The long-term effect of bile duct injuries on health-related quality of life: a meta-analysis

Matthew P. Landman1, Irene D. Feurer1,2, Derek E. Moore1, Victor Zaydfudim3 & C. Wright Pinson1

Departments of 1Surgery and 2Biostatistics, Vanderbilt University Medical Center, Nashville, TN, and 3Division of Gastroenterologic and General Surgery, Mayo Clinic, Rochester, MN, USA

Abstract

Background: The reported effects of biliary injury on health-related quality of life (HRQOL) have varied widely. Meta-analysis methodology was applied to examine the collective findings of the long-term effect of bile duct injury (BDI) on HRQOL.

Methods: A comprehensive literature search was conducted in March, 2012. Because the HRQOL surveys differed among reports, BDI and uncomplicated laparoscopic cholecystectomy (LC) groups’ HRQOL scores were expressed as effect sizes (ES) in relation to a common, general population, standard. A negative ES indicated a reduced HRQOL, with a substantive reduction defined as an ES ≥ -0.50. Weighted logistic regression tested the effects of BDI (versus LC) and follow-up time on whether physical and mental HRQOL were substantively reduced.

Results: Data were abstracted from six publications, which encompass all reports of HRQOL after BDI in the current, peer-reviewed literature. The analytic database comprised 90 ES computations representing 831 patients and 11 unique study groups (six BDI and five LC). After controlling for follow-up time (P = 0.001), BDI patients were more likely to have reduced long-term mental [odds ratio (OR) = 38.42, 95% confidence interval (CI) = 19.14–77.10; P < 0.001] but not physical (P = 0.993) HRQOL compared with LC patients.

Discussion: This meta-analysis of findings from six peer-review reports indicates that, in comparison to LC, there is a long-term detrimental effect of BDI on mental HRQOL.

Introduction

Population-based studies indicate that the rate of bile duct injury (BDI) after laparoscopic cholecystectomy (LC) range from 0.4% to 0.7%, a significant increase in comparison to open cholecystectomy.1–3 Mortality associated with these injuries has been estimated to be two- to three-fold higher than with an uncomplicated cholecystectomy.3,4 Notably, bile duct injuries result in significant long-term morbidity, disability, healthcare utilization, litigation and, ultimately, cost.5–8

The adverse clinical outcomes associated with BDI are well documented. Assessing health-related quality of life (HRQOL), a patient-reported outcome provides a way to quantify the impact of biliary injury during a cholecystectomy from the patient’s perspective. However, studies of HRQOL in patients with BDI continue to result in mixed findings and a clear consensus of the long-term effect of BDI on HRQOL has not emerged in the literature. In 2001, Boerma et al. published the original report examining the impact on HRQOL of BDI occurring during a laparoscopic cholecystectomy.9 Using the Short Form 36 Health

Funding: This study was supported in part by the Institutional National Research Service Award T32 HS 013833 from the Agency for Healthcare Research and Quality, US Department of Health and Human Services, and by an educational award from the Novartis Corporation.

This manuscript was presented in part during the Concurrent Sessions of the Asian Pacific HPBA 3rd Biennial Congress, Melbourne, Australia, 27–30th September, 2011.
Survey® (SF-36), they found worse physical and mental HRQOL after BDI compared with an uncomplicated LC and average general population values.\(^9\) These findings were replicated by the same group in 2008 in an expanded sample and in a subset of the original patients for whom there was longitudinal HRQOL data.\(^11\)

In 2004, Moore et al. also reported worse SF-36 physical and mental HRQOL in BDI patients compared with LC patients.\(^12\) However, Melton et al., in a 2002 study that used a modified City of Hope HRQOL assessment tool, reported worse mental HRQOL in BDI patients but no effect of BDI on physical and social HRQOL compared with a healthy comparison group.\(^13\) Other reports found no differences in HRQOL between BDI and comparison cohorts. In 2004, Sarmiento et al.\(^14\) reported no difference in HRQOL, after a minimum of 5 years, between patients who had surgical reconstruction after BDI and those undergoing a LC. In the most recent report, published in 2009, Hogan et al.\(^15\) concluded that the ‘quality of life of surviving patients after BDI compares favourably with that of an uncomplicated laparoscopic cholecystectomy’. Given the notable variability in HRQOL outcomes reported in the literature, we conducted a meta-analysis to determine whether BDI and LC patients differed in the likelihood of having substantively diminished HRQOL in relation to a common standard that was applicable across all studies.

