Pain 105 (2003) 415–423

PAIN

www.elsevier.com/locate/pain

Preoperative prediction of severe postoperative pain

C.J. Kalkman[a,*], K. Visser[b], J. Moen[a], G.J. Bonsel[c], D.E. Grobbee[d], K.G.M. Moons[a,d]

[a]*Department of Anesthesiology, Division of Perioperative Care and Emergency Medicine, University Medical Center Utrecht, P.O. Box 85500, 3508 GA Utrecht, The Netherlands*
[b]*Department of Anesthesiology, Academic Medical Center, Amsterdam, The Netherlands*
[c]*Department of Public Health, Academic Medical Center, Amsterdam, The Netherlands*
[d]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands*

Received 3 December 2002; received in revised form 25 April 2003; accepted 2 June 2003

Abstract

We developed and validated a prediction rule for the occurrence of early postoperative severe pain in surgical inpatients, using predictors that can be easily documented in a preoperative setting. A cohort of surgical inpatients ($n = 1416$) undergoing various procedures except cardiac surgery and intracranial neurosurgery in a University Hospital were studied. Preoperatively the following predictors were collected: age, gender, type of scheduled surgery, expected incision size, blood pressure, heart rate, Quetelet index, the presence and severity of preoperative pain, health-related quality of life the (SF-36), Spielberger's State–Trait Anxiety Inventory (STAI) and the Amsterdam Preoperative Anxiety and Information Scale (APAIS). The outcome was the presence of severe postoperative pain (defined as Numeric Rating Scale ≥ 8) within the first hour postoperatively. Multivariate logistic regression in combination with bootstrapping techniques (as a method for internal validation) was used to derive a stable prediction model. Independent predictors of severe postoperative pain were younger age, female gender, level of preoperative pain, incision size and type of surgery. The area under the receiver operator characteristic (ROC) curve was 0.71 (95% CI: 0.68–0.74). Adding APAIS scores (measures of preoperative anxiety and need for information), but not STAI, provided a slightly better model (ROC area 0.73). The reliability of this extended model was good (Hosmer and Lemeshow test p-value 0.78). We have demonstrated that severe postoperative pain early after awakening from general anesthesia can be predicted with a scoring rule, using a small set of variables that can be easily obtained from all patients at the preoperative visit. Before this internally validated preoperative prediction rule can be applied in clinical practice to support anticipatory pain management, external validation in other clinical settings is necessary.

Keywords: Postoperative pain; Pain measurement; Prognosis; Prediction model

1. Introduction

Many surgical patients awake from anesthesia in severe pain or develop severe pain in the first hour after arrival in the postoperative anesthesia care unit (PACU). Many studies suggest that postoperative pain is often undertreated as a result of inadequate education of care providers and fear for causing opioid addiction (Donovan et al., 1987; Kotzer, 2000; Kehlet et al., 1996). Severe postoperative pain (typically recorded with visual analog (VAS) or numerical rating scales of 8 or higher) results in extreme patient discomfort and reduces patient satisfaction (Myles et al., 2000). It may also contribute to the development of chronic pain syndromes, such as post thoracotomy pain (Katz et al., 1996; Perttunen et al., 1999). Severe pain also has neurohormonal effects, including increased sympathetic activity and increased levels of stress hormones, which may increase the incidence of postoperative complications such as myocardial infarction or stroke resulting from rupture of vulnerable atherosclerotic plaques (Airaksinen, 1999; Muller, 1999), and it may interfere with normal breathing patterns, thereby promoting atelectasis and postoperative pulmonary infections (Nguyen et al., 2001; Ellstrom et al., 1998; Joris et al., 1998). In the elderly it promotes delirium (Lynch et al., 1998). It would be desirable to preoperatively distinguish patients who are at high risk for developing severe postoperative pain from those who have low risk. Patients at high risk might benefit from aggressive analgesic interventions either pre-emptively or in the early phase of

* Corresponding author. Tel.: +31-30-2509677; fax: +31-30-2541828.
E-mail address: c.j.kalkman@azu.nl (C.J. Kalkman).

0304-3959/03/$30.00
doi:10.1016/S0304-3959(03)00252-5

recovery from anesthesia. This may especially apply to patients at risk for cardiovascular or pulmonary complications following surgery.

In clinical practice, the type of surgical procedure is commonly the single factor determining the initial analgesic strategy, often codified in local protocols and guidelines. Other important (preoperative) predictors of postoperative pain that have been reported include female gender, severity of preoperative pain and younger age (Kotzer, 2000; Thomas et al., 1998; Morin et al., 2000; Bisgaard et al., 2001). Preoperative anxiety too has long been recognized as a significant predictor of postoperative pain (Chapman, 1977; Scott et al., 1983). However, most studies are small and dedicated to specific procedures.

The purpose of the present study was to develop and validate an easily applicable preoperative prediction rule for the occurrence of early postoperative severe pain in surgical inpatients. To allow future implementation in clinical practice, only predictors were selected that can be documented in a preoperative outpatient setting. If successful, such a prediction rule could further optimize perioperative pain treatment and reduce attendant complications, by allowing physicians to preoperatively classify surgical patients according to their risk.

