Anesthesiology
1998; 88:589–600

Bayesian Modeling of Muscle Biopsy Contracture Testing for Malignant Hyperthermia Susceptibility

Julian C. P. Loke, David H. MacLennan†*

***Background*: Phenotyping malignant hyperthermia (MH) by contracture testing has a low but quantifiable degree of inaccuracy, measured by its sensitivity and specificity. Quantifying the limitations inherent in diagnostic testing for MH can help resolve issues in clinical practice, such as the interpretation of a negative test and the apparent lack of complete genetic linkage to *RYR1*.**

***Methods*: Bayesian models, mathematical descriptions of the outcome of diagnostic testing, were constructed. The inputs to the model include patient factors, summarized in a single number called pretest probability (PTP), and sensitivity and specificity that specify the accuracy of the entire test process. The outputs of the model include positive predictive value (PPV) and negative predictive value (NPV), which are numeric expressions of diagnostic certainty of positive and negative test results. A special case was constructed for equivocal results.**

***Results*: The PPV, NPV, and efficiency of contracture testing for MH are functions of PTP, sensitivity, and specificity. The NPV is high for all clinical PTP, whereas PPV is clinically useful for moderate to high PTP.**

***Conclusions*: Diagnostic contracture testing for MH is clinically useful because of high NPV and can exclude MH with near certainty. For MH probands, the clinical grading scale for MH may guide PTP estimation, whereas for relatives of probands, PTP is a function of kinship to a known MH-susceptible relative. A sequential testing strategy optimizes diagnostic information by maximizing PTP within a pedigree. Incomplete testing of parents of an MH susceptible child can pose a significant risk of false-negative results for the untested parent. Even with optimal pedigree testing strategies, the PPV drift effect results in a considerable source of phenotypic uncertainty for genetic linkage studies. (Key words: Bayes; caffeine; halothane; skeletal muscle.)**

MALIGNANT hyperthermia (MH) is an inherited, autosomal dominant, pharmacogenetic disorder in which susceptible patients can exhibit acute, rapidly progressive life-threatening hypermetabolic reactions when exposed to succinylcholine, potent inhalational anesthetic agents, or both.[1,2] In families with MH, prediction of MH susceptibility (MHS) before surgery is achieved by contracture testing of skeletal muscle biopsies for abnormal threshold sensitivity to caffeine[3] and to halothane.[2,4]

Two protocols exist for contracture testing for diagnosis of MHS. The North American Caffeine–Halothane Contracture Test (CHCT)[5] and the European *In Vitro* Contracture Test (IVCT)[6] differ only slightly in the testing technique. Both protocols assign a positive (MHS) and a negative (MHN) diagnosis, but the European protocol allows for an equivocal (MHE) classification that results in the same clinical treatment as a positive diagnosis does.[6] For linkage studies, the MHE classification is considered nondiagnostic.

The discovery of linkage of the calcium-release channel gene, *RYR1*, to human MH,[7,8] the demonstration that a point mutation in *RYR1* is tightly linked to the porcine form of MH,[9,10] and that nine point mutations in *RYR1* can be linked with human MH[11–19] has increased interest in the molecular genetics of human MH. Although only about 50% of human families with MH have a link to the region on human chromosome 19q13.1 containing *RYR1*,[20] many apparently unlinked families have only a single recombinant event.[21]

Several large European MH families, with no link to chromosome 19, have been included in a systematic linkage study using a set of polymorphic microsatellite markers covering the entire human genome. Linkage has been observed in single families for chromosome

* Lecturer, Department of Anesthesia, Malignant Hyperthermia Investigation Unit.

† University Professor, Banting and Best Department of Medical Research.

Received from Banting and Best Department of Medical Research, Charles H. Best Institute, and Department of Anesthesia, Malignant Hyperthermia Investigation Unit, University of Toronto. Submitted for publication April 3, 1997. Accepted for publication November 25, 1997. Supported by the Canadian Genetic Diseases Network of Centres of Excellence, the Muscular Dystrophy Association of Canada, and the Malignant Hyperthermia Association. Dr. Loke is a postdoctoral research fellow of the Medical Research Council of Canada. This work was the basis of a poster presentation at the Annual Meeting of the American Society of Anesthesiologists meeting in San Diego, California, October 18–22, 1997.

Address reprint requests to Dr. Loke: MH Investigation Unit, 200 Elizabeth Street, CCRW 2-830, Toronto, Ontario, Canada M5G 2C4. Address electronic mail to: jul.loke@utoronto.ca

3q13.1,[22] for chromosome 5p and for chromosome 1q.[23,24] Analysis of the chromosome 1q-linked family led to the discovery of a mutation in the *CACNL1A3* gene encoding the α-1 subunit of the dihydropyridine receptor[23] and established a second MH locus in which mutations can be linked to MH.

The major difficulty with linkage studies in human MH pedigrees is their reliance on accurate phenotypic data.[21] Like most other clinical diagnostic tests, phenotyping of MHS by contracture testing of a muscle biopsy has a quantifiable inaccuracy, measured by its sensitivity and specificity. The sensitivity of a diagnostic test is defined as the ability of the test to diagnose disease correctly when disease is known to exist.[25-27] Specificity refers to the ability of the test to exclude disease when disease is known to be absent.[25-27] Although false-negative results are rare[28-34] and false-positive results may occur,[30,31,33,35,36] a complete understanding of the nature and implications of these limitations inherent in diagnostic testing for MH is essential for good clinical practice and genetic linkage studies.

The North American Malignant Hyperthermia Registry has compiled figures on sensitivity and specificity of the CHCT.[37,38] The European Malignant Hyperthermia Group (EMHG) has also published data on sensitivity and specificity of the IVCT.[35,39,40] Both groups have used the International Malignant Hyperthermia Clinical Grading Scale (CGS),[41] devised by collaboration of representative MH clinicians from at least six countries, as a case-finding instrument when establishing sensitivity.

A Bayesian model is a mathematical description of how sensitivity and specificity affect the outcome of a clinical diagnostic test.[25-27] The other input to the model is *pretest probability* (PTP), which is defined as the probability, or clinical suspicion, that a person has disease before testing. The likelihood of disease is modified (conditioned) by the diagnostic test result: It is increased with a positive result and decreased with a negative result. The degree to which PTP is revised by a test result depends on the sensitivity and specificity of that diagnostic test, and it can be calculated.[25-27] The aim of this investigation was to develop a Bayesian model to calculate conditional probabilities starting from the current estimates of sensitivity and specificity of muscle biopsy contracture testing for MH. New values for sensitivity and specificity can be accommodated by simple recalculation.