### Methods

#### Study identification and incorporation

A literature search for relevant publications in PubMed was conducted in March, 2012. Search terms included ‘quality of life’ and ‘bile duct injury’. Study inclusion criteria were: (i) English language publications in the peer-reviewed medical literature; (ii) reports of quantitative HRQOL data in BDI patients, with or without comparison to uncomplicated LC patients; and (iii) the use of generic or condition-specific HRQOL surveys that could be referenced to either general population standards and/or to a healthy reference group. All references from identified studies were examined closely for additional reports that met inclusion criteria. Summary data were extracted and independently verified by two investigators (M.P.L. and I.D.F.) and any initial discrepancies in study data encoding were identified and resolved. This study was reviewed and considered exempt by the Vanderbilt University Medical Center Institutional Review Board.

Analyses were performed in accordance with the guidelines set forth by Stroup et al. for meta-analyses of observational studies,\(^16\) with a few specific exceptions. Given that the number of studies identified in the comprehensive literature review was small and that all available studies were included in the meta-analysis, objective classification of study quality was not performed. While a quantitative index of study heterogeneity was not developed, Table 1 illustrates the heterogeneity of the study findings. All of the studies were performed at tertiary referral centres which, given that less than 20% of all BDIs are repaired by the primary surgeon,\(^4\) suggests that the patients included in these reports are a representative majority of BDI patients from each centre’s referral area. Subgroup analyses were not performed in consideration of statistical power and the limited number of studies available for analysis.

#### Data abstraction and database development

All studies that met inclusion criteria were abstracted. Unique study groups were characterized from each individual study as either BDI or LC. The sample size and follow-up time of each unique study group were recorded. Individual reported HRQOL scales and summary scores were characterized as representing either physical or mental/psychosocial HRQOL on the basis of accepted classifications for each survey.

The heterogeneity of HRQOL survey tools and follow-up time, the between-study variability in how and when HRQOL was measured and recorded, precluded direct comparison of pooled findings between BDI and LC subjects. Between-study measurement heterogeneity was addressed in the meta-analysis by referencing all reported HRQOL summary data in relation to a common standard: the effect size (ES) or standardized difference [in standard deviation (SD) units] between the reported HRQOL scores and the general population standard or healthy reference.
group for the survey. Details of ES, standard error (SE) and observation weight computations are described in the Appendix. Follow-up time at HRQOL determination was referenced to the cholecystectomy procedure.

Effect sizes were computed for each individual reported HRQOL scale and/or summary component score. For example, if a study reported HRQOL data for BDI and LC groups as the eight individual scales and two summary components of the SF-36, that study’s data were represented in the analytic database by 20 individual ES observations (10 BDI and 10 LC ESs). Then each ES observation was classified as representing either physical HRQOL (the physical functioning, role physical, bodily pain, and general health scales; physical component summary score) or mental HRQOL (the vitality, social functioning, role emotional, and mental health scales; mental component summary score). The City of Hope Medical Center Quality of Life Survey is scored as three domains: physical, psychological and social. For the purpose of this study, the psychological and social domains were classified as mental HRQOL, and the physical domain was classified as physical HRQOL.

In order to understand findings from the perspectives of both statistical significance and potential clinical relevance, effect sizes were expressed in relation to an accepted definition of a substantive difference for HRQOL measures. An ES $\leq -0.50$ SD, or HRQOL being at least a half SD below a general population reference standard as indicated by an ED $\leq -0.50$ SD, was considered substantively reduced HRQOL. Given the relatively small number of studies, this approach resulted in an analytic database of 90 individual effect size computations that ensured an optimal statistical power (a sample to covariate ratio of $\approx 15:1$, power $= 0.81$ to detect an incremental $R^2$ of 0.22) for the multivariate analyses.

Statistical analysis
Multivariable weighted logistic regression models were used to test the effects of the cholecystectomy group (BDI versus LC), controlling for follow-up time after a cholecystectomy, on the likelihood of physical and mental HRQOL being substantively reduced in relation to general population/healthy sample reference data (as defined by an ES being $\leq -0.50$ SD). All statistical analyses were carried out using SPSS version 19 (IBM Corp., Armonk, NY, USA).

