Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Peter F. Lyu Date

Impact of Financial Incentives and Guideline Discussions on Decreasing Albumin Use in Critical Care

By

Peter F. Lyu
Master of Science in Public Health

Health Policy & Management

_____

Jason M. Hockenberry, PhD
Committee Chair

_____

Laura Gaydos, PhD
Committee Member

_____

David H. Howard, PhD
Committee Member

_____

David J. Murphy, MD, PhD
Committee Member

Impact of Financial Incentives and Guideline Discussions on Decreasing Albumin Use in Critical Care

By

Peter F. Lyu

Bachelor of Science in Public Health
University of North Carolina at Chapel Hill
2012

Thesis Committee Chair: Jason M. Hockenberry, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Health Policy and Management
2014

Abstract

Impact of Financial Incentives and Guideline Discussions on Decreasing Albumin Use in Critical Care

By Peter F. Lyu

Increasing concern for health care quality and rapidly growing costs has revived debates on the effectiveness of financial incentives for promoting higher value care. However, few financial incentive studies have looked at its effect on reducing overused, controversial, and expensive treatments such as albumin, for which literature generally finds no survival advantages over cheaper alternatives. Using observational data on patients treated in eight intensive care units (ICU) at two Georgia hospitals, this study examines the impact of a financial incentive program and guideline discussions on decreasing albumin use. Over 10 months, providers in eight ICUs in these two hospitals participated in a multi-faceted financial incentive program to decrease albumin use. Simultaneously, five of the eight ICUs directly participated on an Albumin Utilization Taskforce Committee to discuss system-wide guidelines. One year of baseline data was also observed. To identify independent effects for the financial incentives and guideline discussions, we employed a quasi-experimental, pre-post intervention design with non-equivalent comparison groups. Two-part regression models adjusting for health status and other covariates provided estimates of independent effects. Overall, the eight ICUs saw a significant, unadjusted 23.1% decrease in mean albumin orders per admission during the intervention year ($p<0.001$). Financial incentives were significantly associated with 0.38 fewer orders of albumin at a mean of 1.37 orders per encounter ($p<0.001$). Guideline discussions tended to decrease orders by 0.26 orders at the mean, but this overall estimate was not statistically significant ($p=0.059$). Independent effects for both interventions on volume of albumin orders were relatively larger in magnitude and statistical significance compared to effects on probability of any albumin use. Decreases among ICUs with high baseline albumin use constituted the majority of albumin decrease observed across all ICUs. Changes in albumin use were achieved without a significant change in mortality rates. These study findings contribute to the broader discussion on the role of financial incentive programs and organizational interventions in medical care and on their potential implications for costs related to albumin use.

Impact of Financial Incentives and Guideline Discussions on Decreasing Albumin Use in Critical Care

By

Peter F. Lyu

Bachelor of Science in Public health
University of North Carolina at Chapel Hill
2012

Thesis Committee Chair: Jason M. Hockenberry, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Health Policy and Management
2014

Acknowledgments

The author would like to thank the committee members, Jason Hockenberry, PhD, Laura Gaydos, PhD, David Howard, PhD, and David Murphy, MD, PhD, for their invaluable guidance in and outside of conducting this thesis. Additionally, the author thanks Tim Buchman, MD, PhD, for the opportunity to evaluate the quality improvement initiatives at Emory Healthcare; Sara Gregg, MHA, for her time and support pulling data from the Clinical Data Warehouse; Peter Joski, MSPH, for his advice during the data management process; and the entire 2014 Health Policy and Management MSPH cohort at Emory University Rollins School of Public Health for their instrumental feedback over the past year.

Table of Contents

INTRODUCTION 1

LITERATURE REVIEW 3
- Background on Human Albumin 3
- Clinical Effectiveness of Albumin 3
- Cost Implications 5
- Promoting Clinical Practice Change 6
  - Pay-for-Performance 7
  - Institutional Guidelines and Albumin Use 8

OBJECTIVES & RESEARCH QUESTION 10

METHODOLOGY 11
- Data Source 11
- Theoretical Framework 12
- Study Design 17
- Variables 18
  - Dependent Variable 18
  - Independent Variables 18
- Statistical Methods – Main Analysis 19
- Additional Analysis 21
  - Sensitivity Analysis 21
  - Subgroup Analysis 21
  - Mortality Analysis 22

RESULTS 23
- Descriptive Statistics 23
- Financial Incentive & Guideline Discussion Effects 24
  - Unadjusted Analysis 24
  - Adjusted Analysis 26
  - Sensitivity Analysis 30
  - Subgroup Analysis of High and Low Albumin Users 31
- Additional Outcomes 33

DISCUSSION 34
- Summary of Findings 34
- Impact on Provider Decision-Making 35
- Clinical & Policy Implications 38
- Study Limitations 39
  - Internal Validity 39
  - External Validity 41
- Future Directions 42

CONCLUSION 43

REFERENCES 45

APPENDIX 51

INTRODUCTION

Human albumin is a frequently used blood product in a number of health care settings in the U.S., particularly critical care. Experts continue to debate whether albumin, a protein colloid, has clinical advantages over other fluid treatments, specifically crystalloids and non-protein colloids. Albumin is also many times more expensive than these alternative treatments, so an important part of this debate surrounds resource use. The majority of recent literature reviews have shown no significant difference between albumin and lower-cost alternatives in the risk of death, but this conclusion does not necessarily address nuanced differences for specific clinical indications. As a result of this complex evidence base, widely followed standards for albumin use do not exist in the U.S.

Recognizing the substantially higher costs and apparent overuse of albumin, Emory Healthcare's Critical Care Center (ECCC), which includes intensive care units (ICU) from multiple Emory Healthcare hospitals, instituted a multi-faceted financial incentive program for critical care providers to reduce colloid use. At ECCC, albumin represented the vast majority of colloid orders. Providers became aware that albumin was to be the focus for the upcoming fiscal year as early as July 2012. Administrative preparations began at the start of the fiscal year on September 1, 2012. At this point, ICU providers and staff were officially informed of a planned financial incentive program. The incentive program spanned a 10-month period from November 1, 2012, to September 1, 2013, and was offered in three ICUs at Emory University Hospital and five ICUs at Emory University Hospital Midtown. Providers in these ICUs were offered a financial award for either reducing colloid utilization by 25% compared to their previous year's utilization or

maintaining an average colloid cost of $15 per patient day or less. Albumin orders were recorded for all patient encounters except those with liver transplant surgery, which were excluded based on existing literature supporting albumin use. The program also included an audit and feedback component where each ICU received monthly reports on their colloid use relative to other ICUs.

During the same period, Emory Healthcare also began developing institution-wide guidelines for albumin through the Albumin Utilization Task Force Committee. Committee members and advisors included representatives from various specialty and practice areas that frequently used albumin, including representatives from five of the eight ICUs that were involved in the financial incentive program. Committee activity officially began on October 19, 2012, when committee members were asked to submit recommendations for potential albumin guidelines. A series of four in-person committee meetings among providers then took place over an 11-month period between November 19, 2012, and October 7, 2013, to discuss institutional guidelines for albumin use at Emory. The committee ultimately finalized these guidelines during a final fifth meeting on October 24, 2013.

The ECCC is an academic, closed model combined with affiliate providers and includes multiple ICU types (see *Appendix*). With two interventions taking place at the same time in the ECCC, eight ICUs were in a unique setting where all critical care providers were exposed to the incentive program, and five of those eight ICUs had physicians that were also directly involved in the guideline setting process occurring simultaneously. With on-

going debates on how to best promote higher quality, better value care through provider practice change, this study seeks to understand the impact of these financial and organizational mechanisms on changing practice patterns.

LITERATURE REVIEW

Background on Human Albumin

Albumin's multifunctional characteristics have allowed physicians to administer albumin in clinical settings for a wide variety of medical indications, particularly in fluid therapy. Clinical applications first emerged after early studies revealed that human serum albumin had protective advantages and was inversely related to mortality [1]. Since its first widespread use treating burn victims in World War II, albumin has become a frequent treatment option in the U.S. [2]. However, despite frequent use and the existence of various recommendations and guidelines, there are currently no widely followed standards in the U.S. for fluid therapy [2, 3]. Albumin's clinical effectiveness compared to alternate treatment strategies continues to be debated. An extensive, yet inconsistent, literature on the effectiveness of albumin includes both observational studies and randomized controlled trials.

Clinical Effectiveness of Albumin

Studies evaluating albumin effectiveness typically compare albumin against no treatment or treatment with crystalloids or non-protein colloids. Crystalloids are solutions composed of water-soluble compounds and salts. Non-protein colloids include non-water-soluble, non-protein colloidal components, many of which are semisynthetic [4].

These non-protein colloids include hydroxyethyl starch (HES or hetastarch), gelatin, and dextran solutions [5]. Both crystalloids and non-protein colloids have characteristics that can produce similar physiological responses to those produced by albumin, allowing them to function as alternative treatments.

Clinical evaluations of albumin and these alternatives have reported mixed results over the past 20 years, with the most recent studies finding either no difference or higher rates of mortality with albumin use. A meta-analysis by the Cochrane Injuries Group in 1998 first sparked a number of subsequent studies in response when a review of 30 randomized controlled trials actually concluded a greater overall risk of death among patients treated with albumin compared to no albumin [6]. Many researchers later criticized this review for methodological flaws, specifically the exclusion of important, relevant studies [7, 8]. A subsequent meta-analysis by Wilkes and Navickis that included a greater number of randomized controlled trials challenged the Cochrane results, finding no evidence that albumin affects overall mortality, positively or negatively [9]. Three years later, results from the seminal Saline versus Albumin Fluid Evaluation (SAFE) Study further supported this conclusion, finding no evidence that mortality rates differed between patients treated with albumin versus saline solution [10]. Furthermore, Cochrane's most recent meta-analysis in 2013 included 70 randomized controlled trials and found that albumin compared to crystalloids did not predict differential rates of overall mortality when used for fluid resuscitation [11]. That review, however, did find a slightly higher risk of death with the use of HES, a non-protein colloid alternative.

