Article details: 2013-0001	
Title	Early Cholecystectomy for Acute Cholecystitis: A Population-based Analysis of Variability in Practice
Authors	de Mestral et al.
Reviewer 1	Dr. Steven R. Lopushinsky
Institution	Alberta Children's Hospital
General comments	Summary: The authors present a population-based retrospective analysis of patients admitted to Ontario, Canada emergency rooms with acute cholecystitis with the aim to assess variability in early cholecystectomy. They identified 24,437 patients meeting inclusion criteria with 58% undergoing early cholecystectomy consistent with current best evidence. Across 106 hospitals, the authors identified significant variation in the rate of early cholecystectomy with median rate of 51% (IQR 25% - 72%). With a sophisticated statistical analysis, both patient factors and hospital factors were found contributing to this variation. It was suggested that two similar patients presenting to different hospitals had a nearly 4-fold median difference in the odds of undergoing early cholecystectomy. Given that best evidence in the literature supports early cholecystectomy, the authors conclude that hospital-based initiatives to increase early surgery rates are warranted. Major Points: 1) This is a well-written paper further highlighting how best evidence often does not translate into best clinical practice. The strength of this paper lies in its ability to evaluate an entire health system, in a single-payer model, to capture all patients presenting to an emergency room for the first time with acute cholecystitis. The use of large databases such as those used in this study give powerful insight into the use of health services and variability in care provided to individual patients. In the literature, it is clear that patients benefit by early cholecystectomy (shorter hospital stays, fewer readmissions for disease complications, and enjoy the same conversion rate from laparoscopic to open surgical techniques). With the data available, the authors have attempted to account for patient comorbidity and have demonstrated that significant differences remain in how patients are treated from one institution to another. Based on the data presented, it is apparent that an institutional level, rather than solely at the provider level, changes can be made to facilitate improved provision of best care. The biggest weakness of such studies however is in the accuracy of the data and the limitation of variables recorded. The paper is strengthened by a previous validation study supporting the coding of gallstone disease and cholecystectomy procedures. Unfortunately despite a high-level statistical analysis, using multi-level logistic regression, the model was only able to explain about 50% of the variation. The biggest void in the model is the lack of provider-level variables such as admitting specialty, age and/or year of graduation, fellowship training, country of training, and volume of cholecystectomy procedures to name a few. While we can hypothesize hospital factors such as the availability of daytime operating room time influence a surgeon's decision making, at the end of the day, it is the responsibility of that surgeon to make an educated decision. Further, the hospital-level variables are somewhat crude, and likely do not adequately separate hospital policies or culture from one institution to another. Such factors as hospital capacity, availability of ERCP, presence of an ICU and others come to mind as those potentially contributing to hospital variation. Variables presented such as teaching status, annual number of acute cholecystitis admissions or elective cholecystectomies performed do not lend for policy change, as they are somewhat innate to that institution. Having said that, they may provide some insight as what can be learned from high-performing institutions. 2) With respect to the patient population, significant effort is made to capture and account for patient differences (i.e. excluding patients admitted to an ICU or undergoing cholecystostomy tube placement, and including patient variables such as age and comorbidity in the analysis). Given the fact that the primary outcome is early cholecystectomy (<7 days), the study would be strengthened by looking at the number of patients not undergoing early surgery who actually underwent surgery at a later date. I believe that this would further eliminate patient-level confounding or unmeasured bias not currently captured in the model. At the very least, it would suggest what percentage of patients not undergoing early surgery seem to be reasonable surgical candidates. 3) The statistical analysis is clearly sophisticated and appropriate with multi-level logistic regression modeling. As a non-statistician it can be a little difficult to understand, with my primary questions around the funnel plots and the use of ninety-nine percent control