Our Bayesian model assumes only that sensitivity and specificity are correct. Sensitivity and specificity are difficult to estimate because there is no "gold standard" for MHS. Exposing CHCT-positive patients to triggering anesthetics that provoke life-threatening reactions is not only unethical but may require several such anesthetics before confirming MHS.[2,31,33,36] Similarly, looking for anesthetic reactions in a highly selected group of CHCT-negative patients would give an estimate of sensitivity[28,29,31-33,36,42] but would be subject to spectrum bias, test-retest bias, and selection bias, which are difficult to overcome. Analysis of data that were used to study the agreement between CGS and IVCT from two European test centers[43,44] can also yield estimates of sensitivity (1.00) and specificity (0.53-0.66) that differ from the pooled EMHG estimates, suggesting that such biases may truly exist.

Table 1. Sensitivity and Specificity for Currently Used Diagnostic Protocols for Muscle Biopsy Contracture Testing. (95% Confidence Intervals)

	NAMHR CHCT[38]	EMHG IVCT[40]
Sensitivity (95% CI)	0.97 (0.84-1.00)	0.99 (0.948-1.00)
Specificity	0.78 (0.71-0.85)	0.936 (89.2-96.5)

NAMHR = North American Malignant Hyperthermia Registry; CHCT = caffeine-halothane contracture test; EMHG = European Malignant Hyperthermia Group; IVCT = *in vitro* contracture tests.

The effect of inaccuracies in estimating sensitivity and specificity on the Bayesian model can be examined, and robust conclusions can be drawn. If the values for sensitivity and specificity for contracture testing are accepted as published, then the conclusions from this manuscript must hold.

Although limited analyses of such data have been done,[30,35] our more extensive mathematical model can examine the effect of proband CGS rank and of various pedigree testing strategies. This allows us to propose rational applications of contracture testing data, clinically and in future genetic linkage research, by quantifying the degree of uncertainty in phenotypic diagnosis.

Methods

Table 1 summarizes sensitivity and specificity figures for the North American Malignant Hyperthermia Registry[38] and the EMHG.[40] For the North American Malignant Hyperthermia Registry CHCT, the sensitivity is 0.97 and specificity is 0.78. For the EMHG IVCT, the sensitivity is 0.99, and the specificity is 0.936.

Table 2. Bayesian Model of Contracture Testing for Diagnosis of MH, Using a 2 × 2 Contingency Table, Including an Example Calculation for PPV, NPV, and Efficiency Using Sensitivity = 0.97 and Specificity = 0.78, Which Are Fixed Properties of the NAMHR CHCT

	MH	Normal	
CHCT positive	True positive (TP) = Sensitivity × PTP = *0.97 × 0.4* = *0.388*	False positive (FP) = (1 − PTP) − TN = *0.6 − 0.468* = *0.132*	Positive predictive value (PPV) = TP/(TP + FP) = *0.388/(0.388 + 0.132)* = *0.746*
CHCT negative	False negative (FN) = PTP − TP = *0.4 − 0.388* = *0.012*	True negative (TN) = Specificity × (1 − PTP) = *0.78 × (0.6)* = *0.468*	Negative predictive value (NPV) = TN/(FN + TN) = *0.468/(0.012 + 0.468)* = *0.975*
	PTP = *0.4*	1 − PTP = *0.6*	Efficiency = (TP + TN)/(TP + TN + FP + FN) = *(0.388 + 0.468)/(0.388 + 0.468 + 0.132 + 0.012)* = *0.856*

MH = malignant hyperthermia; PTP = pretest probability; TP = true positive; TN = true negative; FP = false positive; FN = false negative; PPV = positive predictive value; NPV = negative predictive value; NAMHR = North American Malignant Hyperthermia Registry; CHCT = caffeine–halothane contracture test; Efficiency = (TP + TN)/(TP + TN + FP + FN). An example calculation is shown in italics, where PTP, the input variable, is set to 0.4. TP and TN are calculated first. These values are used in the calculation of FN and FP. PPV is calculated as the ratio of TP to all positives (TP + FP). If the patient has a positive result, the certainty that this is a true positive is 0.746. NPV is similarly calculated as the ratio of true negatives to all negative (TN + FN) results. Given a negative result, the certainty that this is a true negative is 0.975. The overall efficiency of the test is 0.856.

Construction of Bayesian Models

Construction of a 2 × 2 contingency table using current estimates of sensitivities and specificities allowed the calculation of conditional probabilities associated with varying PTPs as the input variable. The positive predictive Value (PPV) is the calculated conditional probability of a true positive (TP), given a positive test result. The negative predictive value (NPV) is the conditional probability of a true negative (TN), given a negative test result. The PPV and NPV are calculated values that depend on fixed sensitivity and specificity and variable PTP. The efficiency of testing is calculated as the proportion of correctly assigned results (TP + TN) over all results, including false-positive (FP) and false-negative (FN) results (table 2).

In the EMHG "clinical" IVCT, MHE results in the same treatment as MHS: Where a more "specific" IVCT result is required, such as for genetic linkage studies, MHE results are considered nondiagnostic. Accordingly, a 3 × 2 contingency table was constructed to model the "specific" IVCT. Equivocal predictive value is the calculated conditional probability of MH given an MHE result. In this case, efficiency is calculated as the proportion of correctly assigned results (TP + TN) over all results, which now also include equivocal positive (EP) and equivocal negative (EN) results (table 3). The MHE rate was 0.14 in probands, and 0.05 in control patients.[35,40]

All contingency tables were coded with computer spreadsheet software (Borland's Quattro Pro 5; Microsoft's [Redmond, WA] Excel, v. 7.0) and modeled with PTP values from 0 to 1. Selected values were checked by manual calculation. Results were plotted using Origin 3.73 (MicroCal, Northampton, MA).

Results

Figure 1 illustrates the effect of PTP on PPV, NPV, and efficiency of diagnostic contracture testing for MH. In all cases, for PTPs between 0 and 0.75, NPV remains near 1.00. A high NPV is required clinically so that the risk of a false-negative result is acceptably low. The PPV is a measure of the degree of certainty that can be associated with a positive test result. It is a function of PTP and specificity. When PPV is low, such as at low PTPs, a positive result is likely to be falsely positive.

