.

Amarasignham, R., Moore, B.J., Tabak, Y.P., Drazner, M.H., Clark, C.A., Zhang, S., Reed, W.G., Swanson, T.S., Ma, Y. and E.A. Halm. 2010. "An Automated Model to Identify Heart Failure Patients at Risk for 30-Day Readmission or Death Using Electronic Medical Record Data," *Medical Care* 48(11): 981-988.

Anderson, Gerard F., and Earl P. Steinberg. 1985. "Predicting Hospital Readmissions in the Medicare Population," *Inquiry* 22(3): 251-258.

Anderson, D., Golden, B., Jank, W., Price, C. and E. Wasil. 2011. "Examining the Discharge Practices of Surgeons at a Large Medical Center," *Health Care Management Science* 14(4): 338-47.

Anderson, D., Golden B., Jang, W. and E. Wasil. 2012. "The Impact of Hospital Utilization of Patient Readmission Rate," *Health Care Management Science* 15(1): 29-36.

Bowles, K.H., Foust, J.B. and M.D. Naylor. 2003. "Hospital Discharge Referral Decision Making: A Multidisciplinary Perspective," *Applied Nursing Research* 16(3): 134-43.

Bright, T.J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R.R., Samsa, G., Hasselblad, V., Williams, J.W., Musty, M.D., Wing, L., Kendrick, A.S., Sanders, G.D. and D. Lobach. 2012. "Effect of Clinical Decision-Support Systems," *Annals of Internal Medicine* 157(1): 29-44.

Caminiti, C., Meschi, T., Braglia, L., Diodati, F., Iezzi, E., Marcomini, B., Nouvenne, A., Palermo, E., Prati, B., Schianchi, T. and L. Borgis. 2013. "Reducing Unnecessary Hospital Days to Improve Quality of Care Through Physician Accountability: A Cluster Randomised Trial," *BMC Health Services Research* 13:14.

Chandra, A., Gruber, J. and R. McKnight. 2010. Patient Cost-Sharing and Hospitalization Offsets in the Elderly. *American Economic Review* 100(1): 193-213.

Chandra, A., Jena, A.B. and J.S. Skinner. 2011. The Pragmatist's Guide to Comparative Effectiveness Research. *Journal of Economic Perspectives* 25(2): 27-46.

Chandra, A. and J. Skinner. 2012. Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature* 50(3): 645-680.

Cebul, R.D., Rebitzer, J.B., Taylor, L.J. and M.E. Votruba. 2008. Organizational Fragmentation and Care Quality in the U.S. Healthcare System. *Journal of Economic Perspectives* 22(4): 93-114.

Center on Budget and Policy Priorities, 2013. "Policy Basics: Where Do Our Federal Tax Dollars Go? Revised April 12, 2013. http://www.cbpp.org/cms/?fa=view&id=1258

Cox, James C., Schnier, Kurt E., Leeds, Ira, Sadiraj, V. and Sweeney, John F. 2013. "Reducing Hospital Readmissions through Improved Discharge Decision-Making," Working Paper.

Cutler, D.M., McCellan, M. and J.P. Newhouse. 1998. What Has Increased Medical-Care Spending Bought? *American Economic Review, Papers and Proceedings* 88(2): 132-36.

Demir, E. 2012. "A Decision Support Tool for Predicting Patients at Risk of Readmission: A Comparison of Classification Trees, Logistic Regression, Generalized Additive Models, and Multivariate Adaptive Regression Spline," *Working Paper, University of Hertfordshire Business School*.

Feldstein, M.S. 1971. Hospital Care Inflation: A Study of Nonprofit Price Dynamics. *American Economic Review* 61(5): 853-872.

Feldstein, M.S. 1973. The Welfare Loss of Excess Health Insurance. *Journal of Political Economy* 81(2): 251-80.

Feldstein, M.S. 1977. Quality Change and the Demand for Hospital Care. *Econometrica* 45(7): 1681-1702.

Finkelstein, A. 2007. The Aggregate Effects of Health Insurance: Evidence From the Introduction of Medicare. *Quarterly Journal of Economics* 122(1): 1-37.