2. Patients and methods

2.1. Patients

The present study analyzed data from a cohort of patients who participated in a large randomized trial investigating differences in the incidence of postoperative nausea and vomiting after intravenous or inhalational general anesthesia executed in the Academic Medical Center, Amsterdam, The Netherlands. The design and primary results have been reported elsewhere (Visser et al., 2001). In brief, after informed consent was given, 1957 surgical patients (1416 inpatients and 541 outpatients), aged 18–85, were randomized to either surgery with intravenous anesthesia using propofol or surgery with inhalational anesthesia using isoflurane and nitrous oxide. There were no restrictions to the type of surgery, except cardiac surgery and intracranial neurosurgical procedures. Exclusion criteria were emergency surgery, pregnancy, ASA physical status 4 and morbid obesity (weight >120 kg). ASA physical status denotes the American Society of Anesthesiologists preoperative risk classification based on comorbidity, where ASA 1 denotes a patient presenting for surgery who is otherwise healthy, ASA 2 refers to a patient with mild comorbidity, such as adequately treated hypertension, ASA 3 refers to patients with significant comorbidity that interferes with normal activity (coronary artery disease, diabetes, COPD), and finally ASA 4 refers to patients whose comorbidity is a constant threat to life (heart failure, severe COPD).

Induction of anesthesia was achieved with thiopental in patients randomized to isoflurane/nitrous oxide, and with propofol in patients randomized to total intravenous anesthesia with propofol/air. Anesthesia was maintained with propofol or isoflurane in nitrous oxide according to the randomization. The treatment allocation was concealed until immediately before induction of anesthesia. Intraoperatively, the anesthesiologist was free to use opioid analgesics and muscle relaxants as needed. Postoperative analgesics were administered by the PACU nurse as prescribed by the responsible anesthesiologist and dictated by the demands of the patient. This regimen always included paracetamol (1000 mg) four times daily, initially administered as suppositories until the patient was able to drink, and combined with NSAIDS, intramuscular (IM) or intravenous (IV) morphine depending on the type and extent of the surgical procedure. For some larger procedures the anesthesiologist prescribed morphine by continuous infusion, but for most procedures morphine was administered on demand, typically in doses of 10 mg IM, or 2–5 mg IV when the patient complained of severe pain. These protocols reflected standard practice in the hospital at the time of the study. The present analysis was based on data from the 1416 surgical inpatients.

2.2. Outcome: presence of severe pain

A trained research nurse—who was not responsible for postoperative care of the patient—recorded the severity of pain using an 11-point-numerical rating score (NRS) (where 0 indicates no pain at all, and 10 the most severe pain imaginable) every 15 min until discharge from the PACU. A separate item in the case record form allowed recording of whether the patient was sufficiently alert to complete an NRS. The outcome of the present study was the presence of severe pain, defined as a numerical rating score (NRS) ≥ 8, occurring at least once within the first hour after arrival at the PACU. In the present analyses, patients who were too drowsy or sleepy to provide a pain score at the moment of measurement received a NRS-score of 0 (no pain) for that time point.

2.3. Candidate predictors

We selected 26 candidate preoperative predictors of severe postoperative pain. These included age, gender, Quetelet index (indicator of body mass: weight divided by height (in meters) squared), surgical history, type of scheduled surgery (six levels: orthopedic, intra-abdominal, laparoscopic, ear/nose/throat, ophthalmic, other), expected incision size for the scheduled surgical procedure (two levels: small incision, <10 cm, large incision, ≥ 10 cm), systolic and diastolic blood pressure, heart rate, and the severity of preoperative pain as assessed with the pain domain of the quality of life questionnaire Short Form 36 (SF-36) that was transformed to a score of 0 (no pain) to 10

in conformity with numerical rating or visual analog scores (Aaronson et al., 1998; Moseley et al., 2002), duration of surgery (from incision to closure) and anesthetic technique (isoflurane versus propofol). Additional predictors were the remaining domain scores of the SF-36 (seven individual and two general scores (predictors), all ranging from 0 to 100), two predictors of the Spielberger's State–Trait Anxiety Inventory (STAI; the State and Trait score, both ranging from 20 to 80) (Spielberger et al., 1970) and two predictors obtained from the Amsterdam Preoperative Anxiety and Information Scale (APAIS) (Moerman et al., 1996). The APAIS consists of six questions (each scored on a 5-point-Likert scale from 1 (not at all) to 5 (extremely)) specifically designed to assess the patients' preoperative anxiety score (range 4–20) and a score for the patients' need for information regarding the scheduled surgery and anesthesia (information seeking behavior, range 2–10).

2.4. Data analysis

We first estimated the association between each candidate predictor and the outcome (bivariate analysis). However, pre-selection of predictors based on p-values estimated from bivariate analyses may result in unstable prediction models (Chatfield, 1995; Harrell et al., 1996). Therefore, all candidate predictors were considered in the multivariable analysis using logistic regression modeling. Since the final aim in prognostic research is to obtain optimal prediction with minimal burden to the patient, clinician and third party payers, we adopted a hierarchical approach in the modeling in which easily obtainable predictors were included first (Moons et al., 1999; Moons and Grobbee, 2002). The initial (overall) model included all (13) candidate predictors minus the (13) predictors obtained from the self-report questionnaires (SF-36, APAIS, and STAI) as the latter were considered as more cumbersome to measure. In contrast to etiological research, it is common in prediction research to use a more liberal p-value than 0.05, e.g. 0.15 or even 0.20 or 0.25, for keeping variables in the model (Chatfield, 1995; Harrell et al., 1996). Hence, the overall multivariable model was reduced by manually deleting (one by one) predictors with p-value >0.15 based on the log likelihood ratio test. The reduced model thus including only predictors with p-value ≤ 0.15, was then extended by adding, separately and in combination, the SF-36, APAIS and STAI predictors, to estimate their incremental value in the prediction of postoperative pain. Of particular interest was the difference in incremental value, if any, of the shorter APAIS and the much longer STAI questionnaire. In multivariable analysis, 79 subjects (5.6%) were deleted due to missing values on one of the predictors.