In the case of the EMHG, the "specific" interpretation of the IVCT produces a higher PPV and NPV (fig. 1B). The graph of equivocal predictive value against PTP is close to the diagonal, confirming that MHE alone is a poor predictor of MH. The MHE classification implies

Table 3. Bayesian Model of Alternative Interpretation of IVCT for Diagnosis of MH, Using a 3 × 2 Contingency Table

	MH	Normal	
IVCT positive	True positive (TP) = Sensitivity × PTP	False positive (FP) = (1 − PTP) − TN − EN	Positive predictive value (PPV) = TP/(TP + FP)
IVCT equivocal	Equivocal positive (EP)	Equivocal negative (EN)	Equivocal predictive value (EPV) = EP/(EP + EN)
IVCT negative	False negative (FN) = PTP − TP − EP PTP	True negative (TN) = Specificity × (1 − PTP) 1 − PTP	Negative predictive value (NPV) = TN/(FN + TN)

IVCT = *in vitro* contracture test; MH = malignant hyperthermia; PTP = pretest probability; TP = true positive; TN = true negative; FP = false positive; FN = false negative; PPV = positive predictive value; NPV = negative predictive value; EP = equivocal positive; EN = equivocal negative; EPV = equivocal predictive value; Efficiency = (TP + TN)/(TP + TN + FP + FN + EP + EN).

that the likelihood of MH is unchanged by the test and remains at the PTP. The increased proportion of MHE patients who are left undiagnosed contributes to a reduction in efficiency by nearly 0.2 at high PTPs.

Statistical Imprecision in Estimating Sensitivity and Specificity

The PPV is likely to be affected by any imprecision, expressed as the 95% confidence interval, in estimating specificity (table 1). However, even a major change in specificity from 0.82 to 0.92, within the 95% confidence, has a minor effect on the shape of the PPV curve, which retains an elbow at PTP between 0.125 and 0.25 (fig. 2), whereas NPV and efficiency are relatively unaffected. Conclusions about PPV are likely to be reliable for all PTPs.

In contrast, imprecision in estimating sensitivity has a profound effect on NPV at high PTP. When sensitivity is near 1, the NPV of the test will remain high at all PTPs (fig. 1B). However, as little as a 0.01 to 0.03 change in the estimate of sensitivity results in a dramatic change at high PTP from a nearly straight line (fig. 1B) to a curve (fig. 2B), which is similar to the NPV curve for the CHCT (fig. 1A), with little change in PPV and efficiency. Conclusions about NPV at high PTPs should account for the decrease in NPV.

Discussion

The Bayesian model, used in this analysis, does not depend on the mechanism of the test procedure or the

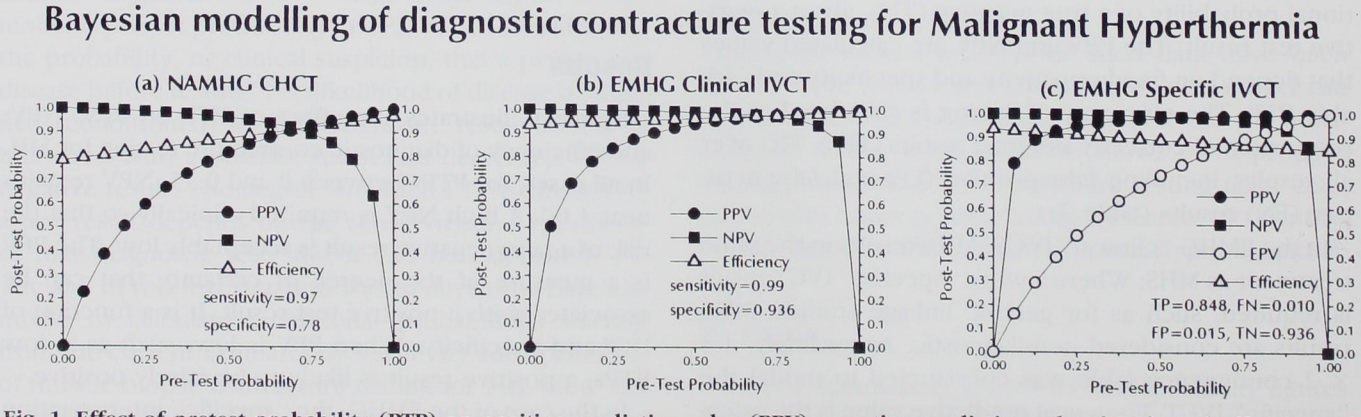

Fig. 1. Effect of pretest probability (PTP) on positive predictive value (PPV), negative predictive value (NPV), and efficiency of diagnostic muscle biopsy contracture testing for malignant hyperthermia (MH) susceptibility. Separate graphs are shown for (*A*) the North American Malignant Hyperthermia Registry (NAMHR) Caffeine–Halothane Contracture Test (CHCT), (*B*) the European MH Group (EMHG) *in vitro* contracture test (IVCT), and (*C*) the alternative "specific" interpretation of the EMHG IVCT in which MHE is considered nondiagnostic. The graph of equivocal predictive value (EPV) shows that MHE alone is not a good diagnostic indicator.

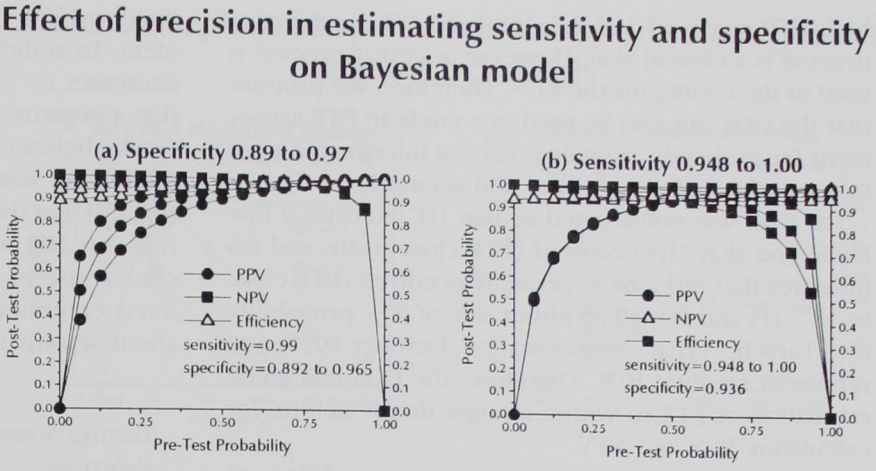

Fig. 2. (*A*) The effect of statistical inaccuracy in estimating specificity over a fractional range of 0.892–0.965, representing the 95% confidence interval for the European MH Group (EMHG) *in vitro* contracture test (IVCT). Positive predictive value is affected slightly but retains an elbow near the pretest probability (PTP) of 0.125–0.25. (*B*) Inaccuracy, within the 95% confidence interval for IVCT, in estimating sensitivity has a profound effect on negative predictive value (NPV) at high PTPs. A sharp elbow exists at PTPs near 0.8, with NPV decreasing at high PTPs.

disease and can be applied where there is a discrete outcome (*e.g.*, health and disease). There is no assumption about probability distributions, so the model is still valid if MH diagnostic tests do not have a bimodal frequency distribution[45] and even if strongly positive results occur in healthy persons for any reason unrelated to MH.