Fuchs, V.R. and Millstein, A. 2011. "The $640 Billion Question — Why Does Cost-Effective Care Diffuse So Slowly?" *New England Journal of Medicine*, 2011, 364, 1855 – 1987.

Graumlich, J.F., Novotny, N.L., Nace, G.S. and J.C. Aldag. 2009a. "Patient and Physician Perceptions After Software-Assisted Hospital Discharge: Cluster Randomized Trial," *Journal of Hospital Medicine* 4(6): 356-363.

Graumlich, J.F., Novotny, N.L., Nace, G.S., Kaushal, H., Ibrahim-Ali, W., Theivanayagam, S., Scheibel, L.W. and J.C. Aldag. 2009b. "Patient Readmission, Emergency Visits, and Adverse Events After Software-Assisted Discharge From Hospital: Cluster Randomized Trial," *Journal of Hospital Medicine* 4(7): E11-E19.

Hall, R.E. and C.I. Jones. 2007. The Value of Life and the Rise in Health Spending. *Quarterly Journal of Economics* 122(1): 39-72.

Heggestad, Torhild. 2002. "Do Hospital Length of Stay and Staffing Ratio Affect Elderly Patients' Risk of Readmission? A Nation-Wide Study of Norwegian Hospitals," *Health Services Research* 37(3): 647-665.

Hunt, D.L., Haynes, R.B., Hanna, S.E., and K. Smith. 1998. "Effects of Computer-Based Clinical Decision Support Systems on Physician Performance and Patient Outcomes," *Journal of the American Medical Association* 280(15): 1339-1348.

Hwang, S.W., Li, J., Gupta, R., Chien, V. and R.E. Martin. 2003. "What Happens to Patients Who Leave Hospital Against Medical Advice?," *Canadian Medical Association Journal* 168: 417-420.

Institute of Medicine, National Academy of Sciences, *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America.* Washington, DC: The National Academies Press, 2012.

Kaiser Health News, Oct 2, 2012, "Medicare Revises Hospitals' Readmissions Penalties". http://www.kaiserhealthnews.org/Stories/2012/October/03/medicare-revises-hospitals-readmissions-penalties.aspx

Kassin, Michael T., Owen, Rachel M., Perez, Sebastian, Leeds, Ira, Cox, James C., Schnier, Kurt, Sadiraj, Vjollca and Sweeney, John F. 2012. "Risk Factors for 30-Day Hospital Readmission among General Surgery Patients," *Journal of the American College of Surgeons*, 215(3), 322-330.

Lee, Eva K., Fan Yuan, Daniel A. Hirsh, Michael D. Mallory, and Harold K. Simmon. 2012. "A Clinical Decision Tool for Predicting Patient Care Characteristics: Patients returning within 72 Hours in the Emergency Department," *AMIA Annual Symposium Proceedings:* 495-504.

Leeds, Ira, Sadiraj, Vjollca, Cox, James C., Schnier, Kurt E. and Sweeney, John F. 2013. "Assessing Clinical Discharge Data Preferences among Practicing Surgeons," *Journal of Surgical Research*, 184, 42-48.

McCellan, M. and J. Skinner. 2006. The Incidence of Medicare. *Journal of Public Economics* 90: 257-76.

Nabagiez, John P., Masood A. Shariff, Muhammad A. Khan, William J. Molloy, and Joseph T. McGinn. 2013. "Physician assistant home visit program to reduce hospital Readmissions." *Journal of Thoracic and Cardiovascular Surgery* 145(1): 225-233.

Newhouse, J.P. 1992. Medical Care Costs: How Much Welfare Loss? *Journal of Economic Perspectives* 6(3): 3-21.

Office of Information Products and Data Analytics, Centers for Medicare and Medicaid Services, "National Medicare Readmission Findings: Recent Data and Trends," 2012. http://www.academyhealth.org/files/2012/sunday/brennan.pdf

Organization for Economic Cooperation and Development, "OECD Health Data 2012: How Does the United States Compare". http://www.oecd.org/unitedstates/BriefingNoteUSA2012.pdf Downloaded December 14, 2012.

