

Appropriateness variation in cholecystectomy

JOSÉ M. QUINTANA, JOSÉ CABRIADA, IGNACIO LÓPEZ DE TEJADA, LUIS PERDIGO, INMACULADA ARÓSTEGUI, AMAIA BILBAO, IDOIA GARAY *

Background: Cholecystectomy use varies moderately across regions and countries. The purpose of this study was to evaluate the appropriateness of the use of cholecystectomy in patients with nonmalignant diseases in different hospitals using explicit criteria developed by an expert panel. **Methods:** Patients on waiting lists to undergo a cholecystectomy for nonmalignant diseases in six public hospitals in Spain were included in this prospective observational study over a 1-year period. Appropriateness criteria were first developed by a panel of experts using the RAND appropriateness method. The appropriateness of the indication was judged by applying explicit criteria developed previously. Complications were recorded 3 months after surgery. **Results:** After evaluation of 960 patients, 7 (0.7%) were considered to have undergone inappropriate procedures and 76 (7.9%) were judged to be uncertain. Differences were found in the rate of appropriateness among some centres (inappropriateness rate ranging from 0 to 2.6%). No differences were found among the appropriateness categories for length of stay or complications, mortality, or readmission up to the third month after discharge. **Conclusions:** This study identified a low percentage of inappropriately performed cholecystectomies in our area. Even so, appropriateness differences among hospitals were found.

Keywords: cholecystectomy, health services research, hospitals, physician's practice patterns, utilization review

Cholecystectomy is a widely performed procedure. The use of cholecystectomy is still increasing in most developed countries,^{1,2} but surgical rates vary moderately across regions and countries,^{3,4} a finding not explained solely by differences in the prevalence of gallbladder disease. Variations in clinical decision making may also contribute. In relation to this, the appropriateness of use of medical procedures has been seen as an important element in quality of care.⁵

Central to this investigation is the determination of what constitutes appropriate indications for a given procedure. A method that combines expert opinion with available scientific evidence was developed by investigators at the RAND Corporation and the University of California at Los Angeles.⁶ This method has been used to evaluate the appropriateness of a variety of medical and surgical interventions, including cholecystectomy. Previous RAND methodology-based studies of this procedure were performed in the early 1980s in the United States,⁷ and in the late 1980s in Israel⁸ and the United Kingdom.⁹ The goal of those studies was to provide managers and clinicians with information that would help them decide whether or not cholecystectomy was appropriate for a specific patient. Ten to 20 years have passed since those panels convened, and important changes have occurred in this field. The introduction of echography and other imaging tests has improved the diagnosis of patients with gallbladder disease or digestive symptoms. At the same time, new pharmacologic or invasive modalities, such as endoscopic retrograde cholangiopancreatography (ERCP), or surgical treatments, such as laparoscopy, have changed the indications for the intervention. Therefore, the criteria developed in the 1980s are no longer useful. In our criteria, we introduced variables such as the use of ERCP or new imaging tests such as echography not included in

those previous studies. Our algorithm tried to incorporate new evidence and new scenarios to present updated explicit criteria of the indication for cholecystectomy. A new field study applying those updated criteria was found necessary.

The goal of this study was to apply explicit criteria, developed using a multidisciplinary approach, to examine the appropriateness of the indications for cholecystectomy in patients with nonmalignant diseases in various hospitals in Spain.

METHODS

Explicit criteria development

The criteria for measuring the appropriateness of the use of cholecystectomy were developed according to a previously described explicit method,⁶ i.e., the RAND appropriateness method, which consists of the following steps.

First, an extensive literature review was performed to summarize existing knowledge concerning efficacy, effectiveness, risks, costs, and opinions about the use of cholecystectomy to treat nonmalignant diseases.