Although more recent literature seems to indicate an equivalent overall survival effect from albumin versus crystalloids, this overall finding does not necessarily translate to conclusions for specific cases for albumin use. Given the wide array of clinical indications for albumin use, individual studies included in systematic literature reviews each tend to focus on one or a small number of these indications. While these larger meta-analyses report subgroup findings for some indications, other individual studies report results for subpopulations of patients that differ from the overall albumin finding. For example, the Sepsis Occurrence in Acutely ill Patients (SOAP) study previously found higher mortality rates associated with albumin use in patients with trauma or sepsis [12]. In contrast, other past studies suggested that albumin might be associated with lower mortality for patients with severe sepsis [10, 13], and these findings even influenced albumin use policy in the UK [14]. However, the most recent randomized controlled trial of severe sepsis patients demonstrated no difference in mortality between albumin and crystalloid groups [15]. These nuanced and evolving findings continue to drive a debate on the clinical effectiveness of albumin. As a result, albumin use in the U.S. has not been standardized.

Cost Implications

In addition to clinical outcome considerations, albumin use has significant cost implications for health care systems and hospitals. Relative to alternative treatments, albumin has significantly higher costs per order, largely due to a smaller supply of these products available to hospitals [16]. By volume, albumin can be two to three times more expensive than non-protein colloids, such as HES and dextran [3, 17]. When compared to

crystalloids, the cost differential is even larger: albumin can be anywhere from 20 to 100 times more expensive [3, 17, 18]. With high costs combined with frequent use and a lack of widely accepted guidelines, clinical decisions on albumin versus alternate treatments have become a targeted space for cost reduction in health systems.

At Emory Healthcare, albumin costs (i.e. wholesaler acquisition price) approximately $170 per liter, while other non-protein colloid and crystalloid average costs are closer to $50 and $1 per liter, respectively. The volume of fluid typically used varies widely depending on the specific patient, but past research has suggested that mean cumulative volume of treatment fluid per patient is only about 1000 mL greater with crystalloids compared to colloids [19].

Promoting Clinical Practice Change

Although widely followed standards for albumin use do not exist on a national level in the U.S., health care institutions such as Emory Healthcare have attempted to reduce use of albumin at a local level. However, changing clinical practice behavior has long been considered a challenging process [20]. Guideline setting and financial incentives are two strategies currently used to implement practice changes [21, 22]. While a wealth of literature on the effectiveness of these approaches exists, few studies have looked at the impact of these strategies on changing clinical practice for a controversial treatment such as albumin.

Pay-for-Performance

Rising health care costs over the past two decades have spurred a rapidly growing body of studies focusing on the financial value of quality care. Pay-for-performance (P4P) methods, which tie a portion of financial remuneration to performance measures, have become a widely discussed strategy for managing health care costs and promoting quality [23]. However, the most comprehensive systematic literature review of P4P studies to date revealed that most of these research efforts have focused on P4P's effect on promoting underused services, specifically preventive services and chronic care management [24, 25]. In contrast, literature evaluating the impact of P4P on reducing overused services, such as albumin, is scarce despite half of P4P measures being targeted at cost efficiency [26].

The rapidly growing area of research debating the effectiveness of financial incentives remains focused in outpatient settings [24, 27-30]. Generalizations of a P4P effect to inpatient settings should be cautious, especially in the context of debated measures that lack a fully accepted evidence base such as albumin. Rosenthal and Dudley claim that, "the basic intent of pay-for-performance [is] to encourage and assist providers in offering the most clinically appropriate care" [26]. However, expected provider response to financial incentives is less clear when widespread standards for "clinically appropriate care" are not fully developed or recognized by practitioners.

In their theoretical discussion of the influence of financial incentives, Town et al. propose that, "incentives that are compatible with…professional values may be more influential

than incentives that conflict with professional values" [31]. Furthermore, qualitative findings from a study in California saw more than 60% of physician organizations respond to P4P implementation by reviewing clinical guidelines [32]. Thus, even in the presence of clear financial incentives, physician decision-making seems to depend at least in part on existing guidelines in the medical literature. Absent widely accepted guidelines, the impact of financial incentives is to this date unexplored.

Institutional Guidelines and Albumin Use

An evidence base does not necessarily translate to widely accepted standards of care, particularly as shown by the albumin case. A small number of observational studies have looked at the impact of albumin guidelines, set locally within institutions or nationally in other countries. These studies revealed an inconsistent impact on albumin use after guideline implementation, and the majority had important study design limitations that limit generalizability.

Several studies outside the U.S. reported between 20-70% reduction in albumin use following implementation of nationally recommended guidelines at the institutional level [33-35]. However, the implications for U.S. providers are unclear, as providers in other countries practice in widely different health care systems. For example, Durand-Zaleski et al. and Debrix et al. observe changes in French public hospitals, which operate on fixed budgets [33, 34]. Due to strong cost restrictions inherent to fixed budget hospitals, guidelines may have had a stronger impact in those environments. Additionally, a descriptive study by Fan et al. in the British Columbia found that, without a national set

of guidelines, hospitals continued to use albumin for indications that were unsupported in the literature [36].

In the U.S., a number of institutional albumin guidelines have been developed over the past 20 years in local settings. Unfortunately, studies that observed the impact of these guidelines were severely limited by their study design, particularly a very short time period of post intervention data (e.g. 1-4 months) or a very small scope of patient populations [2, 3, 17, 18]. Many of the guidelines implemented in these studies were based on or adapted from guidelines developed by the University Hospital Consortium (UHC) [37]. Provider response to UHC guidelines was mixed, often depending on the implementation strategies used at the local level. In a context without organized guideline implementation, Tanzi et al. and Yim et al. found that albumin continued to be used inappropriately in over 50% of patients shortly after publication of UHC guidelines in 1993 and 2000 [2, 3]. In contrast, Charles et al. saw a 54% reduction in overall albumin use when an organized education program was implemented to promote UHC-based guidelines [17]. However, this study did not evaluate appropriateness of use at the patient level and only included patients treated in a single surgical intensive care unit.

According to Grimshaw and Russell, success of guidelines depends on three components: the method in which guidelines were developed, the medium through which guidelines were disseminated to practitioners, and the implementation strategies for encouraging practitioners to follow guidelines [38]. In their literature review, they found that 55 out of 59 studies evaluating guidelines in a number of different clinical areas saw significant

changes in medical processes. These studies looked at both externally and internally developed guidelines but different implementation strategies. Based on their findings, Grimshaw and Russell suggest that externally developed guidelines require more resources at the local level for dissemination and implementation. Conversely, internally developed guidelines must focus resources during the developmental stage to ensure rigorous evaluation of the guidelines.

Grimshaw and Russell also add that guidelines themselves may not affect rapid change without the presence of other incentives. Interestingly, none of their review's guideline studies included financial incentives as a component of the implementation strategy. Amidst a growing nationwide interest in P4P, financial incentives could potentially provide that additional push to overcome the "clinical inertia" needed for practice change [30].

While the financial incentive program and guideline discussions at Emory Healthcare were not officially coordinated, their overlapping implementation offered a unique opportunity to observe how those interventions affected physician decision-making around albumin use.

OBJECTIVES & RESEARCH QUESTION

The objectives of this study are to:

1. Determine whether the financial incentive program at Emory Healthcare ICUs was independently associated with decreased albumin use,

2. Determine whether guideline discussions independently predicted decreased albumin use, and
3. Explore whether there was a combined effect from financial incentives and guideline discussions together.

Given these objectives, this study's primary research question asks: do financial incentive programs and guideline discussions predict lower utilization of albumin in critical care units?

METHODOLOGY

Data Source

Data used in this study was extracted from electronic medical records for all patients admitted to eight ICUs in two Emory Healthcare hospitals on or after September 1, 2011, and discharged before September 1, 2013. The unit of analysis was at the encounter level. This study only observed health and administrative information during patients' first ICU stay as previous research has suggested that patients' in subsequent ICU stays are significantly different from patients in their first ICU stay [39]. Additionally, since liver transplant patients were excluded from Emory Healthcare's albumin financial incentive program, encounters with ICD-9 codes indicating liver transplant were dropped from the study sample. The final sample for the main analysis included 16,117 unique encounters.

Theoretical Framework

This study's design and methods are based on a theoretical framework developed specifically for the context surrounding the albumin-related initiatives at Emory Healthcare and is illustrated in Figure 1. At the highest level, albumin use is driven by a combination of provider and patient factors. The provider factors were the primary focus of this study because the financial incentive program and guideline discussions were targeted at providers and their decisions to order albumin. Furthermore, the nature by which institutional guidelines and financial incentives each affects these provider decisions can also be thought to contain their own theoretical mechanisms. Thus, this study's theoretical framework adapts concepts from two existing frameworks: Grimshaw and Russel's framework for understanding consensus guideline development and Magnus' financial incentive framework [38, 40].

Grimshaw and Russell provide a useful framework for understanding the conceptual steps in the consensus guideline process (development, dissemination, and implementation) and how strategic investment at certain steps can increase impact from the final guidelines [38]. This study's interest in the impact of internal guideline discussions looks specifically at the *development* component. As Grimshaw and Russell note, when more resources and local provider buy-in are obtained while developing internal guidelines, fewer resources are needed for dissemination and implementation. Based on this framework, guideline discussions themselves could theoretically have an independent effect as providers involved on the guideline committee bring back an albumin-conscious mentality and preliminary albumin guidelines to their own ICU.

Likewise, financial incentives can be understood within an early framework developed by Stephen A. Magnus, originally applied in the context of health maintenance organizations (HMO) [40]. Briefly, structural characteristics of financial incentives can be described within five dimensions: (1) percentage of income offered or at stake, (2) level of organization in which the incentive is offered, (3) synergy between multiple financial incentives, (4) synergy between financial and nonfinancial incentives, and (5) incorporation of signaling and psychological effects. Characterizing incentives through this framework allows for a systematic means of understanding intended effects and for comparing different incentives. Thus, these dimensions were used to characterize the financial incentive offered to critical care providers at Emory Healthcare:

Dimension 1: (Incentive Size) The payout to physicians was $750 per week for meeting the reduction threshold and no payout if the goal was not met (i.e. all-or-nothing). The incentive was paid out in a lump sum at the end of the incentive period. In total, the maximum possible award amounted to about a 16% bonus for physicians (assuming a $240K salary). For non-physician providers such as nurse practitioners (NP) and physician assistants (PA), the maximum possible payout was $2,500 per full-time employee (FTE) per six months, offered twice during the incentive period and approximately amounting to a year-end 5% bonus (assuming 1 FTE and a $100K salary).