Pauly, M.V. 1986. Taxation, Health Insurance, and Market Failure in the Medical Economy. *Journal of Economic Literature* 24(2): 629-75.

Phelps, C.E. 1992. Diffusion of Information in Medical Care. *Journal of Economic Perspective* 6(3): 23-42.

Roth, A.E., Sönmez, T. and M.U. Ünver. 2004. "Kidney Exchange," *Quarterly Journal of Economics* 119(2), 457-488.

Roth, A.E., Sönmez, T. and M.U. Ünver. 2007. "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review* 97(3), 828- 51.

Shea, S.; R. V. Sideli; W. DuMouchel; G. Pulver; R. R. Arons and P. D. Clayton. 1995. "Computer-Generated Informational Messages Directed to Physicians: Effect on Length of Hospital Stay," *Journal of the American Medical Informatics Association* 2(1), 58-64.

Weisbrod, B.A. 1991. The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment. *Journal of Economic Literature* 29(2): 523-552.

Appendix A: Description of Data from Electronic Medical Records

The patient data contain information on the procedure conducted at the hospital and the patient's sex, age and ethnicity as well as comorbidity factors (e.g., diabetes or hypertension) and whether or not the patient in question was readmitted within 30-days following their discharge. The vital signs data contain information on objective and subjective measures of the patient's health status while in the hospital. The objective measures include the patient's body mass index (BMI), systolic blood pressure, diastolic blood pressure, heart rate, oxygen saturation level, respiration rate, and temperature. The subjective measures include the patient's functional status, fall risk score, Katz score and pain score. The laboratory test results include the patient's serum albumin, serum bilirubin, blood urea nitrogen (bun), serum creatinine, hematocrit level (hct), platelet count, prothrombin time (pt), partial thromboplastin time (ptt), serum sodium and white blood cell level.

Data on pain medications administered indicated the type of medication prescribed but did not include reliable information on the pill counts or dosage; it did include the number of times that a patient was prescribed a particular drug during their stay and the type of medication (e.g., oxycodone or morphine). The medication data also indicate whether a drug was administered to the patient orally or intravenously. Data on diagnostic images indicated the type of imaging conducted (e.g., CT scan or radiological exam or x-ray) and when it was conducted. We do not observe the physician's interpretation of the diagnostic imaging but we do observe the volume and sequence of diagnostic images conducted for each patient, which is presumably correlated with interpretation of prior imaging results. Usually, if a patient has a poor imaging result during the course of their stay a physician will order further diagnostic imaging before discharging the patient.

The patient's diet status contains information on their diet orders during the course of their stay and whether or not they are on either solid food, full liquids, clear liquids or an NPO diet.[34] The data include information on the patient's return of bowel function (e.g. stool count) for each observation time period during their stay. Lastly, in the case that blood transfusions were administered during or following a patient's surgery, we observe when the transfusions were ordered and how often they were administered.

---

[34] The designation "NPO" is an acronym for *nil per os*, which means "nothing by mouth."