Results
The literature search yielded 39 studies. Of these, 33 were excluded for the following reasons: 1 was published in abstract form only, 5 studies were not in the English language (none of which reported HRQOL), 7 were review articles and the remaining 20 studies either did not mention a cholecystectomy ($n = 7$) or were limited to non-HRQOL outcomes after BDI ($n = 13$) (Fig. 1). Six studies met all inclusion criteria and represent all reports in the literature that present primarily-collected HRQOL data in BDI patients as of March 2012. Data these 6 studies represent 521 patients with BDI (6 unique groups, 1 per study) and 310 patients who underwent LC (5 unique groups). A sixth unique LC group could not be included in these analyses because the findings reported by Boerma et al., which were corrected in a 2007 Letter to the Editor, did not include sufficient information to permit accurate computation of the SEs of the effect sizes or weights for the LC group (Table 1).

Meta-analysis
Because follow-up time at HRQOL determination varied substantially between studies, and sometimes between LC and BDI groups within a given study, follow-up time was treated as an ordinal covariate, with category thresholds determined on the basis of the range of reported average values across all BDI and LC groups (approximately 2 to 20 years). This resulted in three follow-up time categories: $\leq$5 years, 6 to 8 years and $>$8 years after BDI or LC. The number of unique groups and patients by follow-up time period is summarized in Table 2. There were 90 HRQOL scale-level effect sizes in the analytic database. Forty-nine effect size observations corresponded to BDI patient groups and 41 corresponded to LC patient groups (Table 3). Within the BDI groups, 24 effect sizes corresponded to physical HRQOL and 25 corresponded to mental HRQOL. Among the LC groups, 20 effect sizes corresponded to physical HRQOL and 21 corresponded to mental HRQOL. The 90 individual standardized effect size computations are summarized by HRQOL scale and study in Fig. 2. Bars represent (in SD units) the standardized difference between reported group data and the general population standards for the particular HRQOL scale (see Appendix, equation 1). Bars that fall at or below the dashed line ($\leq -0.50$ SD) indicate observations where HRQOL was coded as being substantively reduced for the group. Bars falling above the dashed line represent observations where HRQOL was coded as not being substantively reduced.

The first weighted logistic regression model tested the effects of BDI (versus LC) and follow-up time on the likelihood of physical HRQOL being substantively reduced (Table 4). The overall model and main effect of follow-up time category on HRQOL were statistically significant. Patients who were followed, on average, for greater than 8 years have a 72% lesser likelihood of having a reduced physical HRQOL [odds ratio (OR) $= 0.28$, 95% confidence interval (CI) $= 0.14$–0.56, $P < 0.001$]. After controlling for follow-up time, BDI patients were not more likely to have a reduced physical HRQOL than LC patients ($P = 0.993$). The second weighted logistic regression model tested the effects of BDI (versus LC) and follow-up time on the likelihood of mental HRQOL being substantively reduced (Table 4). The overall model and the effects of follow-up time and BDI were statistically significant. Patients who were followed, on average, for more than 5 years after a cholecystectomy were significantly less likely ($P < 0.001$) to have a reduced mental HRQOL. After controlling for follow-up time, BDI patients were about 38 times more likely to have a reduced mental HRQOL than LC patients (OR $= 38.42$, 95% CI $= 19.14$–77.10, $P < 0.001$).
The Nagelkerke $R^2$ value, which can range from 0 to 1, is a measure of overall strength of association for logistic regression models and represents the proportion of variance explained by a model. Both models had substantive $R^2$ values and non-chance overall $P$-values (both model $P < 0.001$). Importantly, this does not imply that every statistical effect (BDI versus LC, follow-up time, and the individual follow-up time categories) in each model was statistically significant. The $P$-values for individual model
effects indicate which variables are associated with either an increased or decreased likelihood of having a reduced HRQOL. ORs that are greater than 1.00, with CIs that do not span 1.00, indicate variables (or categories) that are independently associated with an increased likelihood of having a reduced HRQOL. Those that are less than 1.00, with CIs that do not span 1.00, indicate variables (or categories) that are associated with a reduced likelihood of having a decreased HRQOL.

The P-values, ORs and CIs for individual model effects lead to the interpretation that, after controlling for follow-up time (which was statistically significant in both models), the effect of BDI was not significant in the model of reduced physical HRQOL but BDI was associated with a significantly increased likelihood of mental HRQOL being substantively reduced. The finding that the OR for the effect of BDI (versus LC) in the physical HRQOL model was not statistically significant (P = 0.993) means that the estimate is completely chance and it should not be interpreted as being meaningful in this model. Similarly, both models indicate that follow-up time is related to a reduced likelihood of HRQOL being reduced, with the effect of follow-up being significant at greater than 8 years (compared with <5 years) in the physical HRQOL model and at both 6 to 8 and greater than 8 years in the mental HRQOL model.