	limits. Could the authors please comment on the fact that the majority of data points actually sit outside of these limits? Can this type of analysis be applied to data that is not normally distributed? When the mean and control limits are generated (3 standard deviations from that mean) from the data, I am not sure how to interpret the chart given that such a majority of points sit outside these limits. 4) At the end of the day, perhaps the most important outcome is to improve patient outcome and perhaps learn from the high performing institutions. Cardiac surgery report cards in the United States are an early example of this where such report cards have been shown to improve performance. Were individual institutions made aware of their results / relative performance? If early cholecystectomy was considered a quality of care indicator, how do we learn from the high performers to re-create their results? Minor Points 1) Cohort definition (page 7 lines 41-47): Why did you choose appendectomy as defining procedure for including a hospital rather than cholecystectomy itself? 2) Variability in rate of early cholecystectomy (page 11 lines 11 -15): Although the variability remained when the data was limited to the young and healthy (the IQR's stable) the median rate increases from 51% to 74% (IQR 41 – 88%) suggesting that the data is somewhat skewed to the right and perhaps indicating lesser variability. 3) Association of patient and hospital-level characteristics (page 11 lines 49 – 56): Can a median odds ratio (MOR) be interpreted in a way similar to small area variation statistics where higher values can be expected where there is less consensus on the indications for surgery (i.e. circumcision) versus those with harder indications such as hip replacement. If so, is there data around the MOR for other procedures to put it into context?
Reviewer 2	Dr. May C Tee
Institution	University of British Columbia
General comments	General Comments: • This is an interesting population-based study examining practice differences between hospitals with respect to early versus delayed cholecystectomy. The authors examine several covariates using multivariable logistic regression models including patient and hospital characteristics to identify predictors of what may explain practice variation. • The authors neglect to examine surgeon characteristics, which practically speaking may be the most important variable in determining whether patients undergo early versus delayed cholecystectomy. Given the fact that cholecystectomy in an emergent setting requires a certain degree of comfort and confidence (on the part of the surgeon) with respect to operating in non-ideal circumstances, surgeon practice preference could quite possibly account for the variation in patient care among the hospitals. • The authors seem almost patronizing that the "best available evidence" suggests early cholecystectomy is preferable to delayed cholecystectomy when those studies may not be generalizable to the authors' diverse study population. I would caution the authors to use more neutral language. Major Revisions: • As stated above, it would be interesting to evaluate surgeon characteristics as a potential variable explaining practice differences in early versus delayed cholecystitis. Ultimately, the decision on operative timing is a highly complex one that must take into account hospital resources, acuity of patient illness, and surgeon experience with emergent cases that may be best managed in the initial period with a trial of non-operative management. The remaining 47% of unexplained variability from this paper might be answered by looking at this particular issue. • The authors use several surrogate markers: (1) appendectomies performed at a hospital as a surrogate marker of presence of a general surgeon, (2) ICU admission and percutaneous cholecystostomy as a surrogate of severe acute cholecystitis, and (3) income based on postal code demographics as a surrogate for socioeconomic status. There are issues with using these markers as surrogates: (1) in very remote areas of Ontario, appendectomies may actually be performed by general practitioners who have extra surgical training, which may bias the results toward more cases of delayed cholecystectomy; (2) ICU admission and percutaneous cholecystostomy tube are not necessarily surrogate markers of severe cholecystitis, in fact, they may be markers of non-surgical disease (e.g. acalculous cholecystitis, which is often associated with critically ill patients who do not actually require