Appendix B: Additional Screens from the Decision Software

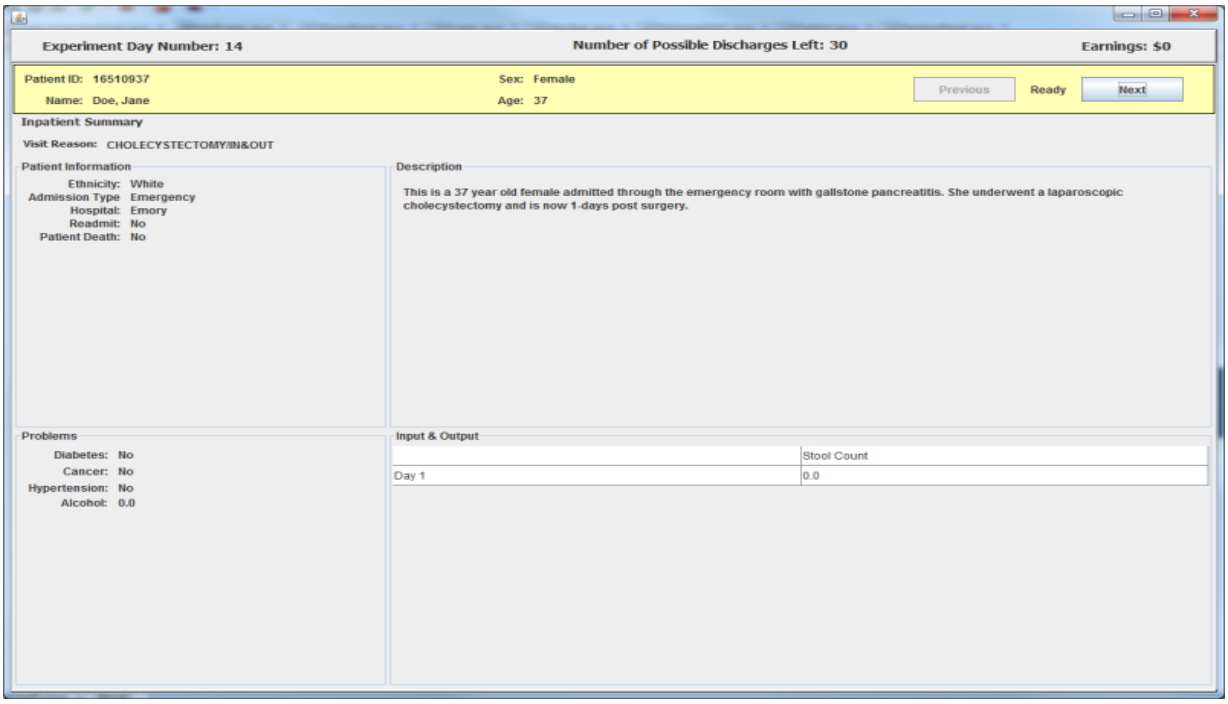

Figure A1: Inpatient Summary