An alternative approach, using pre-test odds and likelihood ratios, although more cumbersome, is also possible. It is less suitable, however, because contracture test results are reported as discrete outcomes (MHS and healthy) rather than as ranks of likelihood (*e.g.*, a scale of 1 to 10). Because all current clinical testing protocols report discrete outcomes, Bayesian modeling is appropriate.

Sensitivity and specificity refer to the ability of diagnostic testing to discriminate between healthy persons and patients at risk for life-threatening reactions to anesthesia. There was no attempt in this model to characterize discrimination for less typical reactions, such as isolated masseter muscle rigidity, sudden hyperkalemic cardiac arrest, or events not related to anesthesia.

Our model is valid as long as estimates of sensitivity and specificity are accurate. Values from the North American Malignant Hyperthermia Registry and the EMHG (table 1) indicate general agreement for sensitivity and specificity. In the event that more accurate estimates of sensitivity and specificity become available in the future, the models can be recalculated.

Any bias during measurement of sensitivity, such as unintentional selection or incomplete case finding would affect NPV at high PTPs. A potential source of bias lies in the selection of patients for study based on diagnostic test results. Strict inclusion and exclusion criteria that do not depend on diagnostic test results should be satisfied to avoid over or underestimates based on unintentional bias.

Clinical history was used to define control and proband groups for estimating sensitivity and specificity.[38,40] The validity of CGS, as applied to probands, relies heavily on laboratory measurements made during the management of an anesthetic reaction. It is considered valid because only those patients tested since current protocols were established can be included for analysis, and only those patients with severe reactions were included in the proband group.

The model can regard MH as a single disorder or as a group of disorders that share the features of a life-threatening anesthetic reaction, identifiable by the CGS and by contracture testing. Acute MH can progress so rapidly that any delay in treatment can result in severe morbidity and death despite valiant measures.[2] Aborting a reaction early with dantrolene treatment is lifesaving and also the standard of care. To determine whether MH might demonstrate a spectrum of severity if left untreated is highly unethical. For practical purposes, human MH is an all-or-none phenomenon.

Estimating Pretest Probability for Probands

Referral of patients for contracture testing occurs after they have survived an anesthetic reaction that suggests MH. After exclusion of differential diagnoses such as sepsis, thyroid storm, and pheochromocytoma, features of the anesthetic reaction may be reviewed and recognized as typical of MH (*e.g.*, generalized rigidity, fever, family history, elevated creatine kinase) or atypical (*e.g.*, lack of metabolic acidosis or absence of myoglobinuria). On this basis, a clinician can assign low, medium, and

high PTP status of MH in a proband. This subjective process is a clinical skill. However, a similar process is used in the coding for the CGS. Therefore, we propose that the CGS can also be used as a guide to PTP assessment for probands, provided that the information used to score the CGS is complete and accurate.

The CGS was constructed so that D1 indicates a low likelihood that MH occurred (PTP close to 0), and D6 indicates that MH almost certainly occurred (PTP close to 1).[41] D3 and D4 fall on either side of 50% probability (MG Larach, verbal communication, October 1997) and represent medium PTP. Therefore, the CGS can guide estimation of PTP to within a range that is helpful for calculating PPV and NPV.

Separate conclusions can be drawn for conditions of low, medium and high PTP. Provided that sufficient data were collected to allow accurate CGS grading, patients with CGS D1 or D2 will have low PTP, between 0 and 0.25. An example of an anesthetic event ranking CGS D2 is unexplained sinus tachycardia in the absence of electrolyte abnormalities, muscle breakdown, hypercapnia, acidosis, or fever.

For low PTP patients, a negative test result would effectively rule out MH, since NPV approaches 1.00. The PPV, however, is only 0.50 or lower, so a false-positive result becomes as likely as, or more likely than, a true-positive result. Although a negative result would be useful, there would be no simple way to distinguish true positive results from false-positive results. Deciding whether to test remains a clinical decision that can only be determined in each individual case.

When a patient is referred for consultation after D3, D4, or D5 reactions, there is medium PTP, resulting in PPV in the range of 0.70–0.95, and NPV approaching 1.00. This is the situation in which the risk of observing a false-negative result is low, with an acceptable false-positive risk. Thus test results are useful, and testing is clinically indicated. Such a situation would occur, for instance, in a case of isolated masseter muscle rigidity (CGS D3).

Patients with CGS D6 will have high PTP. An example of a D6 reaction is an anesthetic event with generalized muscle rigidity, muscle breakdown with a creatine kinase level >20,000 U/l and myoglobinuria, inappropriately elevated end-tidal carbon dioxide, sinus tachycardia, rapid inappropriate increase in body temperature, severe metabolic acidosis, and rapid reversal of respiratory and metabolic acidosis with dantrolene. Based on the CGS, such an episode is "almost certainly MH."

A positive test result in the setting of high PTP produces a high PPV but does not alter clinical management. In addition, in cases of high PTP (>0.85), NPV decreases to <0.80 for the CHCT (fig. 1), so the risk that a negative result is false increases to 0.20 or more. Similar behavior can be seen within the 95% confidence interval for sensitivity of the EMHG IVCT at high PTPs (fig. 2B). For these patients, a negative test result cannot rule out MH with acceptable certainty, defeating the clinical use of contracture testing. They should be offered contracture testing only to provide useful data about sensitivity.

Testing Strategy Optimizes Pretest Probabilities for Relatives

Based on the assumption that MH is inherited as an autosomal dominant trait,[7,11–18,20,21] PTP is a function of the degree of certainty of MH in the proband and the degree of kinship to the closest related proband. Unless a relative has other clinical features (such as having suffered an anesthetic reaction suggestive of MH, elevated resting creatine kinase level, identical siblings, or consanguinity), PTP is fixed at a ratio of 0.5 for first-degree relatives, 0.25 for second-degree relatives, 0.125 for third-degree relatives, and so on. Any uncertainty in the diagnosis of MH in the proband, as indicated by the PPV for the proband, reduces the PTP for the relatives of that proband proportionately.

A Markov chain, a mathematical method of calculating dependent probabilities,[46] was constructed (described subsequently) to model the effect of different pedigree testing strategies. The calculations in the text below apply to the North American Malignant Hyperthermia Registry CHCT (figs. 3A and 3B), and calculations for the EMHG IVCT are depicted in figures 3C and 3D. The same conclusions were obtained for the two tests. For illustrative purposes, proband PTP is set at 1.

