CLINICAL RESEARCH STUDIES

Risk-adjusted analysis of early mortality after ruptured abdominal aortic aneurysm repair

Jeremy R. Harris, MD,[a] Thomas L. Forbes, MD,[a] Stefan H. Steiner, PhD,[b] D. Kirk Lawlor, MD,[a] Guy DeRose, MD,[a] and **Kenneth A. Harris, MD,[a]** *London and Waterloo, Ontario, Canada*

Purpose: **Ruptured abdominal aortic aneurysms (RAAAs) continue to result in early mortality in up to 50% of patients. Additionally, it remains difficult to compare outcomes given the variability in patient comorbidities and presentation. The purpose of this study was to describe an instrument that permits the prospective analysis of outcomes after RAAA repair while adjusting for the variability in preoperative risk.**

Methods: **Consecutive patients undergoing attempted open RAAA repair over a 5-year period (1999 to 2003) at our center were reviewed. Thirty-day or in-hospital mortality was the main outcome variable. Preoperative mortality risk was estimated for each patient by using a validated modification of the POSSUM scoring system (V-POSSUM). A risk-adjusted cumulative sum method (RA-CUSUM) was used to compare observed versus predicted outcomes by assigning a risk-adjusted score, based on log-likelihood ratios, to each patient. These scores were sequentially plotted with preset control limits to allow for "signaling" when results were substantially different from expected (doubling or halving of odds ratios).**

Results: **A total of 136 patients were reviewed, with an early mortality rate of 45.6%. V-POSSUM scores were accurate in predicting mortality for the entire cohort, with an observed-to-predicted mortality ratio of 0.92 ($P = .80$). Each patient's risk-adjusted score was plotted sequentially. In one segment of the resulting plot, the graph adopted a negative slope and crossed the lower control limit, indicating improved results compared with predicted.**

Conclusions: **V-POSSUM scores in this series accurately predicted early mortality after RAAA surgery. The RA-CUSUM method allows for the prospective evaluation of outcomes, while taking into account patient variability. In the current study, this resulted in the identification of a series of patients who had improved outcomes compared with predicted. (J Vasc Surg 2005;42:387-91.)**

Despite advances in surgery and in perioperative intensive care, ruptured abdominal aortic aneurysms (RAAAs) continue to be associated with a notable early risk of mortality. In many previously published studies, these mortality rates have varied depending on the degree of patients' preoperative comorbidities and hemodynamic instability. A prospective national study of ruptured aneurysms in Canada reported a 51.4% rate of early mortality, with such presenting factors as hypotension and elevated creatinine proving to be independent predictors.[1]

Although there are many single and multicenter reports of outcomes after RAAA repair, most are hindered by their retrospective nature. As such, their applicability to continuing quality assurance and practice audits is limited. Optimally, a prospectively utilized tool would identify unacceptable or improved results compared with an acceptable standard and result in changes in clinical practice. To this end, the cumulative sum failure method (CUSUM) has been used in a variety of clinical situations, including our group's experience with ruptured,[2] open,[3] and endovascular[4] aortic aneurysm repair.

The CUSUM method is useful, but it treats all patients identically and fails to take into account differences in clinical presentations and preoperative risk. This is especially important when the variability in presenting factors contributes to different preoperative risks of adverse outcomes, as with ruptured aneurysms. As a result, a risk-adjusted cumulative sum failure method (RA-CUSUM) was introduced with the ability to consider this variability in preoperative instability and risk.[5-7] It has recently been applied to our group's experience with elective open aneurysm repair.[8]

The purpose of the present study was to retrospectively apply this tool to the analysis of our center's contemporary experience with RAAAs while adjusting for the variability in patients' comorbidities and hemodynamic instability. After this retrospective application,

From the Division of Vascular Surgery, London Health Sciences Centre and The University of Western Ontario,[a] and the Department of Statistics and Actuarial Science, University of Waterloo.[b]

Competition of interest: none.

Additional material for this article may be found online at www.mosby.com/jvs

Reprint requests: Thomas L. Forbes, MD, Division of Vascular Surgery, London HSC, University of Western Ontario, 375 South St, N380, London, ON N6A 4G5, Canada (e-mail: Tom.Forbes@lhsc.on.ca).