	cholecystectomy) or surgeon preference (e.g. a surgeon who prefers delayed cholecystectomy and temporizes patients with percutaneous cholecystostomy tubes); (3) income is also a very crude marker of socioeconomic status, which encompasses so many more complex variables and when based on postal code census data, hardly qualifies as a truly precise or accurate measure. • The discussion appears to imply early cholecystectomy for cholecystitis as the standard of care, which is inaccurate. In fact, at the most recent Canadian Association of General Surgeons forum in Calgary (September 2012), there was still ongoing plenary debates about early versus delayed cholecystectomy. The fact is that there is variation in practice because there is no national consensus on what truly is better. The authors ought to amend their discussion to appear less subjective in order to more accurately discuss the results which are that there is still a huge variation in care for patients with acute cholecystitis and that patients who do not receive early cholecystectomy are not deviations from standard of care. The standard of care for this common acute surgical problem is still evolving, as is the concept of an acute surgical care service model for hospitals. Minor Revisions: • The abstract should define the abbreviation AC (acute cholecystitis) prior to its ongoing use.
Author response	**<u>Reviewer: 1</u>** **Comments to the Author** Summary: The authors present a population-based retrospective analysis of patients admitted to Ontario, Canada emergency rooms with acute cholecystitis with the aim to assess variability in early cholecystectomy. They identified 24,437 patients meeting inclusion criteria with 58% undergoing early cholecystectomy consistent with current best evidence. Across 106 hospitals, the authors identified significant variation in the rate of early cholecystectomy with median rate of 51% (IQR 25% - 72%). With a sophisticated statistical analysis, both patient factors and hospital factors were found contributing to this variation. It was suggested that two similar patients presenting to different hospitals had a nearly 4-fold median difference in the odds of undergoing early cholecystectomy. Given that best evidence in the literature supports early cholecystectomy, the authors conclude that hospital-based initiatives to increase early surgery rates are warranted. **Major Points:** This is a well-written paper further highlighting how best evidence often does not translate into best clinical practice. The strength of this paper lies in its ability to evaluate an entire health system, in a single-payer model, to capture all patients presenting to an emergency room for the first time with acute cholecystitis. The use of large databases such as those used in this study give powerful insight into the use of health services and variability in care provided to individual patients. In the literature, it is clear that patients benefit by early cholecystectomy (shorter hospital stays, fewer readmissions for disease complications, and enjoy the same conversion rate from laparoscopic to open surgical techniques). With the data available, the authors have attempted to account for patient comorbidity and have demonstrated that significant differences remain in how patients are treated from one institution to another. Based on the data presented, it is apparent that an institutional level, rather than solely at the provider level, changes can be made to facilitate improved provision of best care. The biggest weakness of such studies however is in the accuracy of the data and the limitation of variables recorded. The paper is strengthened by a previous validation study supporting the coding of gallstone disease and cholecystectomy procedures. Unfortunately despite a high-level statistical analysis, using multi-level logistic regression, the model was only able to explain about 50% of the variation. The biggest void in the model is the lack of provider-level variables such as admitting specialty, age and/or year of graduation, fellowship training, country of training, and volume of cholecystectomy procedures to name a few. While we can hypothesize hospital factors such as the availability of daytime operating room time influence a surgeon's decision making, at the end of the day, it is the responsibility of that surgeon to make an educated decision. Further, the hospital-level variables are somewhat crude, and likely do not adequately separate hospital policies or culture from one institution to another. Such factors as hospital capacity, availability of ERCP, presence of an ICU and others come to mind as those potentially contributing to hospital variation. Variables presented such as teaching status, annual number of acute cholecystitis admissions or elective cholecystectomies performed do not lend for policy change, as they are somewhat innate to that institution. Having said that, they may provide some insight as what can be learned from high-performing institutions.

	1. With respect to the patient population, significant effort is made to capture and account for patient differences (i.e. excluding patients admitted to an ICU or undergoing cholecystostomy tube placement, and including patient variables such as age and comorbidity in the analysis). Given the fact that the primary outcome is early cholecystectomy (<7 days), the study would be strengthened by looking at the number of patients not undergoing early surgery who actually underwent surgery at a later date. I believe that this would further eliminate patient-level confounding or unmeasured bias not currently captured in the model. At the very least, it would suggest what percentage of patients not undergoing early surgery seem to be reasonable surgical candidates. In this analysis, we focused on the decision to perform early cholecystectomy, taken at the time of index admission. Whether delayed elective cholecystectomy ultimately occurs relates to the extent to which a patient's perioperative risk can be attenuated after discharge, and to patient preference. Furthermore, delayed urgent cholecystectomy may occur based on the severity of any residual or recurrent symptoms, despite a patient being initially deemed a poor surgical candidate. Given these consideration, we did not include the differentiation the reviewer suggests in this analysis. We have however described the clinical course of patients discharged without cholecystectomy in a separate paper that is cited in the introduction (ref #13: de Mestral et al. J Trauma Acute Care Surg 2013) **2.** The statistical analysis is clearly sophisticated and appropriate with multi-level logistic regression modeling. As a non-statistician it can be a little difficult to understand, with my primary questions around the funnel plots and the use of ninety-nine percent control limits. Could the authors please comment on the fact that the majority of data points actually sit outside of these limits? Can this type of analysis be applied to data that is not normally distributed? When the mean and control limits are generated (3 standard deviations from that mean) from the data, I am not sure how to interpret the chart given that such a majority of points sit outside these limits. The funnel plot graphically represents whether variation in the rate of early cholecystectomy across hospitals is in excess or within the range expected based on chance alone. Ninety-nine percent control limits frame the range of random variation around the overall mean cholecystectomy rate and are defined as exact binomial confidence intervals that vary as a function of a hospital's volume of acute cholecystitis admissions. A hospital outside the control limits is interpreted as having an early cholecystectomy rate outside the range of random variation that would be expect based on chance alone. The large number of hospitals outside the control limits therefore indicates that variation in the rate of early cholecystectomy across hospitals was in excess of that expected by chance alone. In other words, factors other than chance explain the extent of variation. Clarification has been added to both the methods and results sections of the manuscript. **3.** At the end of the day, perhaps the most important outcome is to improve patient outcome and perhaps learn from the high performing institutions. Cardiac surgery report cards in the United States are an early example of this where such report cards have been shown to improve performance. Were individual institutions made aware of their results / relative performance? If early cholecystectomy was considered a quality of care indicator, how do we learn from the high performers to re-create their results? That is an excellent suggestion. In fact, the idea of using funnel plots to compare early cholecystectomy rates across hospitals in Ontario came from their use in the American College of Surgeon's Trauma Quality Improvement Project that produces report cards on trauma center performance. Restrictions on data use prohibit the authors of this manuscript from identifying individual centers. However, if a hospital's surgeon-in-chief or general surgery division head is aware of their hospital's early cholecystectomy rate, a crude understanding of where they stand relative to others can be gained from the presentation in this manuscript. **Minor Points:** **4.** Cohort definition (page 7 lines 41-47): Why did you choose appendectomy as defining procedure for including a hospital rather than cholecystectomy itself? Appendectomy was chosen as the surrogate for surgeon availability since appendicitis is common and is managed in the majority of cases with urgent appendectomy if surgical expertise is available. Using cholecystectomy would have meant excluding patients based on a characteristic directly related to the outcome and would have hindered our understanding of the extent of variation in practice. An addition has been made to the manuscript to explain the use of appendectomy. **5.** Variability in rate of early cholecystectomy (page 11 lines 11 -15): Although the variability remained when the data was limited to the young and healthy (the IQR's stable) the median rate increases from 51% to 74% (IQR 41 – 88%) suggesting that the data is somewhat skewed to the right and perhaps indicating lesser variability.