A "haphazard" testing strategy describes the situation in which a patient is tested without having first tested the intermediary relatives, leaving gaps in phenotyping of the pedigree (figs. 3A and 3C). For each patient, PTP must be assigned only on the basis of the degree of kinship to the proband. For first-, second-, and third-degree relatives, PTP is sufficiently high that PPV and NPV are informative. For more distant relatives, however, the PPV decreases to <0.5 for the CHCT and the IVCT, although the NPV is high. A PPV of 0.5 indicates moderate risk of MH, and it is sufficiently high that triggering drugs should be avoided during anesthesia. However, the high likelihood of observing a false-posi-

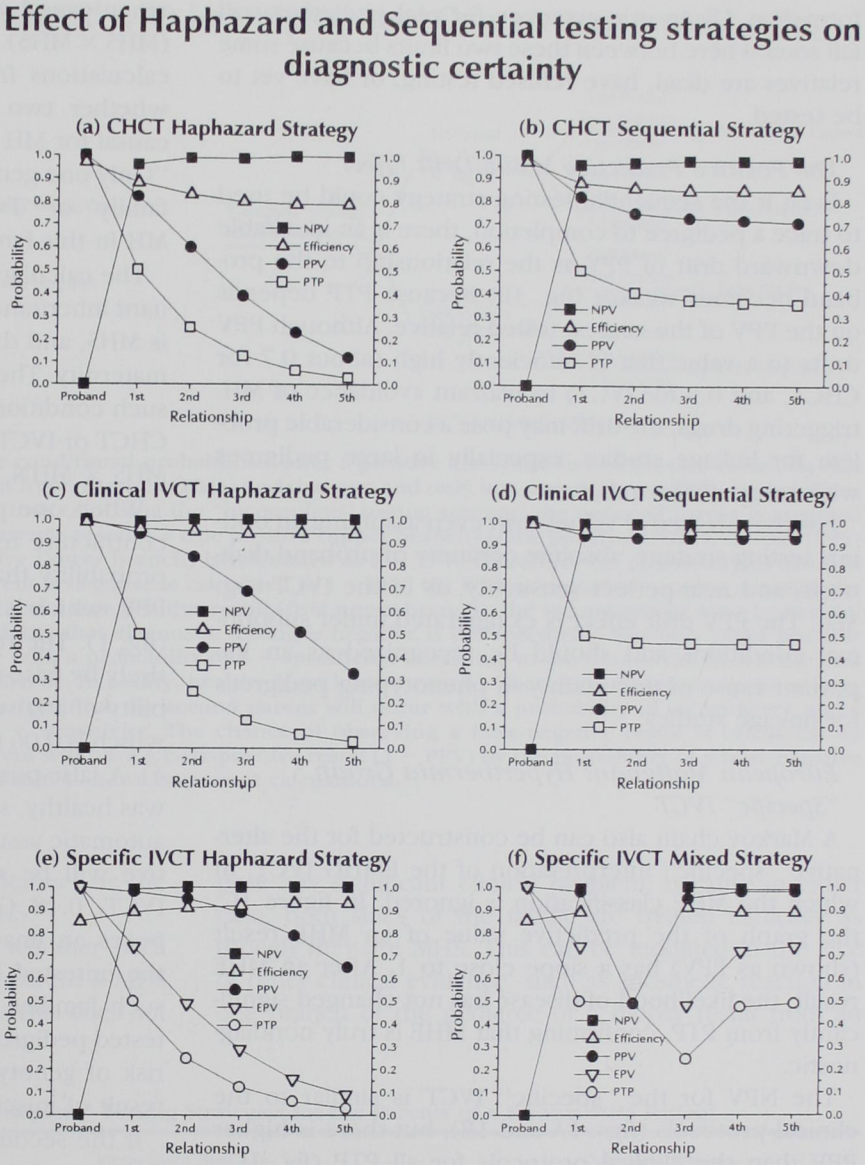

Fig. 3. Effect of testing strategy on pretest probability (PTP), positive predictive value (PPV), and negative predictive value (NPV) for the North American Malignant Hyperthermia Registry (NAMHR) caffeine–halothan contracture test (CHCT; sensitivity = 0.97, specificity = 0.78), the European MH Group (EMHG) *in vitro* contracture test (IVCT; sensitivity = 0.99, specificity = 0.936), and the EMHG "specific" IVCT (true positive = 0.825, false negative = 0; false positive = 0.025, true negative = 0.87). The "haphazard" testing strategy (*A*, *C*, *E*) uses a PTP equal to one half of the PTP from the previous generation, resulting in a dramatic decrease in PPV. In contrast, the "sequential" testing strategy (*B*, *D*) optimizes PTP based on the PPV of the previous generation and limits PPV drift to 0.73 for the CHCT, and 0.93 for the IVCT. In panels *E* and *F*, the proband PTP was set to 0.4, and a gap in the pedigree was produced when a person tests as malignant hyperthermia equivocal or has not been tested. A lower PTP must be assigned, based on the degree of kinship in this "mixed" testing strategy, producing a further decrease in PPVs. For patients other than the proband, NPV approaches 1.00.

tive reaction poses a significant problem for genetic linkage studies.

An alternative to the "haphazard" strategy would be to test close relatives of the proband first and then to offer testing to the nearest relatives of a person who tested positive. Our model confirms that such a "sequential" testing strategy can optimize the information that can be gained from contracture testing (Figs. 3B and 3D) and that it does so by optimizing PTP.

For first-degree relatives of a proband, PTP is 0.5, producing a PPV of 0.82. First-degree relatives have a PTP of one half of 0.82; *i.e.*, 0.41 (fig. 3B). If the initial first-degree relative had not been tested (as in the case of a "haphazard" testing strategy), the PTP for the second-degree relative would be only 0.25, with a PPV of only 0.60 (fig. 3A). By comparison, "sequential" testing of this second-degree relative results in a higher PPV of 0.75 (fig. 3B). This effect is even more pronounced for more distant relatives.

The sequential testing strategy maximizes PTPs, producing the best-case performance for PPV, whereas a haphazard strategy uses PTPs based solely on degree of kinship, resulting in worst-case performance. These represent the upper and lower limits of diagnostic per-

formance. Diagnostic accuracy for real pedigrees will fall somewhere between these two limits because some relatives are dead, have refused testing, or have yet to be tested.

The Positive Predictive Value Drift Effect

Even if the sequential testing strategy could be used to trace a pedigree to completion, there is an inevitable downward drift of PPV as the relationship to the proband becomes weaker (fig. 3B) because PTP depends on the PPV of the nearest tested relative. Although PPV drifts to a value that is sufficiently high (about 0.7 for CHCT, and 0.9 for IVCT) to warrant avoidance of MH-triggering drugs, the drift may pose a considerable problem for linkage studies, especially in large pedigrees with many distant relatives.