Second, from this review, a comprehensive and detailed list of mutually exclusive and clinically specific scenarios (indications) was developed in which cholecystectomy might be performed. This list contained 414 indications for cholecystectomy. Each indication was specified in sufficient detail that patients within a given indication were reasonably homogeneous. A detailed description of the variables and their categories and definitions that were considered to create the indications has been published elsewhere.¹⁰ Variables included were patient age, surgical risk (measured by the ASA),¹¹ diagnosis (symptomatic cholelithiasis without complications or biliary colic single or multiple; with complications, as cholecystitis, choledocholithiasis, pancreatitis, or cholangitis; asymptomatic cholelithiasis; and a miscellaneous category that included imaging findings, porcelain gallbladder, polyps, and cholesterosis), gallbladder and common bile duct imaging studies, previous non-surgical procedures such as ERCP performed (successful or not) or not, and the presence of special circumstances, such as diabetes.

Third, we compiled a national panel of six experienced surgeons and six gastroenterologists. Panelists were nationally recognized specialists whose names were provided by their respective medical societies and members of our research team. The panelists were provided with the literature review and the list of indications, and they rated each indication for the appropriateness of performing cholecystectomy, considering the

* J.M. Quintana¹, J. Cabriada², I. López de Tejada³, L. Perdigo⁴, I. Aróstegui⁵, A. Bilbao¹, I. Garay¹

1 Unidad de Investigación, Hospital de Galdakao, Galdakao, Vizcaya, Spain

2 Servicio de Digestivo, Hospital de Galdakao, Galdakao, Vizcaya, Spain

3 Servicio de Cirugía, Hospital de San Eloy, Vizcaya, Spain

4 Servicio de Cirugía, Hospital de Cruces, Bilbao, Vizcaya, Spain

5 Departamento de Matemática Aplicada, Estadística e Investigación

Operativa, Universidad del País Vasco, Lejona, Vizcaya, Spain

Correspondence: J.M. Quintana, MD, PhD, Unidad de Investigación, Hospital de Galdakao, Barrio Labeaga s/n, 48960 Galdakao, Vizcaya, Spain, tel. +34 94 400 7105, fax +34 94 400 7132,

e-mail: jmquinta@hgda.osakidetza.net

average patient and average physician in the year 2000. Appropriateness was defined as meaning that the expected health benefit exceeds the expected negative consequences by a sufficiently wide margin to make cholecystectomy worth performing, choosing the best surgical alternative available for the patient (open cholecystectomy, mini-cholecystectomy, or laparoscopy).

Ratings were scored on a 9-point scale. The use of cholecystectomy for a specific indication was considered appropriate if the panel's median rating was between 7 and 9 without disagreement, inappropriate if the value was between 1 and 3 without disagreement, or uncertain if the median rating was between 4 and 6 or if the members of the panel disagreed. Disagreement was defined as occurring when at least one third of the panelists rated an indication from 1 to 3 and at least another third rated it from 7 to 9. Agreement if no more than three panelists rated the indication outside the 3-point region^{1-3,4,6,7-9} containing the median; and indeterminate if neither agreement nor disagreement was found. This method did not attempt to force panelists to reach agreement on appropriateness.

It was beyond the scope of the work of our panel to compare the use of open cholecystectomy with laparoscopic cholecystectomy. The panelists were instructed to evaluate the appropriateness of performing any cholecystectomy technique against other non-surgical treatments, taking a watch-and-wait strategy, or doing nothing. The panelists were instructed not to evaluate the appropriate timing of the intervention (urgent or as an interval operation), but only if the intervention itself was appropriate.

The ratings were confidential and took place in two rounds, using a modified Delphi process. The first round was performed before the panel meeting. The results were collated and presented to the panelists at a 1-day meeting. Each panelist received the anonymous ratings of the other panelists as well as a reminder of his own ratings. After extensive discussion, panelists revised the indications according to the definitions. During the first round, each panelist rated 390 indications, and during the second round 414 indications, because a new age category was added in some cases. Results of part of the work of the panel of experts were included in a previous publication.¹⁰

Data collection

This prospective observational study took place in six public hospitals (four university-affiliated and two community-based). All the hospitals belonged to the Basque Health Service-Osakidetza, a local government agency in the Basque Country, and to the Spanish National Health Service. The identities of the hospitals and surgeons were not revealed in the research reports. Physicians in each hospital were blinded to the study goals.