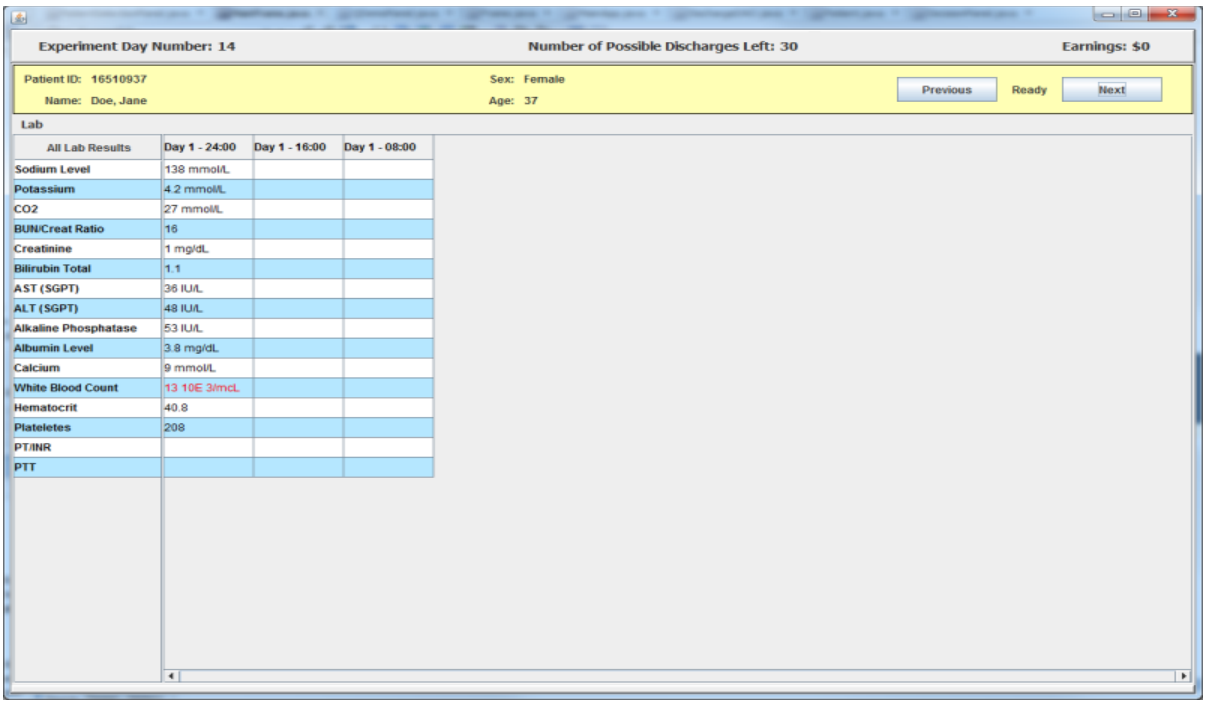

Figure A2: Laboratory Data

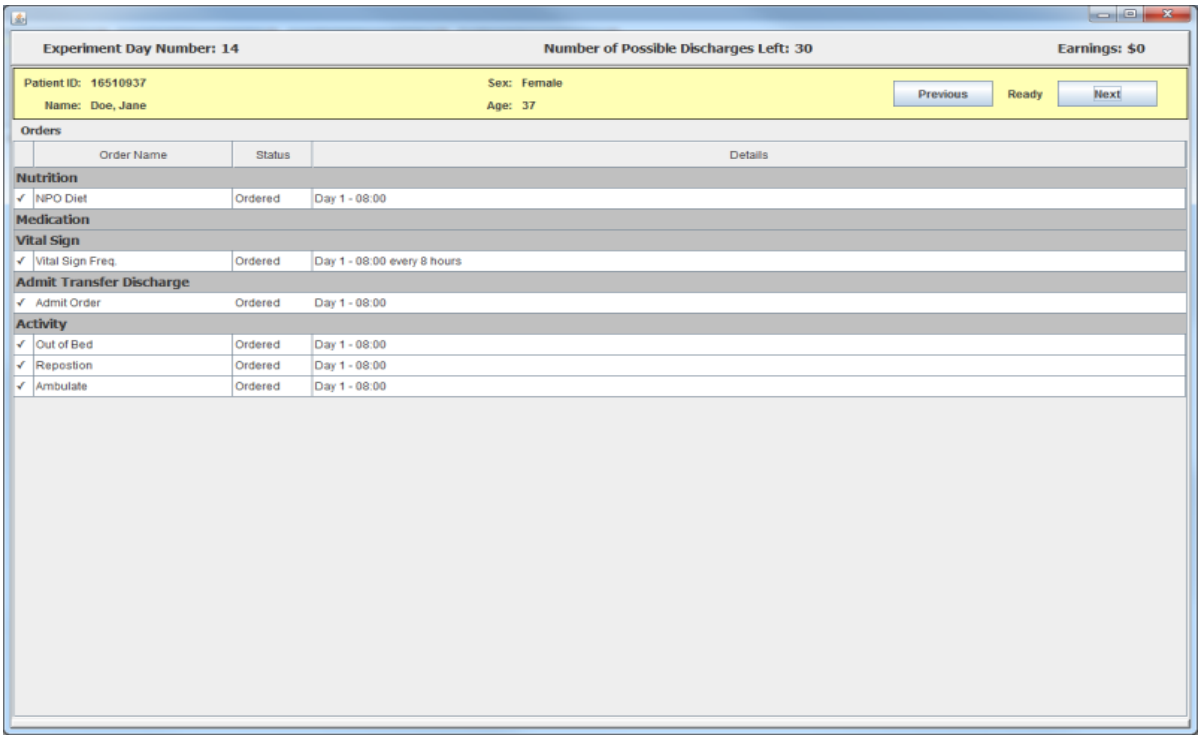

Figure A3: Orders