Positive predictive values drift even assuming an optimal testing strategy, absolute certainty of proband diagnosis, and near-perfect sensitivity, as in the IVCT (fig. 3D). The PPV drift effect is exaggerated under suboptimal conditions and should be recognized as an important cause of uncertainty in phenotyping pedigrees for linkage studies.

European Malignant Hyperthermia Group "Specific" IVCT

A Markov chain also can be constructed for the alternative "specific" interpretation of the EMHG IVCT, in which the MHE classification is ignored. In figure 1C, the graph of the predictive value of an MHE result (shown as EPV) has a slope close to 1. After an MHE result, the likelihood of disease has not changed significantly from PTP, confirming that MHE is truly nondiagnostic.

The NPV for the "specific" IVCT is similar to the clinical protocols (figs. 1A and 1B), but there is higher PPV than the clinical protocols for all PTP (fig. 1C), thereby producing more certain phenotypic information for genetic linkage studies. However, the price paid for this is lower efficiency because a nondiagnostic MHE is likely to be observed, especially at moderate-to-high PTP (fig. 1C). If a MHE results, the PTP remains unchanged, just as if testing had not occurred. When using the "specific" IVCT, any MHE patients leave a gap in phenotyping a pedigree, producing dips and drift in PPV (fig. 3F).

Testing Strategy for Parents of a Child Suspected to Have Malignant Hyperthermia

There is difficulty in interpreting test results for parents of a child suspected to have MH,[47] particularly when causal genes might originate from both parents (MHS × MHS). We examined a hypothetical family using calculations from Bayesian modeling. In determining whether two different genes from two parents are causal for MH in a family, hypotheses can be stated as, "Only one gene from one parent is causal of MH in this family" or "Two genes from two parents are causal of MH in this family."

The calculations that follow assume autosomal dominant inheritance,[7,11-18,20,21] that one and only one parent is MHS, and that there is no dispute about paternity or maternity. The PTP for a parent is set at 0.5. Even under such conditions, we show that it is not possible to use CHCT or IVCT to distinguish between MHS × MHS and MHS × MHN.

When one parent of an MHS patient has a positive contracture test and the other parent is untested, the probability that the result is a true positive is given by PPV, which is 0.94 for the IVCT and 0.82 for the CHCT (fig. 1). The remaining probability is accounted for entirely by false-positive results (fig. 4A). Thus, the probability of a false-positive result is calculated as 1 − PPV, which is 0.06 for the IVCT and 0.18 for the CHCT (table 4). A false-positive result implies that the tested parent was healthy, so the untested parent must be MHS. The automatic assumption that the untested parent is negative will be wrong with this same probability (0.06 IVCT, 0.18 CHCT). "Incomplete" testing, therefore, poses an unacceptable risk of incorrectly identifying the untested half of the pedigree in 1 in every 6-16 such families (table 4). Further, if these incompletely tested pedigrees were subjected to genetic analysis, the risk of genotype-phenotype discordance is high as a result of misclassification.[47]

If the second parent is tested (fig. 4B), the risk of a false-negative diagnosis for that parent is calculated as (1 − PPV)(1 − sensitivity), which is 0.0006 for the IVCT and 0.0054 for the CHCT (table 4). Based on this, a complete testing strategy, in which testing is offered to both parents, is far safer clinically than an incomplete strategy (table 4).

Given the condition that one and only one parent is MHS, the probability that both parents will test positive by chance alone is calculated as sensitivity × (1 − specificity), or 0.063 for IVCT and 0.21 for CHCT (see table 4). This probability, 1 in every 5-15 such families so tested, is sufficiently high to account for reported cases of apparent MHS × MHS without having to invoke the possibility that both parents are MHS, a spontaneous

Conditional probabilities associated with Incomplete and Complete parental testing

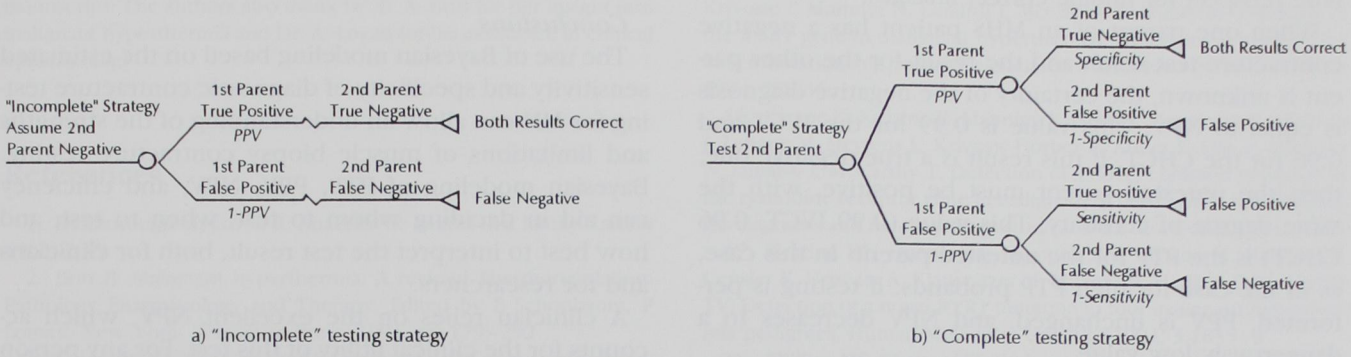

Fig. 4. A decision tree can be used to calculate conditional probabilities after a positive test result in one parent, assuming that a child is known to be suspectible to malignant hyperthermia (MHS) and that one and only one parent can be MHS. Probabilities of a branch are indicated by italics below the line. (*A*) Using the "incomplete" testing strategy, the untested parent is assumed to be negative. The probability that the first parent's result is a true positive (upper branch) is the positive predictive value (PPV; italics). The probability that it is a false positive (lower branch) is calculated as 1 − PPV. Therefore, the probability of error in the automatic assumption that the second parent is negative is calculated as 1 − PPV. (*B*) With a "complete" testing strategy, the second parent is also tested. If the first parent had a true-positive result (first upper branch), the second parent must be healthy under our assumptions. Therefore, the probability that diagnostic testing is negative is the specificity. The only other possible result is a false-positive one, which will occur with a probability of 1 − specificity. Returning to the root of the tree, if the first parent had a false-positive result, he or she must be, in reality, healthy (first lower branch). Therefore, under our assumptions, the second parent must be MHS, so a positive result for the second parent will occur with a probability of (sensitivity), and a false negative result with a probability of 1 − sensitivity. The chance of observing a false-negative result is calculated by multiplying the probabilities that the first parent will have a false-positive result (1 − PPV) by the probability of a false-negative result in the second parent (1 − sensitivity). Table 4 shows important calculations.