The recruitment of patients on the waiting list started in March 1999 and ended in March 2000. All patients undergoing cholecystectomy, who were followed in any of the six hospitals, were included in the study. Patients with malignant, severe organic, or psychiatric diseases, were excluded. A total of 1,009 patients were placed on the waiting list during that time; nine were excluded because they had malignant diseases. Of 1,000 patients who fulfilled the selection criteria, 96.3% of the medical records were accessible and reviewed. Of the 963, three presented with immunodepression (transplant, human immunodeficiency virus) and were excluded from this study.

To collect data and determine appropriateness, we developed a computerized algorithm based on the results of our panel. We also developed data collection questionnaires that included variables before the intervention, admission, and discharge, including the intervention, and complications 3 months after discharge. Besides those variables belonging to the appropriateness algorithm, which were mentioned previously, other

variables collected included sociodemographic data, height, weight, main complaint, 12 comorbidities (diabetes, hypertension, cardiac disease, dementia, depression, chronic obstructive pulmonary disease, stroke, cancer, renal disease, hepatic disease, anemia, and arthritis), previous interventions, intervention characteristics, local and general complications peri- and post-intervention, reintervention, death, and length of hospital stay.

Three trainee reviewers collected the data from the medical record via a standardized questionnaire. The reviewers were blinded to the specific study goals. Three months after discharge, all medical records were again reviewed to determine if the patient had been readmitted, died, or had any complication resulting from the intervention.

Statistical analysis

The unit of study was the patient. Descriptive statistics, frequency tables, mean and standard deviations were included. Chi-square and Fisher's exact tests were used to test for statistical significance among proportions. For continuous variables (e.g. age), the ANOVA test was performed in the univariate analysis. For the logistic regression model, the dependent variable was dichotomized as appropriate intervention versus uncertain or inappropriate. The odds ratios (ORs) of a less-than-appropriate intervention (uncertain and inappropriate) were calculated for each of the hospitals, taking as the reference the one with the lowest rate of uncertain-inappropriate cases (hospital 5). Hospital was the only independent variable, since all other relevant adjustments are already included in the appropriateness definition.

Kappa statistics were calculated to test agreement between reviewers and research team members for some variables (diagnosis, comorbidities, surgical risk and ERCP performed). All effects are significant at $p < 0.05$, unless otherwise noted. All statistical analysis were performed using the SAS for Windows, version 8.0, statistical software.

RESULTS

During the 1-year period, 960 patients were recruited in the six participating hospitals. Ninety-three (9.7%) of them had diabetes and 867 (90.3%) no other comorbidity that directly influenced the decision about performing a cholecystectomy. Most patients (85.8%) presented with uncomplicated cholelithiasis (table 1).

Table 1 Principal sociodemographic and clinical variables

Variable	N n=960	%
Age (mean, SD)	57.9	14.4
Gender: Women	647	67.4
Diagnostic		
Indications in non-diabetics	867	90.3
Symptomatic cholelithiasis	744	85.8
Uncomplicated	528	71
Complicated	216	29
Asymptomatic cholelithiasis	109	12.6
Other indications ^a	14	1.6
Indications in diabetics	93	9.7
Symptomatic cholelithiasis	72	77.4
Uncomplicated	45	62.5
Complicated	27	37.5
Asymptomatic cholelithiasis	21	22.6

a: Other indications category includes: porcelain gallbladder, gallbladder polyps, or cholesterosis.
SD: standard deviation

After applying the explicit criteria to the 960 interventions, we found that 91.4% of the cases were considered appropriate. The indication was judged uncertain in 7.9% of the interventions, and 0.7% of all were inappropriately treated. Of the total 414 theoretical indications scored by the expert panel, only 59 (14.3%) appeared in the patient results.