Dimension 2: (Target Level) Incentives were offered based on each ICU's total colloid utilization as a whole, not on individual provider utilization of colloids. In other words, the incentive was structured as a group-based, all-or-nothing arrangement. However, incentives were awarded only to prescribing providers, which included ICU physicians and affiliates.

Dimension 3: (Other Financial Incentives) Financial incentives for decreasing albumin use may have also synergistically interacted with other financial initiatives that occurred around the same time period.

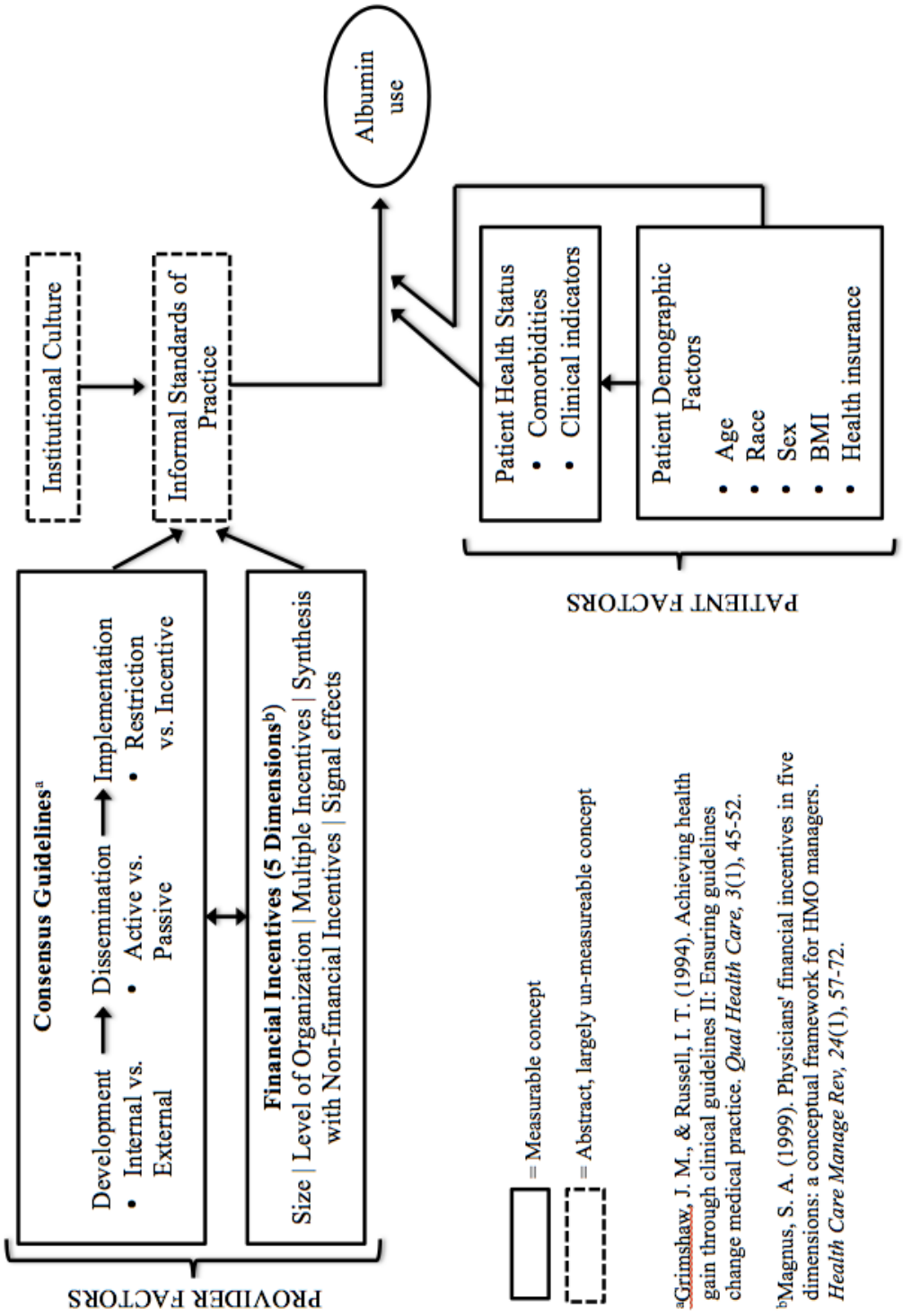
Dimension 4: (Non-financial Incentives) Guideline discussions through the Albumin Utilization Taskforce Committee encourage change based on empirical evidence rather than an explicit financial incentive. Thus, committee discussions may have increased the impact of financial incentives by offering clinical and cost reasons to decrease albumin orders. Additionally, audit and feedback mechanisms involved in the financial incentive program may also impact albumin orders by making physicians aware of their current performance.

Dimension 5: (Specific Signals) By offering a "floor" as one of the performance benchmarks (≤$15 average colloid cost per patient day), the incentive tells providers that colloids use is still expected in some cases. In other words, the incentive targets frequent utilizers of albumin, but also conveys a message that albumin should not be blindly decreased for all patients.

Dimensions 1, 2, and 5 can be especially useful for comparing the impact of this incentive program with incentives in other programs or institutions. For understanding effects within this study, the fourth dimension is most relevant and ties to the development step of the consensus guideline process. Influencing each other, consensus discussions and financial incentives can theoretically influence and change informal standards of practice to ultimately decrease albumin use. Larger institutional culture factors may also play a role and include influences from organizational structure and established protocols. For example, providers in an institution where quality improvement initiatives are widely embraced are more likely to respond to guideline discussions and financial incentives.

On the patient side, health status and demographic factors are the main theoretical moderators of albumin use. For example, those in poorer health and with more complex conditions may be more likely to require fluid therapy and thus be treated with albumin. Other demographic factors, such as race and health insurance, are also hypothesized to have independent effects on albumin use as minority groups, Medicaid, and the uninsured have reported disparities in quality of care, even in critical care settings [41, 42].

Figure 1. Study theoretical framework



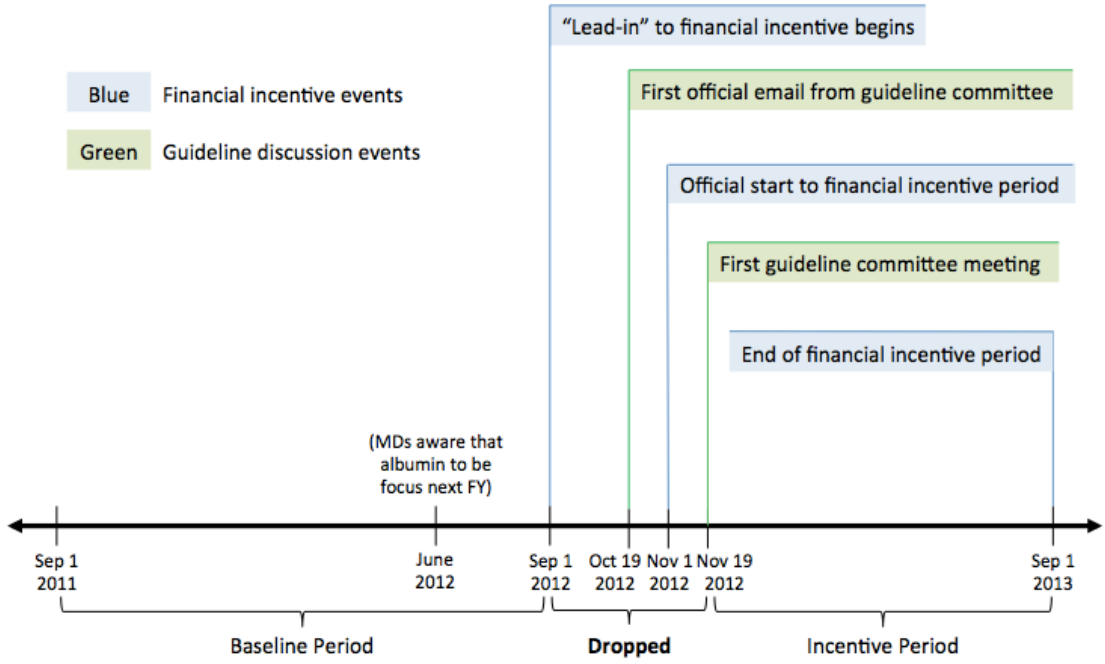
Study Design

The study design used was a quasi-experimental pre-post intervention with nonequivalent comparison groups, using a pre-post comparison to identify independent financial incentive program effects and differences-in-differences analysis to identify independent guideline discussion effects. The two comparison groups were a Guideline Discussion group of five ICUs involved in guideline discussions and a Non-Guideline Discussion group of the remaining three ICUs. This study observes changes in effects before and after the start of the financial incentive period that affected all ICUs. Thus, the interventions affecting each group can be summarized as follows:

- Guideline Discussion Group: guideline discussions + financial incentive program
- Non-Guideline Discussion Group: financial incentive program only

As Figure 2 illustrates, the timeline of early intervention activities are staggered, offering a number of potential dates to mark the beginning of the post-period. Rather than subjectively select one of these dates, the main analysis excluded all admissions from September 1, 2012 to November 19, 2012. As a result, the baseline period (i.e. pre-period) began September 1, 2011, and ended September 1, 2012. The incentive period (i.e. post-period) began November 19, 2012, and ended September 1, 2013, representing the period when the financial incentive program affected all eight ICUs.

Figure 2. Timeline of financial incentive program and guideline discussion events and study time frame



Variables

Dependent Variable

The dependent variable used was number of albumin orders per encounter during the first ICU stay.

Independent Variables

The two primary independent variables captured (1) whether the patient was treated in an ICU in the Guideline Discussion Group or Non-Guideline Discussion Group and (2) whether patients were admitted during the baseline or incentive period.