mutation rate, alternate modes of inheritance, or the existence of alternate genetic loci for MHS.[47]

There is no simple way to determine whether both parents are truly MHS or whether only one parent is MHS and the other has a false-positive diagnosis. This does not pose any significant clinical problem, because in either case, both sides of the family are treated clinically as though they are MHS. This can be modified in the face of other clinical evidence, such as anesthetic reaction in one branch of the pedigree or a biopsy result from an

Table 4. Comparison of "Complete" and "Incomplete" Testing Strategies for the Parents of a Known MHS Patient

	Incomplete Strategy	Complete Strategy
Risk of false negative in second parent	(1 − PPV) CHCT: (1 − 0.82) = 0.18 IVCT: (1 − 0.94) = 0.06	(1 − PPV) × (1 − Sensitivity) CHCT: (1 − 0.82) × (1 − 0.97) = 0.0054 IVCT: (1 − 0.94) × (1 − 0.99) = 0.0006
Efficiency	(Sensitivity + Specificity) ÷ 2 CHCT: (0.97 + 0.78) ÷ 2 = 0.875 IVCT: (0.99 + 0.936) ÷ 2 = 0.963	Sensitivity × Specificity CHCT: 0.97 × 0.78 = 0.7566 IVCT: 0.99 × 0.936 = 0.927
Chance of 2 positive	0	Sensitivity × (1 − Specificity) CHCT: 0.97 × (1 − 0.78) = 0.2134 IVCT: 0.99 × (1 − 0.936) = 0.927

MHS = malignant hyperthermia susceptible; PPV = positive predictive value for pretest probability = 0.5 for a parent. From Figure 1, when PTP is 0.5, PPV = 0.82 for the CHCT (fig. 1a); and PPV = 0.94 for the IVCT (fig. 1b). Sensitivity and specificity for CHCT and IVCT are shown in table 2. The top row refers to the case where one parent has a positive contracture test, and the second parent is yet to be tested. In the "incomplete" testing strategy, the second parent is automatically assumed to be negative. In the "complete" testing strategy, the second parent is offered testing. Probability calculations are derived from figure 4.

identical sibling. However, subjecting families to genetic linkage studies in which both parents test positive holds little prospect for finding correct linkages.

When one parent of an MHS patient has a negative contracture test result and the result for the other parent is unknown, the certainty of the negative diagnosis is equal to NPV. This value is 0.99 for the IVCT and 0.96 for the CHCT. If this result is a true negative one, then the untested parent must be positive, with the same degree of certainty. This value (0.99 IVCT, 0.96 CHCT) is the PTP for the untested parent. In this case, as in the case for high PTP probands, if testing is performed, PPV is unchanged, and NPV decreases to a dangerously low value.

Given our assumptions, testing the second parent provides little extra clinical information and exposes the parent to an unacceptable risk of a false-negative result. A "rational" testing strategy implies that the partner of a negative parent should not be tested, a conclusion that may appear to be counterintuitive. This conclusion may not apply if the assumptions are violated, such as nonpaternity or nonmaternity, spontaneous MH mutation, or alternate genetic modes of inheritance.

Spectrum Bias

The estimates of sensitivity and specificity used in our Bayesian model were made by examining diagnostic test performance in two populations, one with "severe" MH (D6 probands) and the other with an extremely low prevalence of MHS (control patients). Using such extreme populations, "the sickest of the sick" and the "wellest of the well," may overestimate sensitivity and specificity for populations with intermediate prevalence of MHS. This effect is called "spectrum bias."[26] If MHS is truly unequivocal, then spectrum bias may be minimal. If contracture testing for MHS shows a graded spectrum of responses,[48–50] then sensitivity and specificity will be lower in a population with intermediate prevalence than in the extreme populations.

In the absence of an independent "gold-standard" for MHS, there is no simple way to verify true sensitivity and specificity. Any spectrum bias, however, would result in lower PPV and NPV values, but unless true sensitivity and specificity are known, it is not possible to estimate the magnitude of this effect. Even given the highest possible estimates of sensitivity and specificity, limitations in PPV and NPV are inevitable so long as sensitivity and specificity are not both 1.0. By recognizing these limitations, it is possible to make effective use of diagnostic testing information for clinical and research purposes.

Conclusions

The use of Bayesian modeling based on the estimated sensitivity and specificity of diagnostic contracture testing for MH can allow an understanding of the strengths and limitations of muscle biopsy contracture testing. Bayesian modeling of PTP, PPV, NPV, and efficiency can aid in deciding whom to test, when to test, and how best to interpret the test result, both for clinicians and for researchers.

A clinician relies on the excellent NPV, which accounts for the clinical utility of this test. For any person other than a D6 proband, a negative contracture test effectively rules out MH for them and their descendants. To achieve this, there is an acceptably low false-positive rate, measured by high, but imperfect, PPV. In the absence of a better alternative, contracture testing is a useful clinical tool when properly applied with sequential and complete testing strategies.

For genetic analysis, a much higher PPV is required to be certain that a person who had a positive test result is truly MHS. One approach is to alter the test to achieve higher specificity. This is the EMHG "specific" IVCT interpretation. However, the PTP for a relative is fixed by distance in kinship, especially when there are gaps or MHE diagnoses in the pedigree. Even with a highly specific test, an optimal testing strategy, certain proband identification, and perfect sensitivity, a downward drift in PPV is unavoidable.

A more accurate approach is to accept the limitations of contracture testing and to work within them. Previous linkage models have been used to find association with a positive contracture test result, which has a small, but quantifiable, inherent error in predicting MHS. An alternate linkage model that incorporates the estimated errors in phenotyping contained in this article would be less prone to identifying discordance where discordance does not exist.

The current estimate that only 50% of MH families[20] are linked to chromosome 19q13.1 is almost certainly a large underestimate, once allowance is made for phenotypic uncertainty. The frequent finding of about one discordant individual for each six to eight subjected to the CHCT or IVCT[18,21,51,52] would be expected from the modeling presented here. Where such discordance is observed, it is possible that the linkage exists, but it is obscured by phenotypic uncertainty.

The authors thank Drs. M. Larach, G. Allen, and H. Ørding for sharing their data on sensitivity and specificity of the CHCT and IVCT as they developed the model, and for their critical reading of the manuscript. The authors also thank Dr. B. A. Britt for her insight into malignant hyperthermia and Dr. A. Logan for his assistance in clinical epidemiology.