0741-5214/$30.00

doi:10.1016/j.jvs.2005.05.042

the ability of this instrument to be used prospectively would be apparent.

PATIENTS AND METHODS

The vascular surgery database at our university-affiliated medical center was reviewed to identify those patients who underwent attempted open repair of a RAAA during a recent 5-year period (January 1, 1999 to December 31, 2003). All patients were included who were confirmed to have an acute hemorrhage from their infrarenal aneurysm, as determined by preoperative imaging or intraoperative findings. Excluded from analysis were those who died en route to our hospital or before an attempt at repair. Otherwise, all patients who presented to our center were offered repair. The main outcome variable was death during hospitalization or ≤30 days of operation. Through a retrospective review of the patients' hospital charts, various factors apparent upon presentation were recorded to allow for the use of the V-POSSUM scoring system.

V-POSSUM Scoring System. The Physiologic and Operative Severity Score for enUmeration of Mortality and Morbidity (POSSUM) scoring system was originally described in 1991.[9] Since that time, various adjustments have been made to facilitate its applicability to different subsets of surgical patients.[10,11] One of these modifications was undertaken by the Vascular Surgical Society of Great Britain and Ireland in developing a model specific for vascular surgery (V-POSSUM).[11] The original POSSUM model was derived by logistic regression of data obtained from a broad range of general surgery patients, but the modified V-POSSUM model was based solely on data of patients undergoing arterial surgery.

Throughout the many published modifications of the POSSUM scoring system, the data set obtained from each patient has remained identical while the regression analyses have changed. As originally described,[10] the data set consisted of 12 preoperative variables that are weighted and combined to form a physiologic score. A further six intraoperative variables are weighted and combined to form an operative severity score. These variables are listed in Table I. The physiologic and operative severity scores are then used in the appropriate POSSUM logistic regression equation to calculate predicted mortality rates for each patient.

In this study, the V-POSSUM equation was used to predict mortality of each patient undergoing attempted repair of a RAAA. The patients were grouped into risk categories by their V-POSSUM-predicted mortalities. In each risk grouping, actual mortality was compared with predicted mortality to determine observed-to-expected ratios. These were then analyzed for goodness of fit and the difference between proportions with χ^2 testing. Results were deemed statistically significant at $P < .05$.

Risk-adjusted cumulative sum failure method. The RA-CUSUM[5-8] chart, as has been described previously,[8] is fashioned by determining each patient's score, which depends on the estimated risk of early mortality, the observed outcome, and an alternative level of performance to be detected (see Appendix [online only] for formulae). In the present study each patient's risk was derived after the use of the V-POSSUM scoring system and subsequent allocation of an individual mortality risk. As each patient's mortality risk varied, the alternative level of performance was set as an odds ratio to detect a doubling or halving in the observed mortality rate compared with those expected. Although this threshold can be altered, a doubling or halving of the odds of mortality is viewed as a clinically significant outcome.

Table I. POSSUM physiology and operative score variables

Physiology score	*Operative score*
Age	Grade of operation
Cardiac signs	Number of procedures
Respiratory signs	Total blood loss
Systolic blood pressure	Peritoneal soiling
Pulse rate	Presence of malignancy
Glasgow Coma Score	Timing of operation
Serum urea	
Serum sodium	
Serum potassium	
Hemoglobin level	
White blood cell count	
Electrocardiogram	

POSSUM, Physiologic and Operative Severity Score for enUmeration of Mortality and Morbidity.

When the RA-CUSUM chart is being plotted to detect changes in the early mortality rate, patient scores associated with mortality are positive and those associated with survival are negative. With this risk-adjusted analysis the "penalty" for the death of a low-risk patient is larger than that for the death of a high-risk patient. To complete the design of the chart, a control limit is set. For the present study, designed to detect increases or decreases in early mortality rate, the control limit was set at +/−2.5.

The resulting graph includes two curves, with the chart designed to identify a decrease in surgical performance with increased mortality rates placed above that designed to identify an improvement in surgical performance with decreased mortality. On the RA-CUSUM chart, a negative slope on either curve represents improved results compared with those predicted, whereas deterioration in performance results in a positive slope.