Other independent variables included measurable patient demographic and health status factors identified in the theoretical framework that likely also affected providers' decision to use albumin. Demographic variables included age, gender, race, and type of insurance.

Health status variables included body mass index (BMI) at admission, a Charlson Comorbidity Index (CCI) score calculated from ICD-9 diagnosis codes [43], and a Sequential Organ Failure Assessment (SOFA) score at ICU admission calculated from procedure orders and lab test results (see *Appendix* for details) [44], and ICU length of stay.

Statistical Methods – Main Analysis

Given the pre-post, non-equivalent comparison group design, the first econometric model is based on a simple pre-post analysis, and the second model uses a difference-in-differences analysis method.

1. $Y_{Per\ enc\ alb\ use} = \beta_0 + \beta_1 FinInc + \beta_2 GuidDisc + \beta_3 CCI\ Score + \beta_4 SOFA\ Score + \beta_5 Age + \beta_6 White + \beta_7 Male + \beta_8 ICU\ LOS + \beta_9 BMI + \beta_k [InsuranceType]$

   Where Y = number of albumin orders per encounter

   ß$_1$ = independent effect from having an ICU stay during the incentive period

   ß$_2$ = independent association of albumin use linked to the guideline discussions across both baseline and incentive periods

2. $Y_{Per\ enc\ alb\ use} = \beta_0 + \beta_1 (FinInc * GuidDisc) + \beta_2 FinInc + \beta_3 GuidDisc + \beta_4 CCI\ Score + \beta_5 SOFA\ Score + \beta_6 Age + \beta_7 White + \beta_8 Male + \beta_9 ICU\ LOS + \beta_{10} BMI + \beta_k [InsuranceType]$

Where Y = number of albumin orders per encounter

$ß_1$ = independent guideline discussion effects

$ß_2$ = independent association of albumin use with financial incentives among those not in the guideline discussion group

$ß_3$ = independent association of albumin use with guideline discussions among those in the baseline period

Model 1 identifies the independent association of the financial incentive program with albumin use while controlling for any guideline discussion effects and other confounders. Model 2 is similar to Model 1 but instead uses the interaction term "*FinInc*GuidDisc*" to capture the differences in albumin use before and after the financial incentive program between the study groups, subsequently identifying a guideline discussion effect independent to that of the financial incentive (i.e. the difference-in-differences). Both models include insurance status as a series of dummy variables with private insurance as the reference group.

Distribution of the dependent variable, number of albumin orders per encounter, revealed a preponderance of encounters with zero use of albumin and a skewed right distribution for those encounters with any utilization. As a result of this distribution, a two-part model was used to better fit the data and provide two different levels of inference [45]. The first part of the model used a logit regression to capture effects on probability of any albumin use during an encounter. The second part of the model used a negative binomial regression with a log link, capturing effects on volume of albumin use *given any use*. A

negative binomial distribution was selected over the Poisson distribution based on results from the standard over-dispersion test, which indicated that the outcome data were over-dispersed [46]. Finally, overall calculations combined estimates from both parts to predict overall effects on number of albumin orders. Marginal effects were calculated for all models to produce more easily interpretable effect estimates.

Additional Analysis

Sensitivity Analysis

Statistical models used in the main analysis were also applied in three separate cases where the dropped admissions between September 1, 2012, and November 19, 2012, were included and the start date of the incentive time-period was altered. Each case calculated the independent financial incentive and guideline discussion effects and adjusted for the same factors specified in the main analysis. In the first alternative analysis, the post-period began on July 1, 2012, marking the approximate date when critical care providers became aware of an albumin focus for the next fiscal year. In the second alternative analysis, the post-period started on September 1, 2012, the beginning of the fiscal year when rollout of the financial incentive program began. The final analysis set the start of the post-period to November 1, 2012, the official start date of the incentive program.

Subgroup Analysis

This study also sought to determine whether the effect of the financial incentive program and guideline discussions differed between ICUs with higher albumin baseline use and

those ICUs with lower baseline use. At Emory Healthcare, the three ICUs with high baseline albumin use were all involved in guideline discussions. Thus, the subgroup analysis observed only the Guideline Discussion Group ICUs and compared unadjusted utilization between the subgroup with significantly higher baseline utilization and a subgroup of the remaining ICUs with lower baseline utilization.

Mortality and Cost Analyses

Mean mortality rates during first ICU, any ICU, and overall hospital stays were also observed. Differences in mean mortality between the baseline and incentive periods with both study groups combined were compared using two-tailed t-tests. Mean direct albumin costs per encounter were estimated by multiplying the total number of orders by the estimated unit cost of $170 per order and then dividing that total by the total number of encounters for the given period. Mean direct albumin costs per albumin encounter were also calculated using only encounters with any albumin use as the denominator. Additionally, total hospital costs for patients' entire admission beyond first ICU stay were also observed. Two-tailed t-tests were used to compare these three cost measures between the baseline and incentive periods.

Data cleaning and statistical analysis were conducted using Stata 12 (StataCorp, College Station, TX). Estimates with a probability of less than 0.05 were considered significant. This study was approved by the ECCC Institutional Review Board (IRB).

RESULTS

Descriptive Statistics

Table 1 describes the study population's patient mix within the Guideline Discussion and Non-Guideline Discussion study groups over the two-year study timeframe by comparing demographic and health risk factors between the baseline (pre) and incentive (post) time-periods. Overall, the baseline and incentive periods included 9,094 and 7,023 encounters, respectively.

For both study groups, most factors did not differ significantly between the two periods when compared using two-tailed t-tests and two-tailed proportions tests. However, mean Sequential Organ Failure Assessment (SOFA) scores, a measure of patient severity, and health insurance source or status did show statistically significant differences with a two-tailed t-test and chi-squared test, respectively. In the Guideline Discussion Group, post-period patients had a 0.3 higher mean SOFA score compared to patients admitted during the pre-period, suggesting that the post-period had a sicker mix of patients ($p<0.001$). The Guideline Discussion Group also saw a greater proportion of uninsured patients and a lesser proportion of patients covered by public, private or other sources in the post-period ($p=0.001$).

Table 1. Descriptive characteristics of patient encounters in Guideline Discussion ICUs and non-Guideline Discussion ICUs before and during the financial incentive period

	Guideline Discussion Group		Non-Guideline Discussion Group	
Characteristic	Baseline (n=5,554)	Incentive (n=4,354)	Baseline (n=3,540)	Incentive (n=2,669)
Age, mean (se)	60.4 (0.2)	60.7 (0.2)	57.3 (0.3)	57.1 (0.3)
Male (vs. Female), % (se)	55.5 (0.7)	55.1 (0.8)	46.8 (0.8)	46.0 (1.0)
White (vs. non-White), % (se)	49.6 (0.7)	49.7 (0.8)	52.6 (0.9)	52.5 (1.0)
Health insurance, %				
Private	25.2*	25.0*	33.2	32.8
Public	68.2*	67.2*	58.7	57.7
Uninsured	5.6*	7.1*	6.4	8.1
Other	1.1*	0.6*	1.8	1.4
BMI at admission, mean (se)	29.3 (0.2)	28.9 (0.1)	28.9 (0.2)	28.5 (0.2)
Charlson Comorbidity Index score, mean (se)	3.3 (0.0)	3.3 (0.0)	2.7 (0.0)	2.7 (0.0)
Sequential Organ Failure Assessment score, mean (se)	5.6 (0.0)*	5.9 (0.1)*	4.5 (0.1)	4.5 (0.1)
First ICU length of stay (days), mean (se)	3.0 (0.1)	2.9 (0.1)	3.8 (0.1)	3.7 (0.1)
Hospital length of stay (days), mean (se)	10.4 (0.1)	10.2 (0.2)	10.1 (0.2)	9.6 (0.2)

*$p<0.05$ with two-tailed t-test, two-tailed proportions test, or chi-squared test

Table 1 also reveals notable differences between the two study groups for a number of characteristics. Patients treated in Guideline Discussion ICUs tended to be older, more likely to be male and non-White, publicly-insured, have shorter first ICU lengths of stay, and higher Charlson Comorbidity Index (CCI) and SOFA scores. There were also a substantial number of missing values for race and BMI, but these missing values were evenly distributed between the two study groups and between the baseline and incentive periods.

Financial Incentive & Guideline Discussion Effects

Unadjusted Analysis

Table 2 summarizes the unadjusted changes in albumin use from the baseline to incentive period. All ICUs together (i.e. combining both study groups) reported a significant

average decrease of 0.6 albumin orders per patient encounter (p<0.001). When divided into study groups, Guideline Discussion ICUs saw a statistically significant 0.9 decrease in mean, unadjusted albumin orders per encounter in the post-period, a relative 24.3% decrease (p<0.001). Non-Guideline Discussion ICUs reported a smaller 0.1 order decrease, but this mean decrease was not statistically significant (p=0.088).

Table 2. Unadjusted difference in albumin use overall and for Guideline Discussion ICUs versus Non-Guideline Discussion ICUs before and during the financial incentive period.