The predictive accuracy of the prediction models was estimated by their calibration (reliability or goodness of fit) and discrimination. A models' calibration was evaluated by comparing the models' predicted probability of severe postoperative pain with observed proportions and tested using the Hosmer and Lemeshow (1989) method. The models' ability to discriminate between patients with and without severe postoperative pain was estimated by the area under the receiver operating characteristic curve (ROC area) of the model. Differences in discriminative ability between models was estimated using the difference in ROC area taking into account the correlation between models as they were based on the same subjects (Hanley and McNeil, 1983). As the ROC approach assumes equal weight to false positive and false negative predictions and is prevalence independent, it may not directly reflect the clinical value of a model (Moons et al., 1997). Hence, we estimated the absolute number of correctly predicted patients with and without severe pain across various predicted risk categories.

Bootstrapping techniques were used to validate the final prediction model, i.e. to adjust the estimated model performance and regression coefficients (odds ratios) for overoptimism or overfitting (Harrell et al., 1996; Efron and Tibshirani, 1993). The model's performance obtained after bootstrapping can be considered as the performance that can be expected in similar future patients. Random bootstrap samples were drawn with replacement (100 replications) from the data set consisting of all patients with complete data ($n = 1346$). The multivariable selection of variables was repeated within each bootstrap sample. All analyses were performed using S-plus 2000 (Insightful Corp., Seattle, WA, USA).

3. Results

Of all patients, 366 suffered from severe pain within the first hour after arriving at the PACU (incidence 25.8%). Fig. 1 shows the distribution of pain scores at the end of the first hour in the recovery room. All candidate predictors were to some extent associated with the outcome except for Quetelet index, previous surgery, heart rate and duration of surgery. Of the patients randomized to isoflurane/nitrous oxide anesthesia 28% developed severe postoperative pain compared to 24% in the propofol group (p-value $= 0.13$).

The initial multivariable model with all (13) predictors minus the APAIS, STAI and SF-36 yielded an ROC area of 0.71 (95% CI: 0.68–0.74). Of these 13 only five predictors, i.e. age, gender, preoperative pain, incision size, and type of surgery (included as five indicator variables with ophthalmic surgery as reference), independently contributed to the prediction of the outcome defined as a p-value ≤ 0.15 (Table 2, second column). Other predictors, that seemed relevant in univariable sense (Table 1) such as blood pressure and anesthetic technique, were no independent predictors in multivariable analysis. Apparently, their predictive information was already provided for by the five retained predictors. The reduced model including the five predictors yielded the same ROC area (0.71) as the initial model. Further exclusion of these five predictors

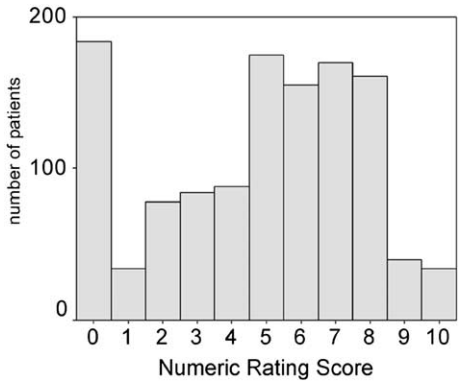

Fig. 1. Histogram of Numerical Rating Scores (NRS) at 60 min after arriving in the Post Anesthesia Care Unit; Y-axis depicts number of patients reporting each score.

significantly decreased the ROC area. None of the SF-36 predictors (when added to the model as separate domain scores or as summary scores) improved this reduced model. Also the two STAI anxiety predictors did not improve the reduced model: the odds ratio of the STATE variable was 1.02 per unit change (p-value = 0.29) and of the TRAIT variable 1.0 per unit (p-value = 0.83). The ROC area remained 0.71. The two APAIS predictors, however, improved the reduced model (Table 2). The ROC area significantly increased to 0.73 (95% CI: 0.70–0.76). Age, gender, type of surgery, incision size, preoperative pain score and the two APAIS scores were therefore considered as the independent predictors of postoperative severe pain, and included in the final model (Table 2).

After bootstrapping and adjustment for overfitting, the ROC area of the final model was 0.71 (overoptimism 0.02). Table 2 (columns 3 and 4) shows the regression coefficients and odds ratio per predictor adjusted for overfitting (i.e. after bootstrapping). The reliability of this final model after bootstrapping was good (p-value Hosmer and Lemeshow test was 0.78).