The procedure is designed to "signal" when the plot falls below the lower control limit or above the upper control limit, indicating a halving or doubling of the odds ratio, respectively. As a form of continuing quality assurance, each signal in this process will prompt an investigation and evaluation of factors not included in the initial risk analysis that may be contributing to this deviation of observed outcomes from those predicted. The plot is reset to baseline zero after each signal to continue the quality assurance process for subsequent patients.

RESULTS

Between 1999 and 2003, 138 patients underwent attempted repair of a RAAA at our center by one of four

Table II. Patient demographics and V-POSSUM data upon patient presentation

Number of patients	*136*
Men	104 (76.5%)
Women	32 (23.5%)
Mean age (range)	73 (43-88)
Mean hemoglobin	109 g/L
Mean serum Na^+	138 mmol/L
Mean serum K^+	4.1 mmol/L
Mean serum urea	8.1 mmol/L
Mean heart rate	89 beats/min
Mean systolic blood pressure	128 mm Hg
Known cardiac comorbidity*	52.9%
Known respiratory comorbidity†	17.6%
Abnormal ECG‡	47.8%
Altered LOC (GCS < 15)	25.7%
V-POSSUM predicted mortality rate	49.6%
Observed early mortality rate	45.6%

ECG, Electrocardiogram; *GCS,* Glasgow coma scale; *LOC,* level of consciousness; *V-POSSUM,* Vascular-Physiologic and Operative Severity Score for enUmeration of Mortality and Morbidity.

*Cardiac comorbidities include hypertension, congestive heart failure, angina, and arrhythmia requiring anticoagulation.

†Respiratory comorbidities include obstructive or restrictive lung diseases.

‡Abnormal ECG includes atrial fibrillation, other abnormal rhythms, or evidence of ischemia or previous infarction.

vascular surgeons, and 1020 nonurgent AAA repairs were performed during the same period. The hospital charts of two patients were unavailable for review and were thus excluded from the analysis, resulting in a 136-patient cohort. All repairs were achieved via the standard open transperitoneal route. The early mortality rate was 45.6%.

Patient demographics and V-POSSUM data are summarized in Table II. This study cohort consisted of 104 men (76.5%) and 32 women (23.5%), with a mean age of 73 years (range, 43 to 88). Over half (52.9%) of the patients had a known cardiac comorbidity, and 17.6 % had known respiratory disease. Nearly half (47.8 %) had an abnormal electrocardiogram on presentation, and 25.7% had an altered level of consciousness.

The overall mortality rate predicted by the V-POSSUM scores was 49.6%. The comparison of V-POSSUM predicted mortality and observed mortality is summarized in Table III. Observed-to-expected ratios were calculated for each of the five separate risk categories. There was no statistically significant difference between these observed and predicted mortality rates ($P = .80$), providing a measure of the applicability of the V-POSSUM scoring system to our patient cohort.

Each patient's risk-adjusted score was then plotted sequentially using a two-sided risk-adjusted cumulative sum graph that is depicted in the figure. During the first part of the curve, there is a steady downward slope of the lower curve such that at patient 45, the plot crosses the lower control limit (–2.5), indicating a halving of the odds of mortality and improved results compared with those expected. After this signal, the graph was reset to zero, and there were no further signals for the duration of this study.

Table III. V-POSSUM predicted mortality versus actual mortality

Predicted mortality group	*Number of patients*	*Actual deaths*	*Predicted deaths*	*O:E ratio**
0%-20%	21	4	2.8	1.43
20%-40%	21	4	6.2	0.65
40%-60%	51	23	25.0	0.92
60%-80%	24	15	16.7	0.90
80%-100%	19	16	16.8	0.95
Total	136	62	67.5	0.92

V-POSSUM, Vascular-Physiologic and Operative Severity Score for enUmeration of Mortality and Morbidity; *O:E,* observed-to-expected.

*$P = .80$.