Patients younger than 75 years old were found equally in the appropriate group (91.3%) than those older than 74 (92.1%). Patients with ASA IV were more likely to be considered inappropriate (7.7%) or uncertain (30.8%) than those with ASA I to III. In addition, differences were found among diagnostic groups, i.e., those with asymptomatic cholelithiasis were considered more frequently to have undergone inappropriate treatment (5.4%) than the other groups (table 2). The most frequently encountered indications and their appropriateness evaluation are shown in table 3. Those asymptomatic were the most common inappropriate indication, and, also, the only one asymptomatic patient classified as ASA IV.

There were no statistically significant differences among the participating hospitals in patient age, sex, recorded comorbidities, and type of intervention. There were differences among them in the surgical risk measured by the ASA and diagnosis group (table 4).

When considering the differences among the centres, hospital 5 had the highest appropriate rate (97.1%) and the lowest uncertain-inappropriate rate (2.9%). Hospital 4 had the highest inappropriate rate (2.6%), though just on 4 patients. Hospital 6 had the highest uncertain-inappropriate rate (12.8%) and the lowest appropriate rate (87.2%). These differences were statistically significant (table 5). Hospital 6 had an OR 4.8 higher than hospital 5, of having an uncertain-inappropriate case, which were significant differences ($p < 0.05$).

We recorded the mortality rate, complications during the intervention, admission, and length of stay at each hospital. At 3 months, we also recorded the mortality rate, complications, and readmissions resulting from the intervention. The mortality rate

Table 2 Appropriateness by diagnostic group

Variables	Appropriateness classification					
	Appropriate		Uncertain		Inappropriate	
Age						
<75 years	795	(91.3)	75	(8.6)	1	(0.1)
≥75 years	82	(92.1)	1	(1.1)	6	(6.7)
Surgical risk						
Low-medium (ASA I/II/III)	869	(91.8)	72	(7.6)	6	(0.6)
High (ASA IV)	8	(61.5)	4	(30.8)	1	(7.7)
Diagnostic group						
Symptomatic uncomplicated cholelithiasis	569	(99.3)	4	(0.7)	0	(0)
Symptomatic complicated cholelithiasis	239	(98.4)	4	(1.7)	0	(0)
Asymptomatic cholelithiasis	60	(46.2)	63	(48.5)	7	(5.4)
Other indications	9	(64.3)	5	(35.7)	0	(0)
Type of intervention						
Laparotomy	126	(90.7)	11	(7.9)	2	(1.4)
Minilaparotomy	23	(85.2)	4	(14.8)	0	(0)
Laparoscopy	715	(91.7)	60	(7.7)	5	(0.6)

Chi square test. All $p < 0.001$ except type of intervention ($p = 0.55$).

Table 3 Most frequently encountered indications and appropriateness category

Indication	N	(%)	Appropriateness
Symptomatic cholelithiasis without complications			
Micro or Macro, NG, NBD + age <75 + low-medium surgical risk	451	(47)	Appropriate
Macro, NG, NBD + age >75 + low-medium surgical risk	25	(2.6)	Appropriate
Micro or Macro, TG, NBD + age <75 + low-medium surgical risk	33	(3.4)	Appropriate
Symptomatic cholelithiasis with complications			
Cholecystitis + age <75 + low-medium surgical risk	75	(7.8)	Appropriate
Pancreatitis + cholelithiasis + bile duct w/cholelithiasis + age <75 + low-medium surgical risk	76	(7.9)	Appropriate
Asymptomatic cholelithiasis			
Silent calculi + age 46–75 + low-medium surgical risk	62	(6.5)	Uncertain
Silent calculi + age >75 + low-medium surgical risk	6	(0.6)	Inappropriate
Pura asymptomatic + age 46–75 + high surgical risk	1	(0.1)	Inappropriate
Incidental finding pre surgical intervention + age <75 + low-medium surgical risk	22	(2.3)	Appropriate
Diabetes and symptomatic cholelithiasis without complications			
Age <75 + low-medium surgical risk	37	(3.9)	Appropriate

All indications encountered in 20 patients (2% of the total sample) or more are included, and those more frequently encountered in each appropriateness category.