Using the regression coefficients of the final prediction model one can estimate for each surgical patient the probability of developing severe postoperative pain after awakening from anesthesia using the formula given in Table 2 (legend) or by using the nomogram (Fig. 2). As an example

Table 1
Bivariate association of each candidate predictor and the presence or absence of severe pain (VAS ≥ 8) within the first hour at recovery room

Preoperative predictor	Severe pain ($N = 366$)	No severe pain ($N = 1050$)	p-Value
Age (years)[a]	42 (14)	46 (15)	<0.001
Female gender	65	55	0.001
Quetelet index (kg/m^2)[a]	24.8 (4.5)	24.8 (4.1)	0.85
Previous surgery	82	83	0.5
Preoperative pain score[a]	4 (3)	2 (3)	<0.001
Heart rate (bpm)[a]	73 (10)	73 (10)	0.6
Diastolic blood pressure (mmHg)[a]	80 (11)	82 (11)	0.04
Systolic blood pressure (mmHg)[a]	130 (18)	133 (19)	0.01
Duration of surgery (hours)[a]	2.13 (1.14)	2.15 (1.15)	0.71
Inhalation anesthesia	53	49	0.072
Type of surgery			<0.001
Ophthalmic	3	10	
Laparoscopy	5	8	
Ear/nose/throat	15	22	
Abdominal	15	7	
Orthopedic	34	19	
Other	27	32	
Medium/large incision size	56	40	<0.001
APAIS anxiety score[a]	10 (4)	9 (4)	<0.001
APAIS need for information[a]	6 (2)	7 (4)	0.12
Spielbergers' STATE anxiety level[a]	40 (12)	42 (11)	<0.001
Spielbergers' TRAIT anxiety level	39 (10)	38 (10)	0.03
SF-36[a]			
Physical functioning	69 (28)	77 (27)	<0.001
Social functioning	72 (28)	78 (25)	<0.001
Physical restraints	50 (44)	64 (49)	<0.001
Emotional restraints	66 (41)	71 (39)	0.04
Mental health	71 (17)	74 (18)	0.003
Vitality	60 (19)	66 (20)	<0.001
General health	67 (21)	70 (20)	0.007
Physical summary score	44 (11)	48 (11)	<0.001
Mental summary score	47 (11)	48 (11)	0.12

Values are percentages unless stated otherwise.
[a] Mean (standard deviation).

Table 2
Reduced and extended (final) model to predict preoperatively the occurrence of severe pain (VAS ≥ 8) within first hour at recovery room

	Reduced model	Extended (final) model		
	Regression coefficient (SE)	Regression coefficient (SE)[a]	OR[a]	*p*-Value
Female gender	0.36 (0.14)	0.22 (0.14)	1.24	0.12
Age (per year)	−0.018 (0.0046)	−0.016 (0.0047)	0.98	<0.001
Pre-operative pain (per unit)	0.16 (0.025)	0.14 (0.025)	1.15	<0.001
Type of surgery				
Ophthalmic[b]	[b]	[b]	[b]	[b]
Laparoscopic	0.36 (0.42)	0.38 (0.43)	1.46	0.37
Ear/nose/throat	0.65 (0.35)	0.59 (0.36)	1.80	0.10
Orthopedic	1.03 (0.35)	0.97 (0.35)	2.64	0.006
Abdominal	1.36 (0.38)	1.37 (0.38)	3.56	<0.001
Other	0.52 (0.34)	0.48 (0.34)	1.62	0.17
Medium/large incision size	0.28 (0.16)	0.23 (0.16)	1.26	0.15
APAIS anxiety (per unit)	[c]	0.053 (0.017)	1.05	0.002
APAIS need for info (per unit)	[c]	−0.080 (0.032)	0.92	0.01
Intercept (constant)	−1.86 (0.40)	−1.74 (0.46)		<0.001

SE, standard error; OR, odds ratio.

Probability of postoperative severe pain by final model = 1/(1 + exp(−(−1.74 + 0.22 × female gender − 0.016 × age + 0.14 × preoperative pain + (0 or 0.38 or 0.59 or 0.97 or 1.37 or 0.48) × type of surgery + 0.053 × APAIS anxiety − 0.080 × APAIS need for info + 0.23 × incision size.

[a] Regression coefficient and corresponding odds ratio after bootstrapping (i.e. adjusted for overfittng).

[b] Reference group.

[c] Not included in the model.

of using this nomogram, a woman (which corresponds to 3 points as determined from the 'Points' scale on the top of the figure), 55 years of age (8 points), with a preoperative pain score of 8 (18 points), scheduled for an orthopedic surgical procedure (14 points) with medium-large incision size (3 points), who does not need any additional information about the surgery or anesthesia (level 2; 9 points), has an anxiety score of 12 (6 points) receives a 'Total Points' score