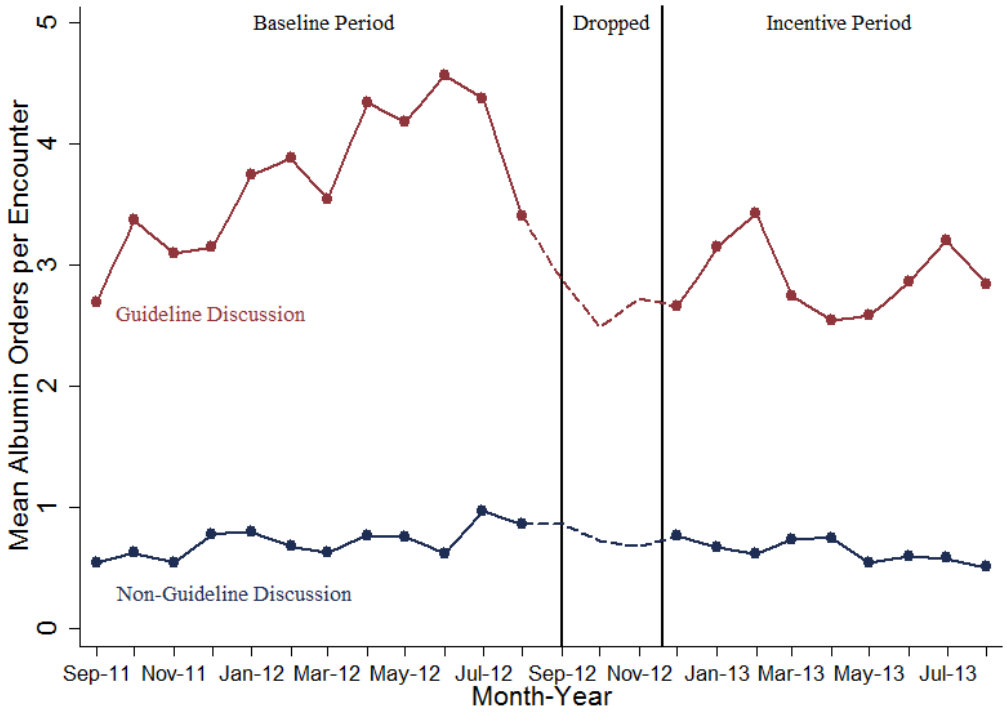
Albumin orders per encounter	Baseline Period Albumin Use	Incentive Period Albumin Use	Difference in Albumin Use
All ICUs (i.e. both study groups), mean (se)	2.6 (0.1)	2.0 (0.1)	-0.6 (0.1)*
Non-Guideline Discussion Group, mean (se)	0.7 (0.0)	0.6 (0.1)	-0.1 (0.1)
Guideline Discussion Group, mean (se)	3.7 (0.1)	2.9 (0.1)	-0.9 (0.1)*
Difference-in-differences			**-0.8 (0.2)***

*p<0.05 with two-tailed t-test

The unadjusted change attributable to guideline discussions is represented in Table 2 as the difference-in-differences, which controls for the financial incentive effect by differencing the change for the two study groups over the baseline and incentive periods (i.e. (-0.9) – (-0.1)). In other words, an unadjusted mean decrease of 0.8 orders per encounter was associated with guideline discussions (p<0.001).

This unadjusted difference in the magnitude of change between the Guideline Discussion and Non-Guideline Discussion groups is also illustrated in Figure 3, which displays monthly average albumin use for the two study groups across the two years. The middle lines from September 1, 2012, to November 19, 2012, are dashed because these admissions were dropped for the main analysis.

Figure 3. Monthly albumin use per encounter between Guideline Discussion and Non-Guideline Discussion ICUs



Adjusted Analysis

Tables 3 and 4 summarize the full, adjusted results estimated with two separate two-part models (TPM). The first TPM identified the financial incentive-specific effect by calculating the pre-post difference while controlling for guideline discussions as one of the covariates (Table 3). The second TPM identified the guideline discussion-specific effect by using a difference-in-differences interaction term (financial incentive variable*guideline discussion group variable) in addition to other covariates (Table 4). All estimates represent marginal effects at the predicted means, which were largely the same predicted values between the two TPMs: the predicted mean probability of using any albumin was 0.28, the predicted mean volume of use given any use was 6.6 orders, and the overall predicted mean use was 1.4 orders.

Part 1 logit regression estimates of the first TPM found that the financial incentive program was independently associated with a 1.9 percentage-point lower probability of using any albumin, or a 6.9% decrease relative to the mean ($p=0.027$). The Part 2 negative binomial regression subsequently found that the incentive program was associated with using 1.36 fewer orders per encounter among patients receiving at least one order of albumin, or a 20.8% decrease relative to the mean ($p<0.001$). Overall, the first TPM estimated that the financial incentive program was independently associated with a decrease of 0.38 orders of albumin during the first ICU stay, or a relative 27.7% decrease ($p<0.001$).

In contrast, the second TPM estimated that guideline discussions tended to decrease overall albumin use by 0.26 orders, but this effect was not statistically significant ($p=0.059$). Guideline discussions also did not seem to impact the probability of whether or not a patient received any albumin, as estimated by the logit model ($p=0.844$). However, the negative binomial regression did find that guideline discussions predicted a significant decrease of 1.19 fewer orders among those patients who received any albumin, or an 18.1% relative decrease from the mean ($p=0.014$).

Several covariates also had statistically significant effects on albumin use that are worth noting. Patients with higher SOFA scores were associated with greater albumin use, specifically 0.35 more orders of albumin for each single point score increase at the mean ($p<0.001$). Demographic characteristics also had significant associations. White patients were more likely to receive albumin, receiving an estimated 0.79 orders of albumin more

per encounter than non-White patients (p<0.001). Publically insured and uninsured patients were less likely to receive albumin, using approximately 0.18 and 0.38 fewer orders per encounter, respectively, than privately insured patients (p=0.01; p<0.001).

Table 3. Adjusted two-part model marginal effects of the financial incentive on albumin use per encounter during the first ICU stay.

	Part 1: LOGIT	*p*-value	Part 2: GLM (NB)	*p*-value	Overall Effects	*p*-value
Independent financial incentive effect	-0.019[a]	0.027	-1.363[c]	<0.001	-0.378[c]	<0.001
Encounter in a guideline discussion ICU (vs. non-guideline discussion)	0.286[c]	<0.001	2.433[c]	<0.001	1.898[c]	<0.001
Patient age	-0.001[c]	<0.001	0.009	0.137	-0.004	0.064
Patient sex is male (vs. female)	0.017	0.056	0.204	0.249	0.127[a]	0.027
Patient race is White (vs. non-White)	0.123[c]	<0.001	0.840[c]	<0.001	0.790[c]	<0.001
Insurance status/source						
Private	(Ref)		(Ref)		(Ref)	
Public	-0.032[b]	0.004	-0.112	0.581	-0.182[b]	0.010
Uninsured	-0.083[c]	<0.001	0.232	0.636	-0.375[c]	0.001
Other	0.033	0.483	-0.513	0.363	0.045	0.858
Patient BMI at Admission	0.0001	0.866	0.023[b]	0.001	0.005[a]	0.032
Weighted CCI score	-0.001	0.446	-0.015	0.715	-0.010	0.417
SOFA score at admission	0.054[c]	<0.001	0.387[c]	<0.001	0.348[c]	<0.001
ICU length of stay (days)	0.009[c]	<0.001	0.256[c]	<0.001	0.098[c]	<0.001
Predicted mean	0.276		6.552		1.367	
Number of encounters	13137		4331		13137	

[a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.001$

Table 4. Adjusted two-part model marginal effects of guideline discussions on albumin use per encounter during the first ICU stay.

	Part 1: LOGIT	*p*-value	Part 2: GLM (NB)	*p*-value	Overall Effects	*p*-value
Independent guideline discussion effect	-0.004	0.844	-1.186[a]	0.014	-0.263	0.059
Encounter during financial incentive period (vs. baseline)	-0.016	0.407	-0.329	0.483	-0.148	0.278
Encounter in a guideline discussion ICU (vs. non-guideline discussion)	0.287[c]	<0.001	2.800[c]	<0.001	1.993[c]	<0.001
Patient age	-0.001[c]	<0.001	0.009	0.123	-0.004	0.069
Patient sex is male (vs. female)	0.017	0.056	0.195	0.270	0.125[a]	0.030
Patient race is White (vs. non-White)	0.122[c]	<0.001	0.829[c]	<0.001	0.787[c]	<0.001
Insurance status/source						
Private	(Ref)		(Ref)		(Ref)	
Public	-0.032[b]	0.004	-0.100	0.623	-0.179[a]	0.011
Uninsured	-0.083[c]	<0.001	0.225	0.645	-0.376[b]	0.001
Other	0.033	0.483	-0.513	0.360	0.045	0.858
Patient BMI at Admission	0.0001	0.867	0.022[c]	<0.001	0.005[a]	0.030
Weighted CCI score	-0.001	0.445	-0.016	0.687	-0.011	0.402
SOFA score at admission	0.054[c]	<0.001	0.390[c]	<0.001	0.349[c]	<0.001
ICU length of stay (days)	0.009[c]	<0.001	0.257[c]	<0.001	0.098[c]	<0.001
Predicted mean	0.276		6.550		1.367	
Number of encounters	13137		4331		13137	

[a]$p < 0.05$, [b]$p < 0.01$, [c]$p < 0.001$

Sensitivity Analysis

Results in Table 5 present the main model and additional adjusted, two-part model results as the definitions of the pre- and post-periods were altered. Comparison of different time-period definitions helps to evaluate the robustness the main model estimates. The first alternative study time-period change included all admissions and set the beginning of the post-period to July 1. This time-period change resulted in a small 0.05 order decrease in the magnitude of the financial incentive marginal effect and did not result in major changes for the guideline discussion effects.

The second alternative pre- and post-period adjustment also included all admissions, but set the post-period start date as September 1. Financial incentive effects with this alternate design did not differ substantially from those effects from the main analysis. However, the guideline discussions effect did suffer a 0.03 order decrease in the magnitude of the marginal effect and larger standard errors, which increased the *p*-value by 0.02.

The final alternative design included all admissions and set the post-period to begin on November 1. The magnitudes of both the financial incentive and guideline discussion effects decreased by approximately 0.05 and 0.06, respectively. The standard errors on the guideline discussion effect estimate also increased substantially, such that the *p*-value increased to 0.128.