DISCUSSION

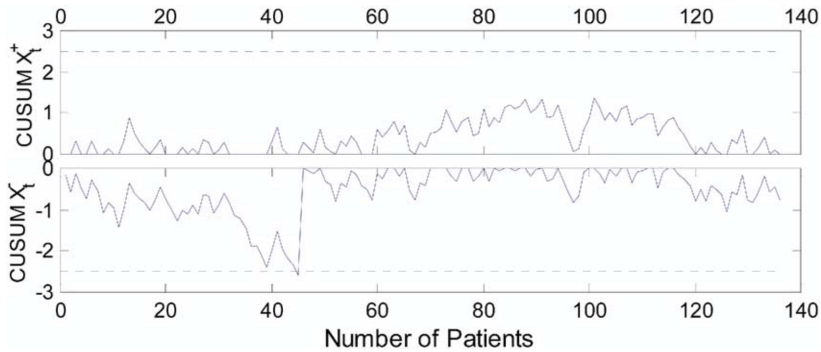
Early mortality after open RAAA repair remains disturbingly high, with recent reports confirming rates as high as 50%.[1,17] Unfortunately, during a time where major advances have been made in emergency medical transportation and in intensive care, a resulting improvement in outcomes after open repair has not been evident. Of course, comparing surgical outcomes by using crude mortality data, without taking into account variabilities in case mix, can be both misleading and unfair. Increasing interest in comparisons of hospital and even individual surgeons' performances, in addition to ongoing emphasis on continuing quality assurance, has highlighted the need for techniques to accurately compare outcomes after appropriate risk adjustment.

There have been innumerable attempts to devise scoring systems that are simple enough to be clinically relevant yet thorough enough to accurately predict patient outcomes. Two prime examples are the acute physiology and chronic health evaluation (APACHE) scoring system[18] and the Injury Severity Score[19] that are used in the intensive care and trauma settings. One of the more popular models seen in the surgical literature has been the POSSUM scoring system.[9]

Calculating a POSSUM score requires collecting 12 relatively routine preoperative data points and recording six intraoperative findings. Many of these factors, including but not exclusive to hypotension, elevated serum creatinine, and transfusion requirements, have been previously identified as independent predictors of mortality after RAAA repair.[1] These data are then used in previously derived logistic regression equations to provide estimates of the likelihood of morbidity and mortality.

Over the past decade, many studies have assessed the applicability of POSSUM-based models to particular subsets of surgical patients. As such, validated POSSUM-based models are now being used to assess outcomes in vascular surgery,[11-16] colorectal surgery,[20] head and neck surgery,[21] and orthopedic surgery.[22]

Multiple studies have confirmed the validity of various POSSUM models in predicting mortality after vascular surgery and, specifically, repair of RAAAs.[12-15] One con-



Rate-adjusted cumulative sum failure *(CUSM)* method of early mortality after ruptured abdominal aortic aneurysm repair, 1999-2003. The control limit *h* was set at ±2.5. The upper curve is set to "signal" when there is a doubling of the odds ratio of early mortality, and the lower curve is set to "signal" with a halving of the odds ratio. This plot "signals" at patient 45 by crossing the lower control limit. The graph is reset at zero and monitoring continues without any further "signaling".

trary study found that the described POSSUM models (P-POSSUM and V-POSSUM) predicted mortality following elective AAA repair but did not adequately predict mortality following repair of RAAAs.[16] The authors then proceeded to develop a further modification to the POSSUM scoring system designed specifically for ruptured aneurysms (RAAA-POSSUM).

In the present study, the V-POSSUM scoring system proved to be an accurate predictor of mortality in our patient cohort and formed the basis of the subsequent risk-adjusted analysis. The utility of these scoring systems is not in determining which patients should be offered or denied lifesaving aneurysm surgery, but rather is to permit an ongoing quality-assurance assessment of surgical practice and outcomes.

Validated methods of risk adjustment have allowed for more accurate comparisons of surgical outcomes, but the efficacy of discovering periods of unacceptable or improved outcomes years later through standard statistical analysis of retrospective data remains questionable. Optimally, a prospectively utilized tool is necessary for the identification of factors, irrespective of those specific to the patients themselves, affecting outcomes that would then result in changes to clinical practice. To this end, CUSUM has been used in a variety of clinical situations to detect periods of improved or deteriorating performance in a timely fashion.[2-4]

Combining a validated model of risk assessment (V-POSSUM) with a prospective analysis of results (CUSUM) allows for a powerful tool for effective practice audit and continuous quality assurance (RA-CUSUM). As a component of this instrument, signals and alarms prompt an investigation into possible contributing factors for the unexpected outcomes, which are irrespective of patient factors or comorbidities that are already considered by the prior risk adjustment.