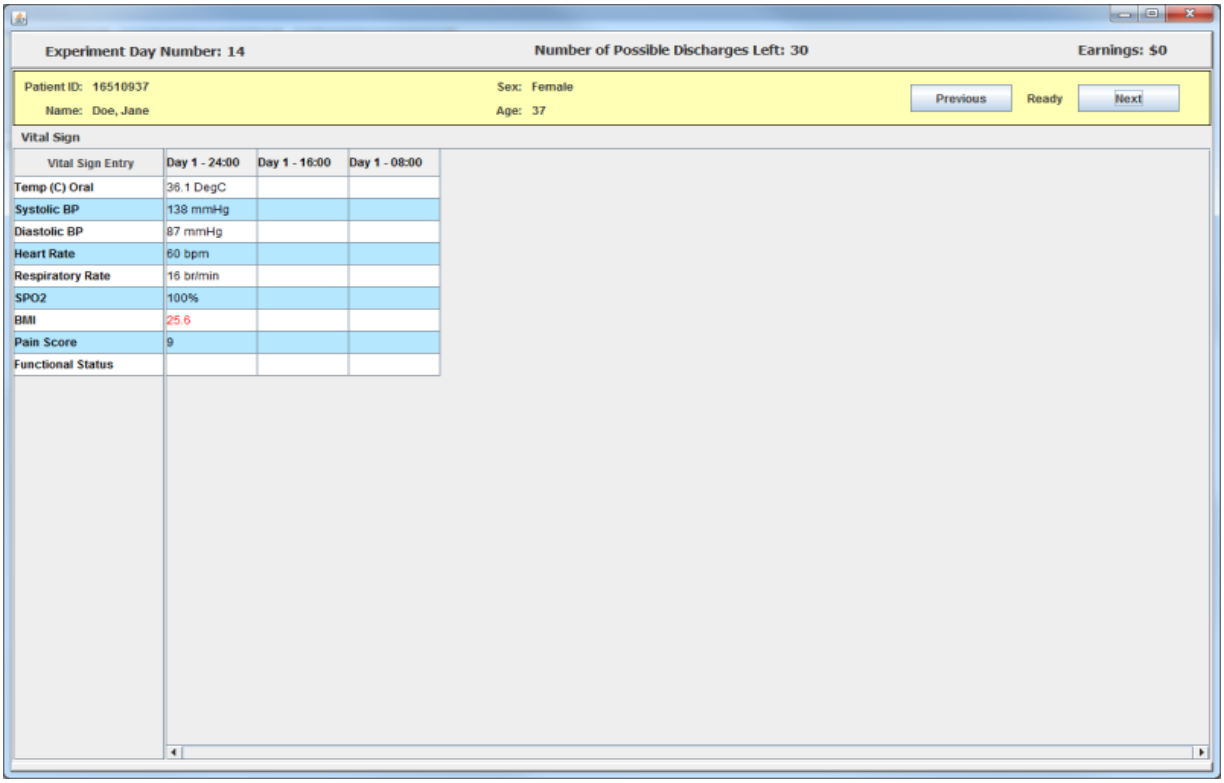

Figure A4: Vital Signs

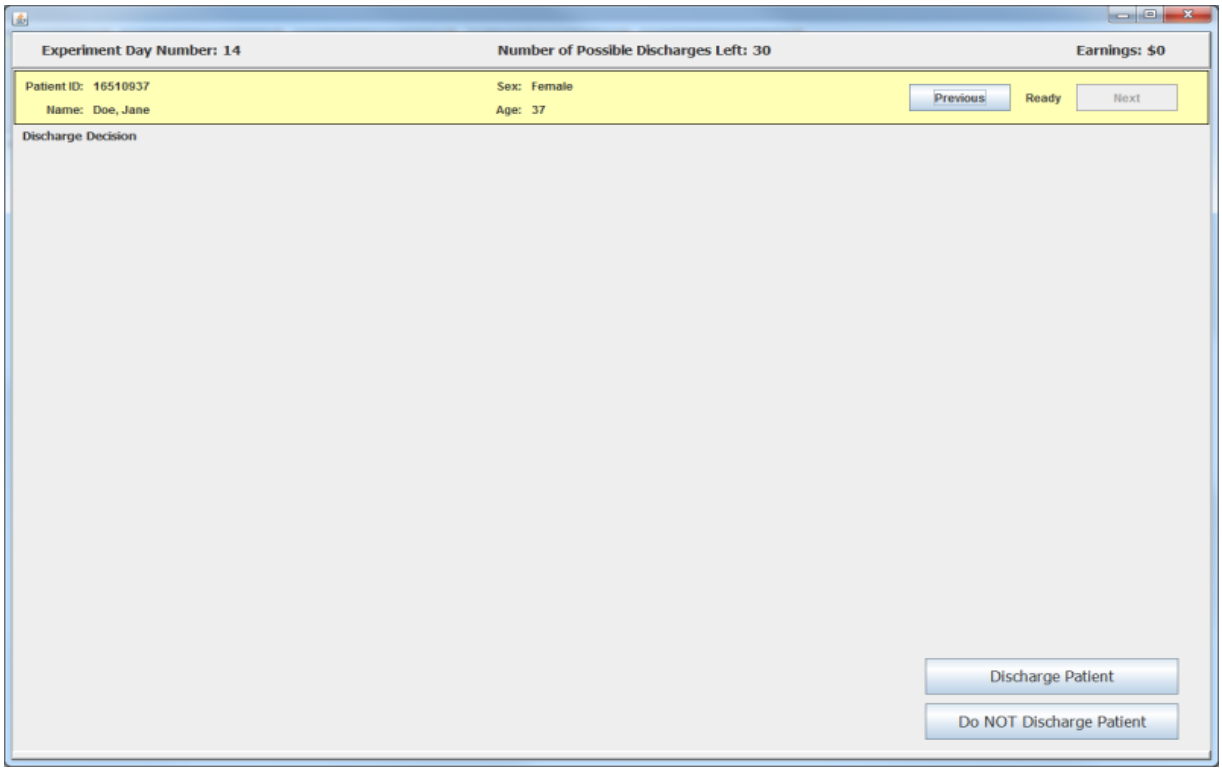

Figure A5: Baseline Treatment