Table 5. Sensitivity analysis of change in study time-period definition on magnitude and *p*-values of financial incentive and guideline discussion independent effects

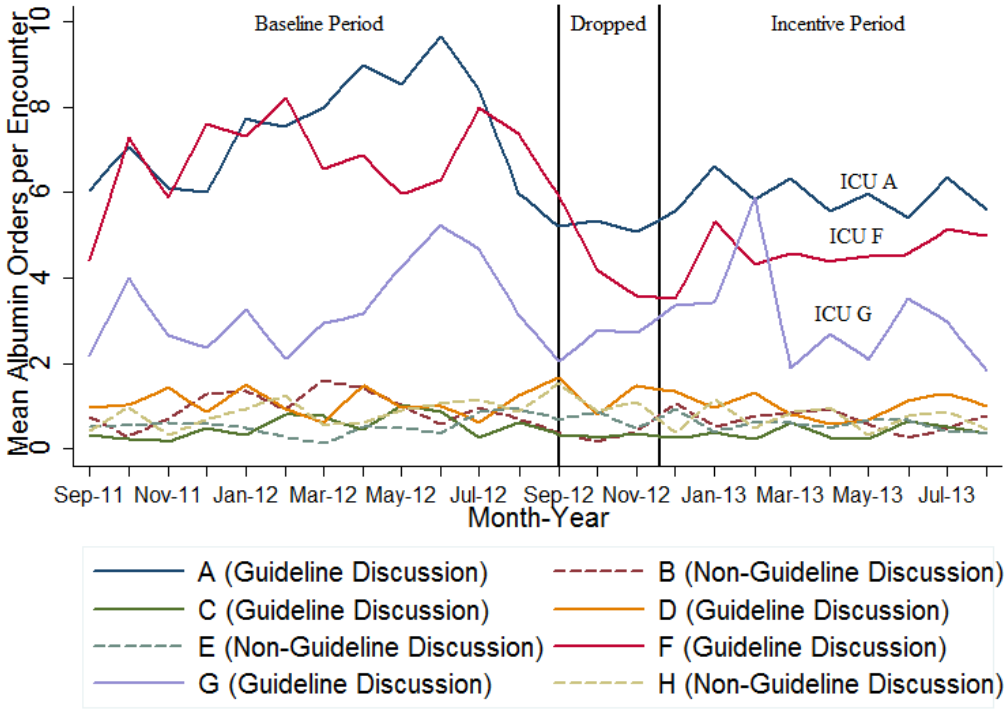
	Independent Financial Incentive Marginal Effect	*p*-value	Independent Guideline Discussion Marginal Effect	*p*-value
Main analysis				
Drop admissions between Sept 1, 2012-Nov 19, 2012	-0.378	<0.001	-0.263	0.059
Alternative models				
Include all admissions; incentive-period begins on July 1, 2012	-0.329	<0.001	-0.261	0.055
Include all admissions; incentive-period begins on September 1, 2012	-0.383	<0.001	-0.231	0.079
Include all admissions; incentive-period begins on November 1, 2012	-0.328	<0.001	-0.203	0.128

Note: marginal effects were estimated at the predicted means

Subgroup Analysis of High and Low Albumin Users

As Figure 4 illustrates, the majority of albumin use variation occurred in only three of the eight total ICUs. These three ICUs all stood out as higher users of albumin in the baseline period and were all involved in the guideline discussions. In order to identify differences in effect based on an ICU's baseline albumin use, the subgroup analysis tested between high users and low users within the Guideline Discussion group.

Figure 4. Monthly albumin use per encounter for all individual ICUs



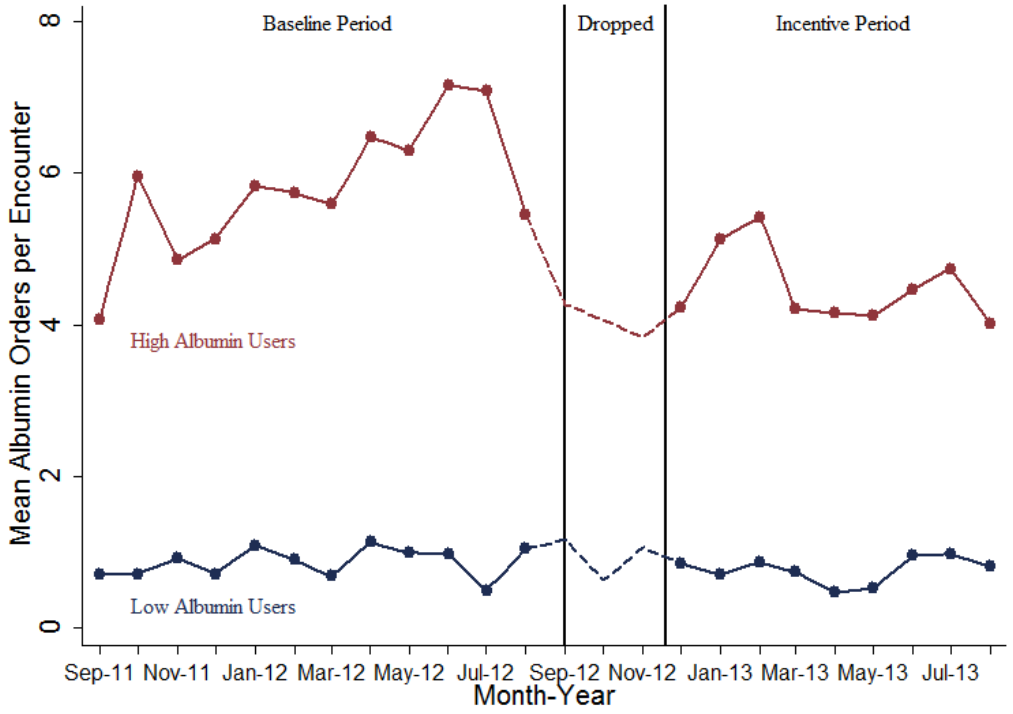
Differences in means from this subgroup analysis represent the combined, unadjusted change associated with both the financial incentive program and guideline discussions. As summarized in Table 6, this combined effect was significantly different between ICUs with high baseline albumin use and low baseline albumin use. Low-baseline-use ICUs saw a small 0.1 decrease in mean albumin orders per encounter ($p=0.152$). In contrast, high-baseline-use ICUs saw a much larger 1.4 decrease in mean albumin orders per encounter ($p<0.001$). These high vs. low user differences within the Guideline Discussion group are also illustrated in Figure 5.

Table 6. Unadjusted sub-group analysis of the financial incentive effect within the Guideline Discussion group between ICUs with lower versus higher baseline albumin use

Orders per encounter, mean (se)	Baseline Period	Incentive/Guideline Discussion Period	Difference in Albumin Use
Albumin use in Guideline Discussion Group	3.7 (0.1)	2.8 (0.1)	-0.9 (0.1)*
ICUs with low baseline albumin use	0.9 (0.1)	0.8 (0.1)	-0.1 (0.1)
ICUs with high baseline albumin use	5.9 (0.2)	4.4 (0.1)	-1.4 (0.2)*

*$p<0.05$

Figure 5. Monthly albumin use per encounter for only the Guideline Discussion Group, comparing ICUs with high versus low baseline albumin use



Additional Outcomes

While the primary outcome of this study is utilization, other clinical and administrative outcomes related to albumin use offers insight into the potential downstream effects of decreased albumin use. Table 7 compares mortality and costs for all eight ICUs between the study's baseline and incentive periods. Decreased albumin use across this 2-year

period did not coincide with any significant change in overall first ICU, any ICU, or hospital mortality.

Table 7. Unadjusted comparison of mean mortality and cost outcomes between the baseline and incentive periods

Outcome Variable (for all ICUs)	Baseline Period (n=9,094)	Incentive Period (n=7,023)	Difference	*p*-value
Mortality, % (se)				
First ICU stay	0.9 (0.1)	1.0 (0.1)	0.1	0.585
Any ICU stay	4.8 (0.2)	5.1 (0.3)	0.2	0.508
Hospital	5.7 (0.2)	5.4 (0.3)	-0.2	0.513
Costs, mean $ per encounter (se)				
Direct albumin fluid costs (all patients)	438 (11)	344 (10)	-94	<0.001
Direct albumin fluid costs (albumin patients only)	1,342 (28)	1,067 (26)	-275	<0.001
Total hospital costs	36,746 (426)	35,853 (532)	-893	0.185

Assuming an estimated albumin order cost of $170.89, mean direct albumin fluid costs per encounter saw a significant $94 decrease among all patients. When looking at only patients that actually used albumin, mean direct albumin fluid costs per encounter decreased by $275. This decrease translated into direct aggregate albumin savings of approximately $782,334 between FY2012 and FY2013, including the middle period of dropped encounters. Mean total hospital costs per encounter, which include but are not limited to albumin costs, did not observe any significant change between the baseline and incentive periods.

DISCUSSION

Summary of Findings

This study takes advantage of a quasi-experimental setting to determine whether two institutional interventions, a multi-faceted financial incentive program and guideline

discussions, led to decreased albumin use in ICU settings. Using two-part models to control for covariates and estimate independent effects, findings revealed that the financial incentive program was significantly associated with decreased probability and volume of albumin use. Similarly, guideline discussions significantly predicted fewer orders for those patients who received any albumin, but did not seem to impact the mean probability of whether any albumin was ordered during the encounter. These findings were robust to changes in the pre- and post-period timeframes. A sub-group analysis revealed that decreases among ICUs with the highest baseline albumin use constituted the majority of the overall decrease observed system-wide. These changes in albumin use were achieved without any significant change in hospital or ICU mortality.

Impact on Provider Decision-Making

Estimates produced in each part of the two-part models can provide additional insights on how these interventions affected providers' decision to order albumin. Among encounters with any albumin use, financial incentive and guideline discussion effects estimated large relative decreases in the number of orders relative to their predicted means. In contrast, relative effects on the probability of any albumin use were much smaller in magnitude and statistical significance. Thus, both interventions demonstrated much larger relative effects on the volume of use among those administered albumin compared to the probability of using albumin at all.

From a practical perspective, this finding could suggest that neither the financial incentive program nor guideline discussions changed the way providers determined

which patients were prescribed albumin. Rather, the interventions may have impacted *how much* albumin was ordered after the decision to use albumin was made. Evidence from this study supports this possibility as the average patient had a similar likelihood of receiving any albumin between providers exposed to the interventions and those not involved in either intervention. However, if a provider decided that albumin should be used, that patient was likely to receive significantly less albumin during the first ICU stay due to the two interventions. This conclusion is consistent with the incentives' intended effect since the incentive program's rules do not specify the types of patients that should receive albumin. On the other hand, the guideline discussion meetings undoubtedly involved talks on which patients should receive albumin. One expectation might be that committee members would bring back these insights to their respective ICUs, but the results suggest that this was not the case. Instead, guideline discussions may have simply raised general awareness about albumin overuse and affected decisions on dosage rather than on who would receive albumin.