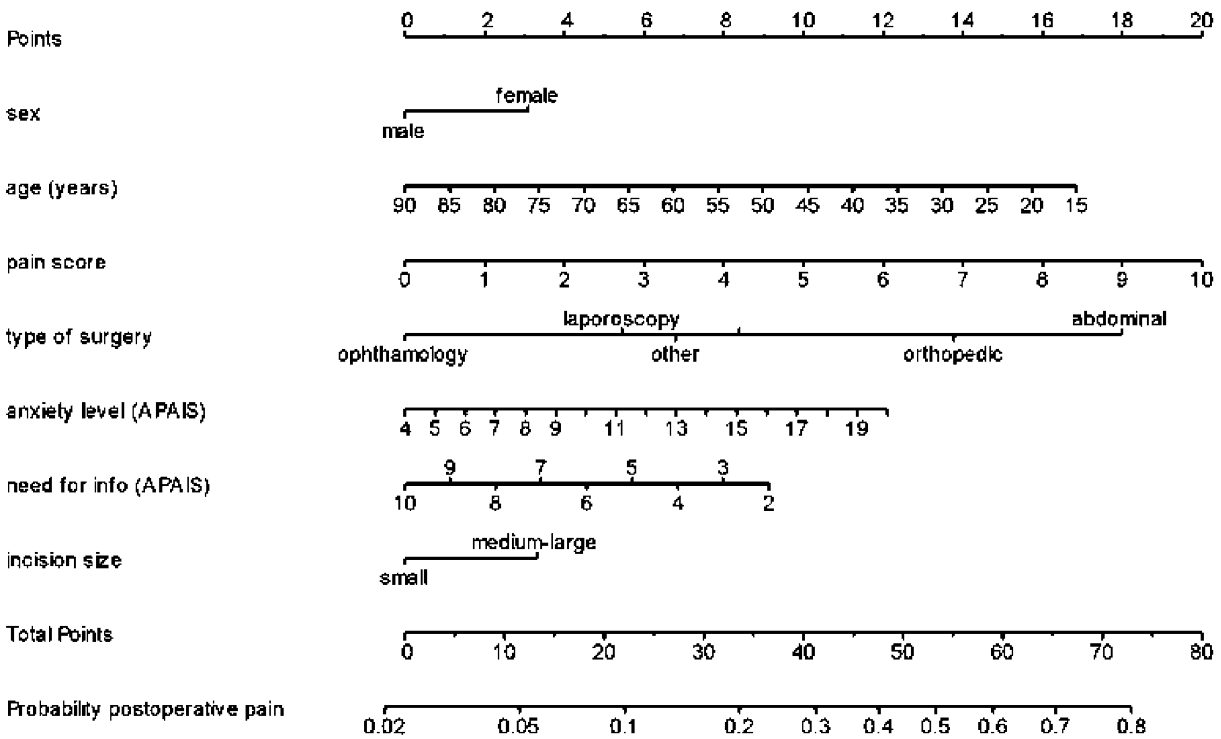

Fig. 2. Nomogram relating the predictors of the final model (Table 2) to the probability of severe pain within the first hour at the Post Anesthesia Care Unit in surgical patients. See text for instructions on how to use such a nomogram.

of $3 + 8 + 18 + 14 + 9 + 6 + 3 = 61$. Using the lower two scales of the figure, this score corresponds to a probability of severe postoperative pain in the PACU of about 0.65. The length of the line of each predictor in the nomogram indicates the relative contribution of the predictor (with the left boundary as anchor) to the probability of severe postoperative pain. A third method to estimate the score and probability of postoperative pain in future surgical patients is to use the score tables in Appendix-A.

Table 3 shows the observed number of patients with and without severe pain across score and risk categories estimated by the final model. The table shows the observed incidence (mean probability) of severe postoperative pain per predefined risk stratum, which increased from 7 to 49%. Reading the table horizontally provides the estimates of the sensitivity and specificity of the scoring rule at different score thresholds (risk subgroups). For example, introducing a threshold at 40, a score ≤ 40 will be considered as 'test-negative' ($n = 867$) and >40 as 'test positive' ($n = 470$). When administering analgesics pre-emptively to patients with a score >40, 61% (21% + 40%) of the 348 patients who actually developed severe pain would be treated correctly (sensitivity or true positive rate), whereas in those with score ≤ 40, in 74% (13% + 39% + 17%) pre-emptive administration of analgesics are correctly withheld (specificity, true negative rate). Using a lower threshold at 35, the sensitivity was 80% and specificity 52%.

4. Discussion

The data from the present study demonstrate that the presence or absence of early severe postoperative pain in surgical patients after awakening from general anesthesia can be predicted using a limited set of simple preoperative patient characteristics, i.e. age, gender, surgical procedure, preoperative pain severity (measured by either the SF-36 pain domain or a verbal rating or VAS), expected incision size and the two APAIS scores. Although they were univariably associated with the outcome, more extensive preoperative measurements, such as the STAI and the SF-36 questionnaires, did not have *added* predictive value. The rule may help care givers to tailor individual analgesic management, and to assist high-risk patients with coping. In addition, there is a growing body of evidence that suggests that adequate pain relief in the immediate postoperative period may prevent the development of chronic postoperative pain (months to years after the operation) (Katz et al., 1996; Perttunen et al., 1999).

All selected predictors in our rule, including the patient's anxiety for the planned surgery, had been reported in previous studies, but their incremental (independent) value and their generalizability was largely unclear (Thomas et al., 1998; Morin et al., 2000; Bisgaard et al., 2001; Scott et al., 1983; Chung et al., 1997; Taenzer et al., 2000; Nelson et al., 1998; Keogh and Herdenfeldt, 2002). Apart from Chung et al. (1997), these studies investigated patients undergoing one specific type of surgery only. Chung et al. found the same predictors in unselected ambulatory surgical patients, except that they observed a higher incidence of severe pain in males and in patients with a higher body mass index. This may have been the result of the patient mix and type of procedures typically performed in ambulatory surgery and because they measured severe pain behaviorally rather than using a subjective pain scoring instrument.