At echography: Micro: microlithiasis; Macro: macrolithiasis; NG: normal gallbladder; TG: thickened gallbladder; NBD: normal bile duct. Low-medium surgical risk: ASA I to III; high surgical risk: ASA IV.

was 0.3%, the mean length of stay was 4 days, and 23.8% developed at least one local complication, while 10.9% presented with a systemic complication from admission until revision at 3 months. When comparing if the rates of any of these adverse outcomes differed among appropriateness groups, we did not find any statistically significant differences among appropriateness groups with respect to these adverse outcomes.

A sample of 121 records was reviewed by members of the research team, for a few relevant variables, to test the accuracy of the data retrieved by the reviewers. We found Kappa correlations higher than 0.99 in all cases.

DISCUSSION

The present study assessed prospectively the appropriateness of surgical indications for patients undergoing cholecystectomy for nonmalignant diseases in a region of Spain in 1999 to 2000, using updated appropriateness criteria developed by a national panel of experts directed by this team. Applied as a screening tool for patients scheduled for surgery, those appropriateness criteria indicated a very low percentage (0.7%) of potentially inappropriate procedures, compared to other studies in which a similar methodology has been used. Of those RAND appropriateness studies done on gastroenterologic problems, inappropriateness rates ranged from 18 to 49% for the use of upper gastrointestinal endoscopy¹²⁻¹⁴ and from 25 to 31% for colonoscopy.¹⁵ From the previous RAND cholecystectomy panels developed in the 1980s in the United States, Israel, and United Kingdom, the explicit criteria were applied only in the Israeli and United Kingdom field studies.^{5,16} The UK study found inappropriateness rates ranging from 2 to 30% depending on which of two different sets of criteria were used to evaluate cases, one compiled by a panel of surgeons and the other a mixed panel. They concluded that those judged as inappropriate had only vague symptoms, which matched our findings. The inappropriateness rates ranged from 17 to 36% in Israel, much higher than ours. Other studies using RAND appropriateness

methodology in other procedures, i.e. cardiac procedures, found a low inappropriateness rate similar to ours; one was a study in the US¹⁷ and another in Sweden.¹⁸

Some authors have included patients in the uncertain category with the inappropriate cases,^{19,20} a group that was called less-than-appropriate. We did so in part because we found very few inappropriate cases and also because for logistic regression analysis either a dichotomous dependent variable is needed or, if more categories are included, ordinal categorical variables are preferred. Because we could not assume the latter for the three appropriateness categories, we chose the first option comparing appropriate versus less-than-appropriate categories.

The percentage of indications considered uncertain in this study was fairly low compared to other RAND studies. Just those mentioned previously, where the inappropriateness rate was similar, had a low uncertain rate as well.^{17,18} Nevertheless, these results indicate that there were still 7.9% of cases for which the panelists could not judge the appropriateness of the indication. Therefore, it would be adequate to study the risk-benefit ratio of performing a cholecystectomy on these patients, which, in our case, were exclusively asymptomatic cholelithiasis with low surgical risk, and one patient with high surgical risk.