As the reviewer suggests, within the more homogeneous younger subgroup, there is less variation as evidenced by a smaller proportion of hospitals with an early cholecystectomy outside the range of random variation outlined by the control limits. This point has been added to the manuscript. As the total number of patients decreases and the overall early cholecystectomy rate approaches 1, the control limits adjust with the skew of the assumed binomial distribution.

6. Association of patient and hospital-level characteristics (page 11 lines 49 – 56): Can a median odds ratio (MOR) be interpreted in a way similar to small area variation statistics where higher values can be expected where there is less consensus on the indications for surgery (i.e. circumcision) versus those with harder indications such as hip replacement. If so, is there data around the MOR for other procedures to put it into context?
The MOR is a relatively novel metric that we used to quantifying the extent of variation across hospitals, adjusted for patient characteristics. The expected magnitude of the MOR is a function of the reasons for differences in practice as well as the scope of the comparison (e.g. comparing all hospitals vs. comparing similar teaching hospitals). In recent comparisons of the use of various breast surgeries across physicians, the MOR values ranged from 1.59 to 2.3 (McCahill et al. JAMA 2013 PMID 22298678; Wang et al. Breast Cancer Res Treat 2013 PMID 23354364; Feigelson et al. J Am Coll Surg 2013 PMID 23490543). The most meaningful interpretation of the MOR's magnitude can however be gained from comparing the MOR directly with the ORs for the fixed-effects in the same study (e.g. age, sex, comorbidity level, etc.). This point has been added to the manuscript.

Reviewer: 2
Comments to the Author

General Comments:

This is an interesting population-based study examining practice differences between hospitals with respect to early versus delayed cholecystectomy. The authors examine several covariates using multivariable logistic regression models including patient and hospital characteristics to identify predictors of what may explain practice variation.

1. The authors neglect to examine surgeon characteristics, which practically speaking may be the most important variable in determining whether patients undergo early versus delayed cholecystectomy. Given the fact that cholecystectomy in an emergent setting requires a certain degree of comfort and confidence (on the part of the surgeon) with respect to operating in non-ideal circumstances, surgeon practice preference could quite possibly account for the variation in patient care among the hospitals.
We agree that an understanding of the important physician-level factors that impact on variation would be informative. As mentioned in the limitations section, we were unable to identify the decision-making surgeon/physician. We have added a suggestion that exploring surgeon-level influences on practice should be the focus of future research.
Understanding variation at the hospital level remains a constructive starting point to generate solutions for quality improvement.