In this study, a signal was given at patient 45, indicating that the cumulative results to that point were better than expected. Further examination of these data reveals one possible explanation for these findings. One of the four surgeons performing the RAAA repairs in this cohort of patients began independent practice 6 months before the beginning of the time period that was being analyzed. Previous work has used a CUSUM analysis to assess this surgeon's results after RAAA repair over the first 2.5 years of his career.[2] A learning curve was demonstrated in that study, with significantly better outcomes in the latter half of the study period. Interestingly, that period of decreased mortality coincides directly with the segment of time during which the RA-CUSUM analysis described here signaled improved results. Otherwise, when this method was applied to the individual surgeons' outcomes, there were no signals or deviation of outcomes from expected.

Shortcomings of this study include its reliance on retrospectively obtained data. This was necessary to confirm the internal validity of V-POSSUM in predicting early mortality after RAAA repair at our institution. Examinations of the external validity and reproducibility of this process and the use of other scoring systems in the model are ongoing. By building on this retrospective analysis, the RA-CUSUM technique can evaluate surgical outcomes in an ongoing, prospective manner.

The value of this type of analysis is in the early identification and subsequent correction of factors leading to unacceptable results. On the other hand, this tool can also be used to reinforce positive changes that have been shown to lead to improved outcomes. It is interesting to envision utilizing the RA-CUSUM methodology to test hypotheses regarding the reduction of mortality after RAAA repair. The necessarily earlier feedback derived from this type of analysis compared with standard statistical methods would certainly be beneficial in this regard.

Finally, in this era of increasing endovascular treatment of AAA, it is preferable to compare outcomes of endovascular repair versus standard open techniques on a risk-adjusted basis. Reports of impressive early outcomes with endovascular repair of RAAAs are becoming more frequent,[23] but it remains unclear whether this can directly be attributed to the less invasive nature of this technique, to a possible preselection of more stable patients, or both. Once again, RA-CUSUM could be used in this setting to evaluate an institution's early results with attempts at endovascular repair of RAAAs. This, along with the development of a corresponding preoperative risk measurement tool, is an ongoing area of study.

CONCLUSION

Effective practice audit and continuous quality assurance monitoring requires appropriate risk adjustment and the prospective evaluation of outcomes. In this study, V-POSSUM scores accurately predicted early mortality after RAAA surgery, and the RA-CUSUM method demonstrated the identification of a series of patients who had improved outcomes compared with predicted.