Results from the sub-group analysis also shed additional light on providers' responsiveness to the interventions. ICUs with lower baseline albumin use may have been subject to a "ceiling effect," a barrier commonly discussed in P4P literature [25]. The ceiling effect proposes that incentives to improve care become less effective the closer baseline performance rises toward the desired level of care, or "ceiling" performance. In the context of this study, the ceiling performance could be represented as albumin use with an average cost of $15 per patient day (the incentive threshold), or approximately 0.1 orders per patient day. ICUs with high baseline albumin use were the

furthest from the ceiling and had the most space to demonstrate improvements. Thus, the ceiling effect may explain why high-baseline-use ICUs reported the largest decreases in albumin use while the remaining low-baseline-use ICUs did not demonstrate a significant change.

Furthermore, the "free rider" effect may have had a differential effect between high- and low-baseline-use ICUs. In this study, a free rider effect was initially expected for all ICUs because the incentive was structured around group performance. However, decreases in high-baseline-use ICUs were much larger than the change seen among low-baseline-use ICUs. Past research by Knight, Durham, and Locke may help to explain these differences [47]. Their research on the relationship between team-based incentives, difficulty-level of goals, and performance has suggested that teams facing more difficult goals are actually more likely to take strategic risks and ultimately perform better. Thus, as ICUs with high baseline albumin use faced a challenging goal of achieving significant reductions, they may have been motivated to implement more ambitious strategies that ultimately decreased albumin use. In contrast, low-baseline-use ICUs needed to achieve a much smaller decrease and may have been less motivated as a team to develop a coordinated strategy. While this theory is certainly limited in its generalizability to health care environments, where risk is characterized by mortality, these findings still offer an additional explanation for why low-baseline-use ICUs responded less to the interventions.

Clinical & Policy Implications

This study's findings are especially relevant in today's modern health care setting where costs have a growing role in clinical decision making. More providers are recognizing high-quality care and cost-efficient care to be congruent goals, but changing traditional standards of practice to achieve these goals will require evidence-based strategies. Explicit financial incentives have long been debated as an effective strategy for changing physician behavior. These results continue that important debate and offer support for the effectiveness of financial incentive programs given the strong associations and causal plausibility found with decreased albumin use. This study also proposes that implementation of financial incentives might be more effective if coupled with guideline discussions, providing both clinical and cost motivations to change practice behavior. Large decreases in albumin use among ICUs affected by both interventions support this hypothesis.

Practitioners and health care administrators in critical care should also consider applying and evaluating these interventions in areas beyond IV fluids. In fact, a report compiled by numerous medical specialty organizations identifies other overused services, such as red blood cell transfusions and chest x-rays, that resemble the high-cost, high-use nature of albumin [48]. For institutions that suspect overuse of these services, financial incentives and guideline discussions at the local level could help to reduce utilization.

On a broader policy level, the optimistic results from this study might signal a potential space for public payers or providers to lower costs. However, pursuing these changes in

the public sector may prove more challenging than initiatives born out of private health care organizations. Policymakers and much of the American public are exceptionally wary of the role of costs in driving medical decisions, especially in the context of government policy [49]. Nevertheless, growing pressure to lower costs in both public and private sectors may help overcome hesitation around government-funded cost effectiveness research in medicine. While cost-effectiveness evaluation was not in the scope of this study, the promising results found in this study will hopefully spur providers and researchers to evaluate the cost effectiveness of financial incentive interventions tied to guideline discussions.

Study Limitations

Internal Validity

This study's primary limitation is the lack of a randomized controlled design. Since this study takes an observational perspective, neither patients nor providers were randomized into intervention groups. Certain providers and ICUs were actually selected into the Guideline Discussion Group because they saw patients more likely to require IV fluid therapy and thus had a higher propensity to use albumin. In other words, these ICUs had providers with more experience with albumin treatment and were subsequently invited to help develop institutional guidelines. Due to this selection bias, Guideline Discussion ICUs had much higher baseline use compared to Non-Guideline Discussion ICUs. Since all ICUs with high baseline use were involved in guideline discussions, the guideline discussion effect could not be separated from an effect related to providers' propensity to use albumin. Nevertheless, this limitation does not take away from the conclusion that the

highest users demonstrated the largest response to guideline discussions and financial incentives, which was the intended goal of these interventions.

Covariates were also included in the regression models to control for confounding differences between the study groups. However, EMR data did not have reliable or detailed information on prescribing providers, so specific provider factors could not be identified or controlled. Fortunately, the difference-in-differences design used to assess guideline discussion effects partly addresses these issues. In difference-in-differences analysis, each study group uses its own baseline year as a control for time-invariant confounders, such as mix of providers. However, such designs also assume that mix of providers remains the same from year to year, but this may not actually be the case with some providers such as medical residents.

In the analysis of financial incentive effects, selection bias was not an issue since all ICUs participated in the incentive. However, the lack of a control group that was not incentivized allows for possible history bias. Without a viable control group, this study could not isolate financial incentive effects from other unmeasured environmental events occurring during the same time period. As such, financial incentive effect estimates derived in this study are interpreted as associations with plausible causality given the confounding factors that could be captured and controlled.

Study groups may have also experienced a crossover effect due to specialists on the Albumin Utilization Taskforce Committee who were not officially associated with a

specific ICU but still consulted in both Guideline Discussion and Non-Guideline Discussion ICUs. The direction of this potential bias, however, is unclear. Consulting specialists likely do not have a direct, system- or culture-changing impact in the ICUs since they are only involved with a smaller portion of all ICU patients. On the other hand, consulting specialists help determine the course of treatment for patients, including IV fluid therapy strategy. Future qualitative investigation could help to clarify the impact of consulting specialists around quality improvement initiatives in the ICU.

Each ICU also had different patient populations due to their different critical care specialty. Unfortunately, with such a small number of ICUs, cluster analysis was deemed inappropriate [50]. Since data could not be clustered around ICUs, standard errors may have been underestimated. Though not a direct adjustment for intra-cluster correlation, robust standard errors were calculated using the Huber-White sandwich estimator in Stata, which does account for expected heteroskedasticity and model misspecification.

External Validity

This study can function as an example for evaluations of other cost-effectiveness interventions, but implementation of programs or initiatives based on these results should be considered with caution. This investigation was conducted in the critical care setting, so other types of providers in non-critical care settings (e.g. outpatient settings) may respond differently to financial incentives and guideline discussions. Differences between the patient populations at Emory Healthcare and those at other hospital systems should also be recognized, particularly by safety net hospitals that primarily serve high

concentrations of lower income, higher-risk populations. Finally, comparisons with other P4P studies should make note of the structural differences between this study's incentive program and that of other pay-for-performance initiatives. Importantly, ECCC's incentive program not only focused on reducing overuse, but also included mechanisms that gave providers monthly feedback on their albumin utilization relative to other ICUs. As a result, this study's findings reflect the impact of the incentive program as a whole, not of the financial incentives alone.

Future Directions

Results from this study raise a number of questions for further research. As briefly mentioned before, qualitative investigation could offer insightful explanations as to why some ICUs reported great decreases in use while others did not. This type of qualitative assessment would provide a powerful compliment to this study's quantitative focus. A future study should also assess the overall cost-effectiveness of this intervention on a broader scope, including costs of the incentive program, readmission rates, and post-discharge mortality.

Assessment of appropriateness of use can also be a more precise measure of guidelines' effectiveness in subsequent studies. This study chose not to assess a change in the appropriateness of albumin use because the official guidelines were not finalized until after the end of the incentive period. At Emory Healthcare, the final albumin guidelines were disseminated and implemented in late October 2013 along with a change in the computerized order systems, forcing prescribing providers to explicitly identify the

reason for a given albumin order. Future research is planned to assess whether this guideline implementation strategy continued the downward trend in use and changed the way providers identified patients for albumin treatment.

CONCLUSION

Debates on the effectiveness of financial incentives and pay-for-performance schemes have a long history, but the issue has experienced a new resurgence over the past decade with national attention and public policy focused on rising health care costs. Much of the debate remains inconclusive due to the wide variability in financial incentive designs and settings. This study adds to this debate by looking at the effect of a financial incentive program and guideline discussion meetings at Emory Healthcare's Critical Care Center.

Both the financial incentive program and guideline discussions were independently and significantly associated with decreased albumin use. Furthermore, ICUs with high baseline use that had the most space to demonstrate improvements responded with the largest decreases. These findings provide support for the effectiveness of explicit financial incentive initiatives in driving clinical practice change in critical care environments. Findings from this study also suggest that responsiveness to financial incentive programs may be greater when tied to empirical discussions on the efficacy of desired practices. However, since this study faced important limitations due to its observational nature, more rigorous studies are needed to confirm these findings. Researching the mechanisms for practice change will be especially important in cases

like albumin where existing medical evidence alone does not drive providers to make the most clinically valuable decisions.

REFERENCES

1. Goldwasser, P. and J. Feldman, *Association of serum albumin and mortality risk.* J Clin Epidemiol, 1997. **50**(6): p. 693-703.
2. Tanzi, M., et al., *Evaluation of the appropriate use of albumin in adult and pediatric patients.* Am J Health Syst Pharm, 2003. **60**(13): p. 1330-5.
3. Yim, J.M., et al., *Albumin and nonprotein colloid solution use in US academic health centers.* Arch Intern Med, 1995. **155**(22): p. 2450-5.
4. Myburgh, J.A. and M.G. Mythen, *Resuscitation fluids.* N Engl J Med, 2013. **369**(13): p. 1243-51.
5. Hall, R., *Artificial colloids: whither go we now?*.* Crit Care Med, 2013. **41**(11): p. 2651-3.
6. Cochrane Injuries Group Albumin Reviewers, *Human albumin administration in critically ill patients: systematic review of randomised controlled trials.* BMJ, 1998. **317**(7153): p. 235-40.
7. Cole, R.P., *The UK albumin debate.* Burns, 1999. **25**(7): p. 565-8.
8. Ferguson, N.D., T.E. Stewart, and E.E. Etchells, *Human albumin administration in critically ill patients.* Intensive Care Med, 1999. **25**(3): p. 323-5.
9. Wilkes, M.M. and R.J. Navickis, *Patient survival after human albumin administration. A meta-analysis of randomized, controlled trials.* Ann Intern Med, 2001. **135**(3): p. 149-64.
10. Finfer, S., et al., *A comparison of albumin and saline for fluid resuscitation in the intensive care unit.* N Engl J Med, 2004. **350**(22): p. 2247-56.