Decision Screen

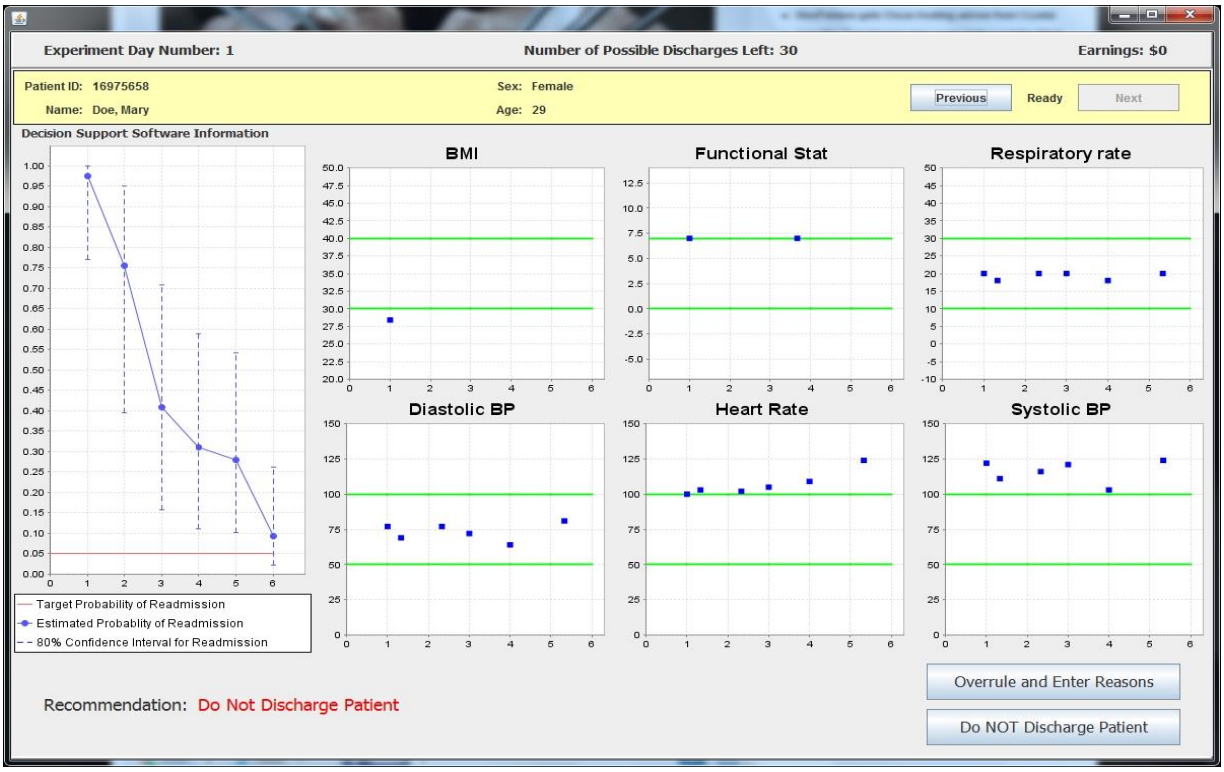

Figure A6: Default Treatment Decision Screen with Negative Recommendation