REFERENCES

1. Johnston KW. Ruptured abdominal aortic aneurysm: six-year follow-up results of a multicenter prospective study. Canadian Society for Vascular Surgery Aneurysm Study Group. J Vasc Surg 1994;19:888-900.
2. Forbes TL, DeRose G, Harris KA. A CUSUM analysis of ruptured abdominal aortic aneurysm repair. Ann Vasc Surg 2002;16:527-33.
3. Forbes TL. A cumulative analysis of an individual surgeon's early experience with elective open abdominal aortic aneurysm repair. Am J Surg 2005;189:469-73.
4. Forbes TL, DeRose G, Kribs SW, Harris KA. Cumulative sum failure analysis of the learning curve with endovascular abdominal aortic aneurysm repair. J Vasc Surg 2004;39:102-8.
5. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. Biostatistics 2000;1:441-52.
6. Steiner SH, Cook RJ, Farewell VT. Risk-adjusted monitoring of binary surgical outcomes. Med Decis Making 2001;21:163-9.
7. Cook DA, Steiner SH, Cook RJ, Farewell VT, Morton AP. Monitoring the evolutionary process of quality: risk-adjusted charting to track outcomes in intensive care. Crit Care Med 2003;31:1676-82.
8. Forbes TL, Steiner SH, Lawlor DK, DeRose G, Harris KA. Risk adjusted analysis of outcomes following elective open abdominal aortic aneurysm repair. Ann Vasc Surg 2005;19:142-8.
9. Copeland GP, Jones D, Walters M. POSSUM: a scoring system for surgical audit. Br J Surg 1991;78:356-60.
10. Whiteley MS, Prytherch DR, Higgins B, Weaver PC, Prout WG. An evaluation of the POSSUM surgical scoring system. Br J Surg 1996;83:812-5.
11. Prytherch DR, Ridler BM, Beard JD, Earnshaw JJ, on behalf of the Audit and Research Committee of the Vascular Surgical Society of Great Britain and Ireland. A model for national outcome audit in vascular surgery. Eur J Vasc Endovasc Surg 2001;21:477-83.
12. Shuhaiber JH, Hankins M, Robless P, Whitehead SM. Comparison of POSSUM with P-POSSUM for prediction of mortality in infrarenal abdominal aortic aneurysm repair. Ann Vasc Surg 2002;16:736-41.
13. Neary WD, Crow P, Foy C, Prytherch D, Heather BP, Earnshaw JJ. Comparison of POSSUM scoring and the Hardman Index in selection of patients for repair of ruptured abdominal aortic aneurysm. Br J Surg 2003;90:421-25.
14. Wijesinghe LD, Mahmood T, Scott DJA, Berridge DC, Kent PJ, Kester RC. Comparison of POSSUM and the Portsmouth predictor equation for predicting death following vascular surgery. Br J Surg 1998;85:209-12.
15. Midwinter MJ, Tytherleigh M, Ashley S. Estimation of mortality and morbidity risk in vascular surgery using POSSUM and the Portsmouth predictor equation. Br J Surg 1999;86:471-4.
16. Prytherch DR, Sutton GL, Boyle JR. Portsmouth POSSUM models for abdominal aortic aneurysm surgery. Br J Surg 2001;88:958-63.
17. Bown MJ, Sutton AJ, Bell PRF, Sayers RD. A meta-analysis of 50 years of ruptured aortic aneurysm repair. Br J Surg 2002;89:714-30.
18. Knaus WA, Zimmerman JE, Wagner DP. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. Crit Care Med 1981;9:591-7.
19. Baker SP, O'Neill B, Haddon W Jr. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma 1974;14:187-96.
20. Tekkis PP, Kessaris N, Kocher HM, Poloniecki JD, Lyttle J, Windsor AC. Evaluation of POSSUM and P-POSSUM scoring systems in patients undergoing colorectal surgery. Br J Surg 2003;90:340-5.
21. Griffiths H, Cuddihy P, Davis S, Parikh S, Tomkinson A. Risk-adjusted comparative audit. Is Possum applicable to head and neck surgery? Clin Otolaryngol Allied Sci 2002;27:517-20.
22. Mohamed K, Copeland GP, Boot DA, Casserley HC, Shackleford IM, Sherry PG, et al. An assessment of the POSSUM system in orthopaedic surgery. J Bone Joint Surg Br 2002;84:735-9.
23. Lee WA, Hirneise CM, Tayyarah M, Huber TS, Seeger JM. Impact of endovascular repair on early outcomes of ruptured abdominal aortic aneurysms. J Vasc Surg 2004;40:211-5.

Submitted Feb 1, 2005; accepted May 26, 2005.

APPENDIX (online only)

The rate-adjusted cumulate sum (RA-CUSUM) chart is fashioned by plotting X_t versus patient number t where $X_t = \max(0, X_{t-1} + w_t)$. The patient's score ($w_t$) depends on the estimated risk of early mortality (p_t), the patient's outcome (y_t), where $y_t = 0$ for survival and $y_t = 1$ for mortality, and OR_A, or an alternative level of performance to be detected. Each patient's risk, p_t, is derived after the V-POSSUM scoring system is used and the subsequent allocation of an individual mortality risk. As each patient's mortality risk varies, OR_A is set as an odds ratio.

Each patient's score (w_t) is derived by using a log-likelihood ratio. This is defined by the logarithm of the ratio of the probability of the observed outcome to that expected by the estimated risk, defined as p_t if the outcome is mortality and $(1 - p_t)$ if the outcome is survival.

$$w_t = log[OR_A/(1 - p_t + OR_A p_t)],\ \text{if}\ \ y_t = 1\ \ \text{(mortality)}$$

or

$$w_t = log[1/(1 - p_t + OR_A p_t)],\ \text{if}\ \ y_t = 0\ \ \text{(survival)}$$

To complete the design of the chart, a control limit (h) is set. For the present study, which was designed to detect increases or decreases in early mortality rate, the control limit (h) was set at ∀2.5. The procedure is designed to "signal" when X_t falls below the lower control limit ($h = -2.5$) or above the upper control limit ($h = 2.5$), indicating a halving or doubling of the odds ratio respectively.