11. Perel, P., I. Roberts, and K. Ker, *Colloids versus crystalloids for fluid resuscitation in critically ill patients.* Cochrane Database Syst Rev, 2013. **2**: p. CD000567.

12. Vincent, J.L., et al., *Is albumin administration in the acutely ill associated with increased mortality? Results of the SOAP study.* Crit Care, 2005. **9**(6): p. R745-54.

13. Delaney, A.P., et al., *The role of albumin as a resuscitation fluid for patients with sepsis: a systematic review and meta-analysis.* Crit Care Med, 2011. **39**(2): p. 386-91.

14. Padhi, S., et al., *Intravenous fluid therapy for adults in hospital: summary of NICE guidance.* BMJ, 2013. **347**: p. f7073.

15. Caironi, P., et al., *Albumin Replacement in Patients with Severe Sepsis or Septic Shock.* N Engl J Med, 2014.

16. Rozga, J., T. Piatek, and P. Malkowski, *Human albumin: old, new, and emerging applications.* Ann Transplant, 2013. **18**: p. 205-17.

17. Charles, A., et al., *Albumin use guidelines and outcome in a surgical intensive care unit.* Arch Surg, 2008. **143**(10): p. 935-9; discussion 939.

18. Suarez, J.I., et al., *Effect of human albumin administration on clinical outcome and hospital cost in patients with subarachnoid hemorrhage.* J Neurosurg, 2004. **100**(4): p. 585-90.

19. Annane, D., et al., *Effects of fluid resuscitation with colloids vs crystalloids on mortality in critically ill patients presenting with hypovolemic shock: the CRISTAL randomized trial.* JAMA, 2013. **310**(17): p. 1809-17.

20. Grol, R. and J. Grimshaw, *From best evidence to best practice: effective implementation of change in patients' care.* Lancet, 2003. **362**(9391): p. 1225-30.

21. Armour, B.S., et al., *The effect of explicit financial incentives on physician behavior.* Arch Intern Med, 2001. **161**(10): p. 1261-6.

22. Cabana, M.D., et al., *Why don't physicians follow clinical practice guidelines? A framework for improvement.* JAMA, 1999. **282**(15): p. 1458-65.

23. James, J., *"Health Policy Brief: Pay-for-Performance"*, C.L. Damberg, Ryan, A. M., Agres, T. Schwartz, A., Dentzer, S., Editor 2012, Health Affairs.

24. Eijkenaar, F., et al., *Effects of pay for performance in health care: a systematic review of systematic reviews.* Health Policy, 2013. **110**(2-3): p. 115-30.

25. Van Herck, P., et al., *Systematic review: Effects, design choices, and context of pay-for-performance in health care.* BMC Health Serv Res, 2010. **10**: p. 247.

26. Rosenthal, M.B. and R.A. Dudley, *Pay-for-performance: will the latest payment trend improve care?* JAMA, 2007. **297**(7): p. 740-4.

27. Bardach, N.S., et al., *Effect of pay-for-performance incentives on quality of care in small practices with electronic health records: a randomized trial.* JAMA, 2013. **310**(10): p. 1051-9.

28. Chen, J.Y., et al., *The effect of a PPO pay-for-performance program on patients with diabetes.* Am J Manag Care, 2010. **16**(1): p. e11-9.

29. Chung, S., et al., *Effect of physician-specific pay-for-performance incentives in a large group practice.* Am J Manag Care, 2010. **16**(2): p. e35-42.

30. Petersen, L.A., et al., *Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial.* JAMA, 2013. **310**(10): p. 1042-50.

31. Town, R., et al., *Assessing the influence of incentives on physicians and medical groups.* Med Care Res Rev, 2004. **61**(3 Suppl): p. 80S-118S.

32. Damberg, C.L., et al., *Taking stock of pay-for-performance: a candid assessment from the front lines.* Health Aff (Millwood), 2009. **28**(2): p. 517-25.

33. Debrix, I., et al., *Clinical practice guidelines for the use of albumin: results of a drug use evaluation in a Paris hospital. Tenon Hospital Paris.* Pharm World Sci, 1999. **21**(1): p. 11-6.

34. Durand-Zaleski, I., et al., *Usefulness of consensus conferences: the case of albumin.* Lancet, 1992. **340**(8832): p. 1388-90.

35. Mirici-Cappa, F., et al., *How albumin administration for cirrhosis impacts on hospital albumin consumption and expenditure.* World J Gastroenterol, 2011. **17**(30): p. 3479-86.

36. Fan, C.P., K.; Selin S., *Serum albumin: New thoughts on an old treatment.* BCMJ, 2005. **47**(8): p. 438-444.

37. Vermeulen, L.C., Jr., et al., *A paradigm for consensus. The University Hospital Consortium guidelines for the use of albumin, nonprotein colloid, and crystalloid solutions.* Arch Intern Med, 1995. **155**(4): p. 373-9.

38. Grimshaw, J.M. and I.T. Russell, *Achieving health gain through clinical guidelines II: Ensuring guidelines change medical practice.* Qual Health Care, 1994. **3**(1): p. 45-52.

39. Rosenberg, A.L. and C. Watts, *Patients readmitted to ICUs* : a systematic review of risk factors and outcomes.* Chest, 2000. **118**(2): p. 492-502.

40. Magnus, S.A., *Physicians' financial incentives in five dimensions: a conceptual framework for HMO managers.* Health Care Manage Rev, 1999. **24**(1): p. 57-72.

41. Lyon, S.M., et al., *The effect of insurance status on mortality and procedural use in critically ill patients.* Am J Respir Crit Care Med, 2011. **184**(7): p. 809-15.

42. Soto, G.J., G.S. Martin, and M.L. Gong, *Healthcare Disparities in Critical Illness.* Crit Care Med, 2013.

43. Charlson, M.E., et al., *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation.* J Chronic Dis, 1987. **40**(5): p. 373-83.

44. Vincent, J.L., et al., *The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine.* Intensive Care Med, 1996. **22**(7): p. 707-10.

45. Bahn, G.D., Massenburg, R., *Deal with Excess Zeros in the Discrete Dependent Variable, the Number of Homicide in Chicago Census Tract*, in *Joint Statistical Meeting*2008, American Statistical Association. p. 3905-3912.

46. Wooldridge, J.M., *Introductory Econometrics: A Modern Approach*. Fifth ed. 2013, Mason, OH: South-Western.

47. Knight, D., C.C. Durham, and E.A. Locke, *The relationship of team goals, incentives, and efficacy to strategic risk, tactical implementation, and performance.* Academy of Management Journal, 2001. **44**(2): p. 326-338.

48. ABIM Foundation, *Choosing Wisely: An initiative of the ABIM Foundation*, 2013.

49. Garber, A.M. and H.C. Sox, *The role of costs in comparative effectiveness research.* Health Aff (Millwood), 2010. **29**(10): p. 1805-11.

50. Angrist, J.D., Pischke, J., *Mostly Harmless Econometrics: An Empiricist's Companion*. 2009, Princton, NJ: Princeton University Press.

APPENDIX

Table A1. Summary of individual ICUs and their respective specialty type and study group

ICU Identifier	Study Group	Type
A	Guideline Discussion	Cardiothoracic
B	Control	Neuroscience
C	Guideline Discussion	Coronary Care
D	Guideline Discussion	Medical
E	Control	Neuroscience
F	Guideline Discussion	Cardiovascular
G	Guideline Discussion	Surgical
H	Control	Medical

Table A2. Summary of final Emory Healthcare indication guidelines for albumin use

Albumin 5%	Albumin 25%
Hemorrhagic Shock	Intradialytic Hypotension
	Acute Respiratory Distress Syndrome
Septic Shock	(ARDS)
ENT Free Flap Postop	Diuretic Resistant Nephrotic Syndrome
Post Cardiopulmonary Bypass	Large Volume Paracentesis
	Immediate Post Liver Transplantation
	Spontaneous Bacterial Peritonitis (SBP)
	Suspected Hepato-Renal Syndrome (HRS)
	Confirmed Hepato-Renal Syndrome (HRS)

Table A3. Summary of Sequential Organ Failure Assessment (SOFA) scoring

	Score				
SOFA Component	0	1	2	3	4
Respiratory					
PaO_2/FiO_2, mmHg	≥400	300-399	200-299	100-199	<100
	&	&	or	&	&
Any mechanical ventilation, yes/no	NO	NO	YES	YES	YES
Renal					
Creatinine, mg/dL	<1.2	1.2-1.9	2.0-3.4	3.5-4.9	≥5.0
	&	&	&	or	or
Urine output, mL/day	≥500	≥500	≥500	200-499	<200
Hepatic					
Bilirubin, mg/dL	<1.2	1.2-1.9	2.0-5.9	6.0-11.9	≥12.0
Hematologic					
Platelets, $x10^3/mm^3$	≥150	100-149	50-99	20-49	<20
Neurologic					
Glasgow coma score	15	13-14	10-12	6-9	<6
Cardiovascular					
Mean arterial pressure (MAP), mmHg	≥70	<70	Any MAP	Any MAP	Any MAP
	&	&	&	&	&
Any vasopressor use	NO	NO	YES, either dopamine or dobutamine	YES, but only ONE of: norepinephrine, epinephrine, phenylephrine, or vasopressin	YES, but MORE THAN ONE of: norepinephrine, epinephrine, phenylephrine, vasopressin, or dopamine

TOTAL SOFA SCORE = respiratory + renal + hepatic + hematologic + neurologic + cardiovascular