2. The authors seem almost patronizing that the "best available evidence" suggests early cholecystectomy is preferable to delayed cholecystectomy when those studies may not be generalizable to the authors' diverse study population. I would caution the authors to use more neutral language.
Thank you for the suggestion. We agree that the issue remains contentious. However, we do feel that based on existing evidence and guidelines, early intervention can be justified for most patients without severe cholecystitis. The wording in the manuscript has been reviewed, the importance of surgical expertise has been acknowledged and the following section has been added to the discussion detailing the major limitations of current knowledge: "In addition, variation in practice may in part reflect the need to address the remaining gaps in evidence comparing the outcomes of early and delayed cholecystectomy. In fact, one of the factors hindering the uptake of early cholecystectomy may be concern that early intervention is associated with a higher rate of major bile duct injury, a rare but devastating operative complication. Adequately powered studies assessing whether this is true, as well as a comparison of real world rates of conversion from laparoscopic to open cholecystectomy are required."

Major Revisions:

3. As stated above, it would be interesting to evaluate surgeon characteristics as a potential variable explaining practice differences in early versus delayed cholecystitis. Ultimately, the decision on operative timing is a highly complex one that must take into account hospital resources, acuity of patient illness, and surgeon experience with emergent cases that may be

	best managed in the initial period with a trial of non-operative management. The remaining 47% of unexplained variability from this paper might be answered by looking at this particular issue. We agree with the reviewer's comments and additions have been made to the manuscript. **4.** The authors use several surrogate markers: (1) appendectomies performed at a hospital as a surrogate marker of presence of a general surgeon, (2) ICU admission and percutaneous cholecystostomy as a surrogate of severe acute cholecystitis, and (3) income based on postal code demographics as a surrogate for socioeconomic status. There are issues with using these markers as surrogates: (1) in very remote areas of Ontario, appendectomies may actually be performed by general practitioners who have extra surgical training, which may bias the results toward more cases of delayed cholecystectomy; (2) ICU admission and percutaneous cholecystostomy tube are not necessarily surrogate markers of severe cholecystitis, in fact, they may be markers of non-surgical disease (e.g. acalculous cholecystitis, which is often associated with critically ill patients who do not actually require cholecystectomy) or surgeon preference (e.g. a surgeon who prefers delayed cholecystectomy and temporizes patients with percutaneous cholecystostomy tubes); (3) income is also a very crude marker of socioeconomic status, which encompasses so many more complex variables and when based on postal code census data, hardly qualifies as a truly precise or accurate measure. (1) The extent of any potential bias would likely be minimal given the small number of patients managed at this type of remote hospital. (2) As you suggest, cholecystostomy drain placement or ICU admission may reflect a range of patient types including those with severe cholecystitis, serious comorbidity precluding operative intervention or acalculous cholecystitis. Furthermore, the vast majority of patients with uncomplicated calculous cholecystitis receiving a cholecystostomy drain were likely considered at high perioperative risk for urgent early surgery. All of these patients should reasonably be excluded. The description in the methods has been reviewed and the wording in Figure 1 has been changed. (3) We agree that neighborhood income level is a crude marker of socioeconomic status, as stated in the methods. Nevertheless its use is supported by the fact that, in other clinical contexts, differences in health resource utilization and outcomes have been found across income levels in Ontario: o Kapral et al. *Neighborhood income and stroke care and outcomes.* Neurology 2012 – Pubmed ID 22895592, o Booth et al. *Universal Drug Coverage and Socioeconomic Disparities in Major Diabetes Outcomes.* Diabetes Care 2012 – Pubmed ID 22891257 o Booth et al. *The impact of socioeconomic status on stage of cancer at diagnosis and survival.* Cancer 2010 – Pubmed ID 20681012 o Liu et al. *Social disparity and the use of intrapartum epidural analgesia in a publicly funded health care system.* Am J Obstet Gynecol 2010 – Pubmed ID 20045506 **5.** The discussion appears to imply early cholecystectomy for cholecystitis as the standard of care, which is inaccurate. In fact, at the most recent Canadian Association of General Surgeons forum in Calgary (September 2012), there was still ongoing plenary debates about early versus delayed cholecystectomy. The fact is that there is variation in practice because there is no national consensus on what truly is better. The authors ought to amend their discussion to appear less subjective in order to more accurately discuss the results which are that there is still a huge variation in care for patients with acute cholecystitis and that patients who do not receive early cholecystectomy are not deviations from standard of care. The standard of care for this common acute surgical problem is still evolving, as is the concept of an acute surgical care service model for hospitals. Thank you for the suggestion. Revisions have been made to the wording throughout the manuscript and a section added to the discussion. **Minor Revisions:** **6.** The abstract should define the abbreviation AC (acute cholecystitis) prior to its ongoing use. The abbreviation has been removed throughout the manuscript as per the editors' request.