It is interesting that preoperative pain was one of the best predictors of severe pain in the early postoperative period. There are several possible explanations for this observation. First, it could well be that chronic noxious afferent input from the area to be operated upon, has produced neuroplastic changes in the spinal cord (sensitisation by upregulation of receptor subsystems) that become manifest as a relatively hyperpathic state in the postoperative period. Second, it could also be that the patient's preoperative pain and focus on the operation—including possible anxiety and worry—are to some extent reflected in the preoperative pain score as well, given the smaller odds ratios of the two APAIS predictors compared to the odds ratio of the preoperative pain predictor. When patients report severe pain, caregivers sometimes tend to resort to psychological or emotional explanations for the pain. Although our data confirm that preoperative anxiety and worry as measured

Table 3
Empirical distribution of patients with and without severe pain within first hour at the recovery room, according to the score of the rule (and to the corresponding probability) as estimated by the final prediction model (Table 2)

Estimated score[a]	< 25	25–35	36–40	41–45	> 45	Total
Estimated probability[b]	<0.15	0.15–0.25	0.25–0.30	0.31–0.35	>0.35	
Observed % severe pain	7	13	28	39	49	
Severe pain	12 (3)	60 (17)	65 (19)	72 (21)	139 (40)	348 (100)
No severe pain	170 (17)	390 (40)	170 (17)	113 (11)	146 (15)	989 (100)
N	182	450	235	185	285	1337

Values are presented as absolute numbers (percentages of the 'Total' column).
N = number of subjects per score (probability) category.
[a] Categories of the score as estimated from the scoring rule (nomogram and Appendix-A).
[b] Probability or risk of severe pain as estimated by the final prediction model (Table 2) that correspond to the score.

with the APAIS questionnaire are significant, they account for relatively little variance when compared with the biomedical variables.

With respect to the patients' anxiety predictors we found that the APAIS, a six item preoperative anxiety score developed specifically to quantify preoperative anxiety, outperformed the more elaborate 2×20 item STAI. The APAIS was developed in the Netherlands by Moerman et al. (1996), and validated recently in the USA (Miller et al., 1999) and Japan (Nishimori et al., 2002). One possible explanation for this might be that STAI identifies a person's global anxiety trait and present anxiety state, but does not address coping strategies. In contrast, two of the six APAIS questions address information seeking behavior as one element of coping. The results of the present study suggest that the anxiety/'worry' component of the APAIS is positively associated with the occurrence of early postoperative pain, whereas a strong information seeking behavior reduces the incidence of severe postoperative pain. Munafo (1998) recently criticized the unitary conception of anxiety, because it inappropriately combines affective, cognitive and behavioral components. These components may influence the postoperative pain response in opposite ways. The present data are consistent with this hypothesis and suggest that future studies should focus on assessing these components of preoperative 'anxiety' separately. The fact that anxiety and coping behavior as measured with APAIS was a better predictor than STAI is also fortunate from a practical perspective, because the six APAIS questions can be easily incorporated into the routine preoperative questionnaire administered to the patient on the preoperative anesthesia clinic or via the Internet.

To further appreciate the present results a few issues need to be addressed. First, in Table 2, Fig. 2 and in Appendix-A we reported a scoring rule that can directly be applied to future surgical patient populations to preoperatively classify patients according to their *absolute* risk of postoperative pain. However, if the overall pain incidence is different from ours (26%), the intercept in Table 2 should be adjusted using standard techniques (Hosmer and Lemeshow, 1989; Poses et al., 1986; Wigton et al., 1986). Second, the presented estimates of the sensitivity and specificity of the scoring rule (Table 3) may appear rather low. However, it should be noted that these estimates apply to a prognostic setting and should not be confused with the much higher estimates obtained from a diagnostic setting. Third, to measure preoperative pain we used the pain domain of the SF-36 instead of the more commonly used VAS. The first can be considered as a reasonable proxy of NRS or VAS measures (Moseley et al., 2002; Ware and Sherbourne, 1992; Kantz et al., 1992). Hence, there is no reason to assume that our choice of preoperative pain measurement has affected the predictive ability of our rule. Fourth, we recoded patients who were sleeping or too drowsy to rate their pain as having no pain. Although this approach reflects clinical practice, it remains a matter of conjecture whether sleeping patients suffer from pain. For obvious ethical reasons we opted not to wake—by physical stimulation—patients who did not respond when asked about their pain. Fifth, we decided to use only data from the PACU period to determine the outcome 'severe pain', as the primary goal of the present study was to test the hypothesis that it is feasible to predict the probability of severe pain shortly after the surgical procedure. We refrained from using pain observations on the wards, since there were known differences in postoperative pain management on the wards in the hospital. Furthermore, we defined severe pain as an NRS-score of 8 or higher. We have repeated the entire analysis after defining the outcome 'severe pain' by an NRS-score of 7 or higher. Expectedly, the incidence increased to 41.3%, but the final predictors as well as the predictive accuracy of the model remained the same. Also, a small proportion of patients who were in apparent severe pain immediately after arrival at the PACU (3.5%) received morphine IV. As this could have influenced the results we repeated the analysis after exclusion of these patients, which yielded no different results. Finally, we used bootstrapping techniques to validate our final model rather than split-sample or cross-validation methods, because it has been shown in the statistical literature that the former is superior (Harrell et al., 1996; Efron and Tibshirani, 1993). With bootstrapping techniques all data is used to develop *and* validate a prediction model, obviously yielding better and more precise estimates of the predictor's odds ratios and the model's ROC area. However, as our data set was relatively large we also conducted the cross-validation method, in which 75% of the data (randomly selected) was used for model development (derivation set) and the remaining 25% for model testing (validation set). This was done four times and the results were averaged. The results were similar to the results obtained after bootstrapping. Exactly the same seven predictors were selected in each of the four final models with similar odds ratios, and the average ROC area of the four models when tested in the corresponding validation sets was 0.71 only with a wider 95% confidence interval: 0.64–0.78.