Indications that were more likely to be considered inappropriate were those in which the surgical risk was high (ASA IV), the patients were older than 75 years, or the patient was diagnosed with asymptomatic cholelithiasis. This is because our criteria did not consider the intervention appropriate in patients with a high surgical risk (either because of age or ASA) or in those with asymptomatic gallstones. Some surgical interventions with outstanding patient benefits, i.e. hip prosthesis for patients with hip osteoarthritis, are performed regardless of ASA stage or age.²¹ However, with cholecystectomy even previous studies showed that the panel criteria were highly correlated with the degree of comorbidity of patients, and their conclusions were similar to ours.²² Finally, asymptomatic cholelithiasis has not been considered an indication for prophylactic cholecystectomy

Table 4 Differences in patients clinical variables by hospitals

Variables	Hospital 1		Hospital 2		Hospital 3		Hospital 4		Hospital 5		Hospital 6	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
Surgical risk												
Low-medium (ASA I-III)	106	(99.1)	138	(99.3)	239	(100)	154	(98.1)	68	(100)	242	(96.8)
High (ASA IV)	1	(0.9)	1	(0.7)	0	(0)	3	(1.9)	0	(0)	8	(3.2)
Diagnosis												
Symptomatic cholelithiasis	95	(88.8)	99	(71.2)	210	(87.9)	133	(84.7)	61	(89.7)	218	(87.2)
Uncomplicated	68	(71.6)	64	(64.7)	156	(74.3)	83	(62.4)	47	(77.1)	155	(71.1)
Complicated	27	(28.4)	35	(35.3)	54	(25.7)	50	(37.6)	14	(22.9)	63	(28.9)
Asymptomatic cholelithiasis	11	(10.3)	37	(26.6)	26	(10.9)	20	(12.7)	6	(8.8)	30	(12.0)
Other indications	1	(0.9)	3	(2.2)	3	(1.3)	4	(2.6)	1	(1.5)	2	(0.8)

All chi square test: p<0.05. There were no statistically significant differences in sex, age, and type of intervention by hospital.

Table 5 Appropriateness by hospital

	Hospital 1		Hospital 2		Hospital 3		Hospital 4		Hospital 5		Hospital 6		Total	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
Appropriate	96	(89.7)	133	(95.7)	222	(92.9)	142	(90.5)	66	(97.1)	218	(87.2)	877	(91.4)
Uncertain	11	(10.3)	6	(4.3)	16	(6.7)	11	(7.0)	2	(2.9)	30	(12.0)	76	(7.9)
Inappropriate	0	(0)	0	(0)	1	(0.4)	4	(2.6)	0	(0)	2	(0.8)	7	(0.7)
Total	107	(11.2)	139	(14.5)	239	(24.9)	157	(16.4)	68	(7.1)	250	(26.0)	960	

Chi square test. p<0.05

because of its benign course. Asymptomatic cholelithiasis becomes symptomatic annually in from 1% to 4% of cases, reaches 20% at 20 years, and is almost always presented with previous symptoms before complications develop.²³⁻²⁶

Why did we find such a low inappropriateness rate compared to previous studies? We do not have a response to that, though some hypothesis can be put forward. Our panel criteria could be more liberal with respect to performing the intervention. There is no way to compare our panel criteria with previous panels since we have introduced new variables and others have changed significantly. We have tried to avoid panel bias by having a mixed panel of surgeons and gastroenterologists. Another possible explanation is the quality of the data from the medical records. Inappropriateness is led mainly by surgical risk and diagnosis. We did not record directly those two variables but relied on the classification done by the anesthesiologist and surgeon respectively on the medical record. A misclassification on any of those two variables could lead to a different appropriateness classification. Also, it is feasible that our surgeons are taking decisions in a proper way. Waiting lists for many surgical procedures, a common problem in health care systems such as the Spanish, are pushing physicians to prioritize constantly and so probably choosing those patients with clearly better defined indications. Finally, the introduction of laparoscopic cholecystectomy has lowered the threshold for surgery.²⁷