**Discussion**

The effect of BDI on long-term HRQOL has not been well established. Some previous studies found no difference in HRQOL between patients who had a BDI during a laparoscopic cholecystectomy and those who had an uneventful laparoscopic cholecystectomy. Other studies demonstrate a significant long-term reduction in both physical and mental HRQOL after BDI. We used meta-analysis to consolidate data from these reports and test the long-term effect of biliary injury during a laparoscopic cholecystectomy on HRQOL.

While all six studies included in the meta-analysis examined HRQOL in patients treated for a BDI and compared them with either patients having an uncomplicated LC or healthy/normal subjects, a number of methodological differences prevented the direct comparison of pooled BDI and LC groups’ HRQOL scores and necessitated a standardized outcome measure. First, the studies reported scores from several different HRQOL surveys or versions of a survey. Second, while studies differed in the number of HRQOL scales that were reported, our approach enabled the use of all reported HRQOL data that could be referenced to a general population standard. Third, because follow-up time was not consistent between, or in some instances within, studies, a direct comparison of LC and BDI scores would be impossible to interpret. Fourth, while unlikely, it would be theoretically possible for HRQOL scores after LC and BDI not to differ significantly from each other when both groups’ scores were substantively below the general population. Fifth, in order to understand findings from the perspectives of both statistical significance and clinical relevance, effect sizes were evaluated in relation to an accepted definition of a substantive difference for HRQOL measures. Finally, given the relatively small number of primary studies, our approach resulted in an analytic database of 90 individual effect size computations that ensured optimal statistical power.

Clinical outcomes after iatrogenic BDI are well documented. In their series of 200 patients who underwent surgical reconstruction for BDI, Sicklick et al. found that the mean time from BDI to definitive care was 42 weeks. Additionally, nearly 50% of these patients experienced a post-operative morbidity ranging from wound infection to a biliary leak. Walsh et al. noted a 40% overall morbidity rate in their retrospective cohort. At a mean follow-up of 67 months, 11% of patients had developed a biliary stricture. Sahajpal et al. found a rate of 14% for long-term biliary strictures. Post-operative biliary leaks and strictures often lead to subsequent re-interventions that may be associated with patients’ perceptions of HRQOL during the first several years after injury. A prolonged treatment course and additional interventions lead to higher costs and could lead some patients to pursue litigation, which could impact perceptions of their injuries and HRQOL.

Moore et al. reported that the 22% of BDI patients in their cohort who reported filing a lawsuit had significantly lower physical and mental component scores of the SF-36 compared with those BDI patients who did not pursue litigation. A similar reduction in HRQOL in patients filing malpractice claims was noted by de Reuver et al. In an evaluation of the cost of bile duct injuries, Andersson et al. noted that the primary cost of major bile duct injuries was driven by the loss of production (time out of work).
Figure 2 The 90 individual effect size observations that comprised the meta-analytic database are depicted by primary publication and health-related quality of life (HRQOL) scale. These data illustrate the variability of HRQOL scales employed and the heterogeneity of outcomes reported across these studies. Observations that fall at or below the dashed line [effect sizes (ES) ≤ −0.5] were classified in the logistic regression models as representing substantively reduced HRQOL and those that fall above the dashed line were classified as representing HRQOL that was not reduced. Abbreviations for the Short Form 36 Health Survey® (SF-36) scales and components are: PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health; PCS, physical component summary; MCS, mental component summary. The City of Hope (COH) scales are: phys, physical; psy, psychological; social, social. The psychological and social scales were considered mental HRQOL for the purpose of these analyses.
Table 4 Weighted logistic regression models of physical and mental HRQOL

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Outcome = reduced physical HRQOL</th>
<th>Outcome = reduced mental HRQOL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model $R^2 = 0.44$ ($P &lt; 0.001$)</td>
<td>Model $R^2 = 0.41$ ($P &lt; 0.001$)</td>
</tr>
<tr>
<td></td>
<td>($n = 44$ effect sizes)</td>
<td>($n = 46$ effect sizes)</td>
</tr>
<tr>
<td>BDI (ref. LC)</td>
<td>OR $= 3.95 \times 10^4$ P-value $= 0.993$ 95% CI $= 0.00 : -$</td>
<td>OR $= 38.42$ P-value $&lt; 0.001$ 95% CI $= 19.14 : 77.10$</td>
</tr>
<tr>
<td>Follow-up time</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>6 to 8 years (ref. $\leq 5$ years)</td>
<td>0.00</td>
<td>0.992</td>
</tr>
<tr>
<td>$&gt;8$ years (ref. $\leq 5$ years)</td>
<td>0.28</td>
<td>$&lt; 0.001$</td>
</tr>
</tbody>
</table>

BDI, bile duct injury; LC, laparoscopic cholecystectomy; HRQOL, health-related quality of life; OR, odds ratio.