There is yet no agreement on the ideal way to measure postoperative pain. Postoperative pain has behavioral components that can affect the impression of pain severity by care givers, and some authors have used behavioral observations to assess the presence of severe pain rather than the scoring of pain intensity by the patient himself (Chung et al., 1997). For the assessment of pain by subjective self-report, verbal descriptors (adjective scale), a four-category verbal rating scale (VRS-4), 11-point-numeric rating scale (NRS-11) or a 100-mm VAS can be used. In general there is good agreement between the VAS scale and the 11-point-numeric rating scale (Breivik et al., 2000), although patients and nurses expressed a preference for verbal ratings (Briggs and Closs, 1999). However, the level of residual sedation after general anesthesia may influence the patients' ability to use such scales. The issue of optimal pain endpoints is clouded further by the emphasis

that some have recently put on measuring patient satisfaction rather than measures of pain severity, but patient satisfaction may be an unreliable indicator of effective pain relief (Sartain and Barry, 1999). Finally it has become clear that patient expectations also have a considerable impact on satisfaction with pain management (Svensson et al., 2001).

We have presented and discussed the results when using two scoring thresholds. The implementation of this scoring rule into practice will rest on finding a clinically tenable threshold, translating sensitivity and specificity into weights assigned to over- and undertreatment. For example, one might favor prevention of undertreatment (calling for a higher sensitivity) over prevention of overtreatment (higher specificity). Using Table 3 one could select alternative thresholds in the risk score leading to other percentages of correctly and incorrectly predicted patients. Definition of treatment thresholds requires a proper cost-effectiveness analyses accounting for the acceptable consequences and costs related to misclassifications. This was beyond the scope of this paper but requires further research.

In conclusion, we have demonstrated that the occurrence of severe postoperative pain early after awakening from general anesthesia after various types of surgery can be predicted with a scoring rule, using a small set of variables that can be easily obtained at the preoperative visit. Although we were able to quantify the robustness of our scoring rule using bootstrapping techniques, external validation studies in new patients in various clinical settings are necessary before application of this preoperative prediction rule into clinical practice.

Appendix-A

The assigned points per value of each predictor included in the final model (Table 2). The points are assigned by multiplying the adjusted regression coefficient (Table 2, third column) of each predictor times 15 and rounded to the nearest integer. The models' intercept (constant) is already converted in the score assigned to the predictor age.

Sex	Points	Preoperative Pain score	Points	Anxiety level (APAIS)	Points	Total points	Probability of postoperative pain
Male	0	0	0	4–5	0	0	0.03
Female	3	1	2	6–7	2	11	0.05
		2	4	8–9	3	22	0.10
Age (years)		3	6	10–11	5	34	0.20
15–19	17	4	8	12–13	6	41	0.30
20–24	16	5	10	14–15	8	48	0.40
25–29	15	6	12	16–17	9	53	0.50
30–34	13	7	14	18–19	11	59	0.60
35–39	12	8	16	20	12	65	0.70
40–44	11	9	18			73	0.80
45–49	10	10	20	Need for info (APAIS)			
50–54	9			2	9		
55–59	8	Surgery type	0	3	8		
60–64	7	Ophthamology	5	4	7		
65–69	6	Laporoscopy	8	5	6		
70–74	4	Ear/nose/throat	14	6	5		
75–79	3	Orthopedic	18	7	3		
80–84	2	Abdominal	7	8	2		
85–89	1	Other		9	1		
≥ 90	0		0	10	0		
		Incision size	3				
		Small					
		Medium-large					

Appendix A (*continued*)

References

Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A, Verrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055–68.

Airaksinen KE. Autonomic mechanisms and sudden death after abrupt coronary occlusion. Ann Med 1999;31:240–5.

Bisgaard T, Klarskov B, Rosenberg J, Kehlet H. Characteristics and prediction of early pain after laparoscopic cholecystectomy. Pain 2001; 90:261–9.

Breivik EK, Bjornsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. Clin J Pain 2000;16:22–8.

Briggs M, Closs JS. A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. J Pain Symptom Manage 1999;18:438–46.

Chapman CR. Psychological aspects of pain patient treatment. Arch Surg 1977;112:767–72.

Chatfield C. Model uncertainty, data mining, and statistical inference. J R Stat Soc 1995;158:419–66.

Chung F, Ritchie E, Su J. Postoperative pain in ambulatory surgery. Anesth Analg 1997;85:808–16.

Donovan M, Dillon P, McGuire L. Incidence and characteristics of pain in a sample of medical-surgical inpatients. Pain 1987;30:69–78.

Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York, NY: Chapman and Hall; 1993.

Ellstrom M, Olsen MF, Olsson JH, Nordberg G, Bengtsson A, Hahlin M. Pain and pulmonary function following laparoscopic and abdominal hysterectomy: a randomized study. Acta Obstet Gynecol Scand 1998; 77:923–8.