In any case, besides the low rate of inappropriateness, we found some differences among our hospitals in the rate of appropriateness. Hospital 5 or 2, which have the lowest inappropriateness rate, seems to match the panel criteria in all diagnosis groups, even though no member of the panel belongs to those centres nor were surgeons of that centres aware of the study goals. Limitations related to the work of the panel were described in a previous study.¹⁰ Additionally, data collection presented some limitations. The three blinded reviewers were physicians trained to assess and record the main variables of the algorithm in a standardized manner to reduce the chances of bias. We checked during their training for their reliability with good results, and about their accuracy by checking some variables with excellent correlations. However, the quality of the data of some important variables necessary for the appropriateness algorithm can be questionable when depending exclusively on the medical record. In our case, most relevant variables (age, surgical risk, diagnosis, comorbidities) were properly recorded.

Laparoscopy is a surgical option that has rapidly increased in the last few years. In this study, we did not evaluate the appropriateness of the different surgical options currently available, such as open cholecystectomy vs. laparoscopy, which could be the subject of another panel to create explicit criteria for those specific surgical techniques, and, later, another field study. Even so, it was inevitable that panelists were thinking of laparoscopy as the first surgical option. Also, our results showed that in appropriate indications laparoscopy was more likely to be used, mainly because those were patients with lower surgical risk.

This study was performed in six public hospitals in the Spanish health care system, a public, universal coverage system. The generalizability of our results to other different health care systems is questionable though it is quite likely that results in other Spanish public centres are the same.

In conclusion, this prospective observational study demonstrates that the rate of inappropriate-uncertain cholecystectomies in our public hospitals, though low, showed variations in appropriateness among centres. Surgery was more likely to be considered inappropriate in patients with asymptomatic cholelithiasis, or high surgical risk, which matches current state-of-the-art criteria about the indication of cholecystectomy for patients with gallbladder disease. Even so, as suggested by the authors of the RAND appropriateness method,²⁸ if

appropriateness is to be improved it must be assessed at the level of each patient, hospital, and physician.

This study was partially supported by a grant from the Fondo de Investigación Sanitaria (98/002-03). Amaia Bilbao received a grant from the Department of Health of the Basque Government.

We thank the following individuals for their contributions to this study: Drs Aguado, Arenas, Atín, Baile, Barrios, Casanova, Hinojosa, Lanas, Monés, Orive, Pons, Robles, Sarabia, Suárez Alzamora, Valdivieso, Varona, and Zaballa. We are also grateful to Y. Echeverria, A. Higuelmo, I. Lafuente, A. Rodriguez, and I. Vidaurreta for their contribution to the development of the panel of experts, data retrieval, and data introduction. We thank the staff members of the different services, research and quality units, as well as the medical records sections of the participant hospitals for their support.