Conflicts of interest
None declared.

References

for the patient. These monetary costs, which primarily affect the patient and their family, can be substantive and protracted and could affect the way they perceive HRQOL, particularly mental HRQOL, after BDI. This meta-analysis demonstrates the long-term detrimental effect of BDI on mental HRQOL. This finding differs from those of Sarmiento et al.14 and Hogan et al.15 who reported no effect of BDI on either physical or mental HRQOL, and is consistent with the findings of Melton et al.13 who reported a detriment in mental HRQOL only. The detrimental effects on mental HRQOL described by Boerma et al., Moore et al. and de Reuver et al.9,11,12 are supported by this meta-analysis. However, after controlling for follow-up time, we did not find BDI to be associated with a substantively reduced long-term physical HRQOL. These disparate findings may be related to differences between studies and individual BDI and LC groups in follow-up time and the way follow-up time was indexed (whether to the original cholecystectomy or presentation for repair of a biliary injury). While the individual statistical effects of follow-up time category differed in the physical and mental HRQOL models, the overall likelihood of HRQOL being substantively reduced declines over time in LC and BDI patients. Variations in reported associations between BDI and long-term physical and mental HRQOL may be related to differences in survey tools, follow-up time, and differences in patient referral patterns and sampling. Additionally, variability in clinical management protocols such as the choice of endoscopic or surgical management and the timing of repair after BDI may also contribute to differences in patient-reported outcomes such as HRQOL.

Ultimately, while the endoscopic management and surgical reconstruction of BDI are generally successful, patients’ perceptions of their long-term HRQOL may not be without consequence. While several individual studies have reported detriments in physical and/or mental health-related quality of life after a bile duct injury, after controlling for follow-up time that spans over a decade, this meta-analysis indicates that patients who experience bile duct injuries report detriments in mental quality of life compared with the general population or healthy samples. These findings illustrate the importance of addressing both technical and patient-reported outcomes such as health-related quality of life in the immediate post-operative period and beyond.


Appendix: Computation of effect sizes, standard errors and weights

Effect sizes were computed as the standardized difference between reported means of: (i) BDI patients and general population values and (ii) LC patients and general population values (equation 1). Negative ESs indicated a diminished HRQOL, with a substantive reduction defined as an ES $\leq -0.50$ (at least half a standard deviation below general population average values).

Equation 1: Computation of effect size (standardized mean difference)

$$ES = (\text{group average} - \text{general population average}) / \text{pooled standard deviation}$$

The SE of the effect size was calculated using previously described methods (equation 2).

Equation 2: Computation of the standard error of the effect size

$$\text{Standard Error}_{ES} = \left\{ \left( \frac{N_{BDI \ or \ LC \ group} + N_{general \ population \ sample}}{N_{BDI \ or \ LC \ group} \times N_{general \ population \ sample}} \right)^{1/2} \right\}^{1/2}$$

To account for imprecision in ES computations resulting from differences in sample sizes, the inverse variance was calculated and used to weight each ES (equation 3).

Equation 3: Computation of inverse standard error for weighting effect sizes observations

$$\text{General Weight}_{ES} = \frac{1}{\text{Standard Error}_{ES}}$$

To adjust for within-group non-independence of effect sizes when multiple HRQOL scales were reported for a BDI or ULC group, ES observations were also weighted by the inverse of the number of physical and/or mental HRQOL domains reported for the group (equation 4).

Equation 4: Computation of final weights for physical and mental HRQOL ES observations

$$\text{Final Weight}_{ES-physical} = \frac{\text{General Weight}_{ES}}{\text{number of physical HRQOL scales and summary scores reported for the unique group}}$$

$$\text{Final Weight}_{ES-mental} = \frac{\text{General Weight}_{ES}}{\text{number of mental HRQOL scales and summary scores reported for the unique group}}$$