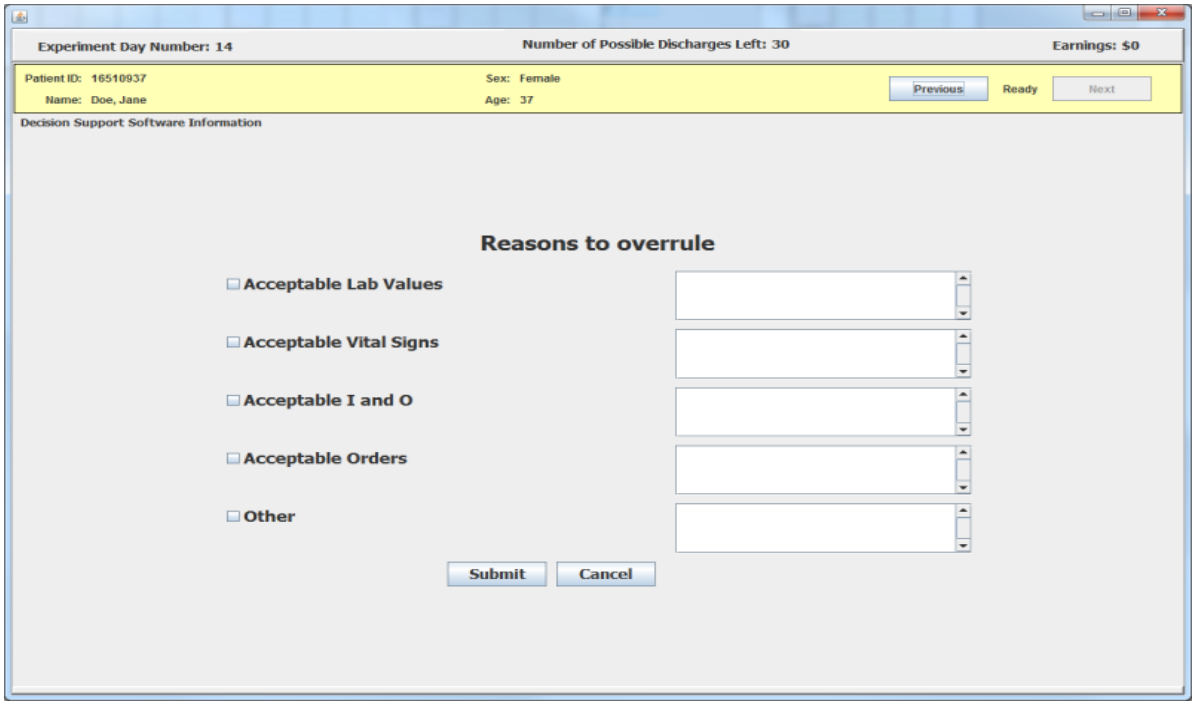

Figure A7: Default Treatment Screen: Reasons for Overruling a Negative Recommendation