REFERENCES

- Lam CM, Murray FE, Cuschieri A. Increased cholecystectomy rate after the introduction of laparoscopic cholecystectomy in Scotland. *Gut* 1996;38:282-4.
- Johanning JM, Gruenberg JC. The changing face of cholecystectomy. *Am Surg* 1998;64(7):643-7.
- McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. *N Engl J Med* 1982;307(21):1310-4.
- Pilpel D, Fraser GM, Kosecoff J, Weitzman S, Brook RH. Regional differences in appropriateness of cholecystectomy in a prepaid health insurance system. *Public Health Rev* 1992;20(1-2):61-74.
- Brook RH. Appropriateness: the next frontier. *BMJ* 1994;308(6923):218-9.
- Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986;2:53-63.
- Park RE, Fink A, Brook RH, et al. Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health* 1986;76(7):766-72.
- Fraser GM, Pilpel D, Hollis S, Kosecoff J, Brook RH. Indications for cholecystectomy: the results of a consensus panel approach. *Qual Assur Health Care* 1993;5(1):75-80.
- Scott EA, Black N. Appropriateness of cholecystectomy in the United Kingdom: a consensus panel approach. *Gut* 1991;32(9):1066-70.
- Quintana JM, Cabriada J, López de Tejada I, et al. Development of Explicit Criteria for Cholecystectomy. *Qual Saf Health Care* 2002;11(4):320-6.
- Schneider AJL. Assessment of risk factors and surgical outcome. *Surg Clin North Am* 1983;63:1113-26.
- Froehlich F, Burnand B, Pache I, et al. Overuse of upper gastrointestinal endoscopy in a country with open-access endoscopy: a prospective study in primary care. *Gastrointest Endosc* 1997;45(1):13-9.
- Quine MA, Bell GD, McCloy RF, Devlin HB, Hopkins A. Appropriate use of upper gastrointestinal endoscopy: a prospective audit. Steering Group of the Upper Gastrointestinal Endoscopy Audit Committee. *Gut* 1994;35(9):1209.
- Seematter-Bagnoud L, Vader JP, Wietlisbach V, Froehlich F, Gonvers JJ, Burnand B. Overuse and underuse of diagnostic upper gastrointestinal endoscopy in various clinical settings. *Int J Qual Health Care* 1999;11(4):301-8.
- Froehlich F, Pache I, Burnand B, et al. Performance of panel-based criteria to evaluate the appropriateness of colonoscopy: a prospective study. *Gastrointest Endosc* 1998;48(2):128-36.
- Scott EA, Black N. Appropriateness of cholecystectomy: the public and private sectors compared. *Ann R Coll Surg Engl* 1992;74(4Suppl):97-101.
- Leape LL, Hilborne LH, Park RE, et al. The appropriateness of use of coronary artery bypass graft surgery in New York State. *JAMA* 1993;269(6):753-60.
- Bengtson A, Herlitz J, Karlsson T, Brandrup-Wognsen G, Hjalmarson A. The appropriateness of performing coronary

angiography and coronary artery revascularization in a Swedish population. *JAMA* 1994;271(16):1260-5.

19 Hilborne LH, Leape LL, Bernstein SJ, et al. The appropriateness of use of percutaneous transluminal coronary angioplasty in New York State. *JAMA* 1993;269(6):761-5.

20 Leape LL, Hilborne LH, Schwartz JS, et al. The appropriateness of coronary artery bypass graft surgery in academic medical centers. Working Group of the Appropriateness Project of the Academic Medical Center Consortium. *Ann Intern Med* 1996;125(1):8-18.

21 Quintana JM, Arostegui I, Azkarate J, et al. Evaluation of explicit criteria for total hip joint replacement. *J Clin Epidemiol* 2000;53(12):1200-8.

22 Kahn KL, Park RE, Brook RH, et al. The effect of comorbidity on appropriateness ratings for two gastrointestinal procedures. *J Clin Epidemiol* 1988;41(2):115-22.

23 National Institutes of Health Consensus Development Panel: Gallstones and laparoscopic cholecystectomy. *JAMA* 1993;269:1018-24.

24 Strasberg S, Clavien PA. Overview of therapeutic modalities for the treatment of gallstone diseases. *Am J Surg* 1993;165:420-6.

25 Friedman GD. Natural history of asymptomatic and symptomatic gallstones. *Am J Surg* 1993;165:399-404.

26 Mulvihill SJ, Somberg KA. Surgical management of biliar lithiasis and postoperative complications. In: Sleisenger MH, Fordtran JS, editors. *Gastrointestinal disease: pathology, diagnosis, management*. Philadelphia: WB Saunders, 1994:1881-9.

27 Escarce JJ, Chen W, Schwartz JS. Falling cholecystectomy thresholds since the introduction of laparoscopic cholecystectomy. *JAMA* 1995;273(20):1581-5.

28 Brook RH, Park RE, Chassin MR, Solomon DH, Keesey J, Kosecoff J. Predicting the appropriate use of carotid endarterectomy, upper gastrointestinal endoscopy and coronary angiography. *N Engl J Med* 1990;323(17):1173-7.

Received 13 November 2002, accepted 12 May 2003