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–43.

Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.

Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley and Sons, Inc.; 1989.

Joris JL, Hinque VL, Laurent PE, Desaive CJ, Lamy ML. Pulmonary function and pain after gastroplasty performed via laparotomy or laparoscopy in morbidly obese patients. Br J Anaesth 1998;80:283–8.

Kantz ME, Harris WJ, Levitsky K, Ware Jr. JE, Davies AR. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. Med Care 1992;30:MS240–52.

Katz J, Jackson M, Kavanagh BP, Sandler AN. Acute pain after thoracic surgery predicts long-term post-thoracotomy pain. Clin J Pain 1996;12: 50–5.

Kehlet H, Rung GW, Callesen T. Postoperative opioid analgesia: time for a reconsideration? J Clin Anesth 1996;8:441–5.

Keogh E, Herdenfeldt M. Gender, coping and the perception of pain. Pain 2002;97:195–201.

Kotzer AM. Factors predicting postoperative pain in children and adolescents following spine fusion. Issues Compr Pediatr Nurs 2000; 23:83–102.

Lynch EP, Lazor MA, Gellis JE, Orav J, Goldman L, Marcantonio ER. The impact of postoperative pain on the development of postoperative delirium. Anesth Analg 1998;86:781–5.

Miller KM, Wysocki T, Cassady Jr. JF, Cancel D, Izenberg N. Validation of measures of parents' preoperative anxiety and anesthesia knowledge. Anesth Analg 1999;88:251–7.

Moerman N, van Dam FS, Muller MJ, Oosting H. The Amsterdam Preoperative Anxiety and Information Scale (APAIS). Anesth Analg 1996;82:445–51.

Moons KG, Grobbee DE. Diagnostic studies as multivariable, prediction research. J Epidemiol Community Health 2002;56:337–8.

Moons CKGM, van Es GA, Stijnen T, Bak AAA, Hofman A, Jonker JC, Habbema JDF, Grobbee DE. Efficiency optimization of the selection period in therapeutic trials. J Clin Epidemiol 1997;50:793–800.

Moons KG, van Es GA, Michel BC, Buller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999; 10:276–81.

Morin C, Lund JP, Villarroel T, Clokie CM, Feine JS. Differences between the sexes in post-surgical pain. Pain 2000;85:79–85.

Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002; 347:81–8.

Muller JE. Circadian variation and triggering of acute coronary events. Am Heart J 1999;137:S1–S8.

Munafo MR. Perioperative anxiety and postoperative pain. Psycho Health Med 1998;3:429–34.

Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM. Patient satisfaction after anaesthesia and surgery: results of a prospective survey of 10,811 patients. Br J Anaesth 2000;84:6–10.

Nelson FV, Zimmerman L, Barnason S, Nieveen J, Schmaderer M. The relationship and influence of anxiety on postoperative pain in the coronary artery bypass graft patient. J Pain Symptom Manage 1998;15: 102–9.

Nguyen NT, Lee SL, Goldman C, Fleming N, Arango A, McFall R, Wolfe BM. Comparison of pulmonary function and postoperative pain after laparoscopic versus open gastric bypass: a randomized trial. J Am Coll Surg 2001;192:469–76. discussion 476–7.

Nishimori M, Moerman N, Fukuhara S, van Dam FS, Muller MJ, Hanaoka K, Yamada Y. Translation and validation of the Amsterdam preoperative anxiety and information scale (APAIS) for use in Japan. Qual Life Res 2002;11:361–4.

Perttunen K, Tasmuth T, Kalso E. Chronic pain after thoracic surgery: a follow-up study. Acta Anaesthesiol Scand 1999;43:563–7.

Poses RM, Cebul RD, Collins M, Fager SS. The importance of disease prevalence in transporting clinical prediction rules. Ann Intern Med 1986;105:586–91.

Sartain JB, Barry JJ. The impact of an acute pain service on postoperative pain management. Anaesth Intensive Care 1999;27:375–80.

Scott LE, Clum GA, Peoples JB. Preoperative predictors of postoperative pain. Pain 1983;15:283–93.

Spielberger CD, Gorush RL, Lushene RE. State Trait Anxiety Inventory manual. Palo Alto, CA: Consulting Psychologists Press; 1970.

Svensson I, Sjostrom B, Haljamae H. Influence of expectations and actual pain experiences on satisfaction with postoperative pain management. Eur J Pain 2001;5:125–33.

Taenzer AH, Clark C, Curry CS. Gender affects report of pain and function after arthroscopic anterior cruciate ligament reconstruction. Anesthesiology 2000;93:670–5.

Thomas T, Robinson C, Champion D, McKell M, Pell M. Prediction and assessment of the severity of post-operative pain and of satisfaction with management. Pain 1998;75:177–85.

Visser K, Hassink EA, Bonsel GJ, Moen J, Kalkman CJ. Randomized controlled trial of total intravenous anesthesia with propofol versus inhalation anesthesia with isoflurane-nitrous oxide: postoperative nausea with vomiting and economic analysis. Anesthesiology 2001; 95:616–26.

Ware Jr. JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Med Care 1992;30:473–83.

Wigton RS, Connor Jl, Centor RM. Transportability of a decision rule for the diagnosis of streptococcol pharyngitis. Arch Intern Med 1986;146: 81–3.