References

1. Denborough M, Lovell R: Anaesthetic deaths in a family. Lancet 1960; 2:45

2. Britt B: Malignant hyperthermia: A review, Thermoregulation: Pathology Pharmacology and Therapy. Edited by E Schönbaum, P Lomax. New York, Pergamon Press, 1991, pp 179-292

3. Kalow W, Britt B, Terreau M: Metabolic error of muscle metabolism after recovery from malignant hyperthermia. Lancet 1970; 2:895-8

4. Ellis F, Keaney N, Harriman D, Sumner D, Kyei-Mensah K, Tyrrell J, Hargreaves J, Parkh R, Mulrooney P: Screening for malignant hyperthermia. BMJ 1972; 3:559-61

5. Larach M, for The North American Malignant Hyperthermia Registry: Standardization of the caffeine halothane muscle contracture test. Anesth Analg 1989; 69:511-5

6. European MH Group: A protocol for the investigation of malignant hyperthermia (MH) susceptibility. Br J Anaesth 1984; 56:1267-9

7. MacLennan D, Duff C, Zorzato F, Fujii J, Phillips M, Korneluk R, Frodis W, Britt B, Worton R: Ryanodine receptor gene is a candidate for disposition to malignant hyperthermia. Nature 1990; 343: 559-61

8. McCarthy T, Healy J, Heffron J, Lehane M, Duefel M, Lehmann-Horn F, Faralli M, Johnson K: Localization of the malignant hyperthermia susceptibility locus to human chromosome 19q12-13.2. Nature 1990; 343:562-4

9. Fujii J, Otsu K, Zorzato F, deLeon S, Khanna V, Weiler J, O'Brien P, MacLennan D: Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. Science 1991; 253:448-51

10. Otsu K, Khanna V, Archibald A, MacLennan D: Co-segregation of porcine malignant hyperthermia and a probable causal mutation in the skeletal muscle ryanodine receptor gene in backcross families. Genomics 1991; 11:744-50

11. Gillard E, Otsu K, Fujii J, Khanna V, deLeon S, Derdemezi J, Britt B, Duff C, Worton R, MacLennan D: A substitution of cysteine for arginine-614 in the ryanodine receptor is potentially causative of human malignant hyperthermia. Genomics 1992; 11:751-5

12 .Gillard E, Otsu K, Fujii J, Duff C, deLeon S, Khanna V, Britt B, Worton R, MacLennan D: Polymorphisms and deduced amino acid substitutions in the coding sequence of the ryanodine receptor (*RYR1*) gene in individuals with malignant hyperthermia. Genomics 1992; 13:1247-54

13. Quane K, Healy J, Keating K, Manning B, Couch F, Palmucci F, Doriguzzi C, Fagerlund T, Berg K, Ørding H, Bendixen D, Mortier W, Linz U, Müller C, McCarthy T: Mutations in the ryanodine receptor gene in central core disease and malignant hyperthermia. Nat Genet 1993; 5:51-5

14. Zhang Y, Chen H, Khanna V, deLeon S, Phillips M, Schappert K, Britt B, Brownell A, MacLennan D: Identification of a mutation in human ryanodine receptor associated with central core disease. Nat Genet 1993; 5:61-5

15. Quane K, Keating K, Healy J, Manning B, Krivosic-Horber R, Krivosic I, Monnier N, Lunardi J, McCarthy T: Mutation screening of the *RYR1* gene in malignant hyperthermia: Detection of a novel Tyr to Ser mutation in a pedigree associated with central cores. Genomics 1994; 23:236

16. Quane K, Keating K, Manning B, Healy J, Monsieurs K, Heffron J, Lehane M, Heytens L, Krivosic-Horber R, Adnet P, Ellis F, Monnier N, Lunardi J, McCarthy T: Detection of a novel common mutation in the ryanodine receptor gene in malignant hyperthermia: Implications for diagnosis and heterogeneity studies. Hum Mol Genet 1994; 3:471

17. Keating KE, Quane KA, Manning BM, Lehane M, Hartung E, Censier K, Urwyler A, Klausnitzer M, Muller CR, Heffron JJ, McCarthy TV: Detection of a novel RYR1 mutation in four malignant hyperthermia pedigrees. Human Molecular Genetics 1994; 3:1855-8

18. Phillips MS, Khanna VK, De Leon S, Frodis W, Britt BA, MacLennan DH: The substitution of Arg for Gly2433 in the human skeletal muscle ryanodine receptor is associated with malignant hyperthermia. Human Molecular Genetics 1994; 3:2181-6

19. Lynch PJ, Krivosic-Horber R, Reyford H, Monnier N, Quane K, Adnet P, Haudecoeur G, Krivosic I, McCarthy T, Lunardi J: Identification of heterozygous and homozygous individuals with the novel RYR1 mutation Cys35Arg in a large kindred. ANESTHESIOLOGY 1997; 86:620-6

20. Ball SP, Johnson KJ: The genetics of malignant hyperthermia. J Med Genet 1993; 30:89-93

21. MacLennan DH: Discordance between phenotype and genotype in malignant hyperthermia. Curr Opin Neurol 1995; 8:397-401

22. Sudbrak R, Procaccio V, Klausnitzer M, Curran JL, Monsieurs K, van Broeckhoven C, Ellis R, Heyetens L, Hartung EJ, Kozak-Ribbens G, Heilinger D, Weissenbach J, Lehman-Horn F, Mueller CR, Deufel T, Stewart AD, Lunardi J: Mapping of a further malignant hyperthermia susceptibility locus to chromosome 3q13.1. Am J Human Genet 1995; 56:684-91

23. Monnier N, Procaccio V, Stieglitz P, Lunardi J: Malignant-hyperthermia susceptibility is associated with a mutation of the alpha 1-subunit of the human dihydropyridine-sensitive L-type voltage-dependent calcium-channel receptor in skeletal muscle. Am J Human Genet 1997; 60:1316-25

24. Robinson RL, Monnier N, Wolz W, Jung M, Reis A, Nuernberg G, Curran JL, Monsieurs K, Stieglitz P, Heytens L, Fricker R, van Broeckhoven C, Deufel T, Hopkins PM, Lunardi J, Mueller CR: A genome wide search for susceptibility loci in three European malignant hyperthermia pedigrees. Human Molecular Genetics 1997; 6:953-61

25. Dawson-Saunders B, Trapp R: Basic and Clinical Biostatistics. Norwalk, Connecticut, Prentice Hall, 1990

26. Sox HJ, Blatt M, Higgins M, Marton K: Medical Decision Making. Newton, MA, Butterworth-Heinemann, 1988

27. Weinstein M, Fineberg H: Clinical Decision Analysis. Toronto, Canada, WB Saunders, 1980

28. Isaacs H, Badenhorst M: False-negative results with muscle caffeine halothane contracture testing for malignant hyperthermia. ANESTHESIOLOGY 1993; 79:5-9

29. Islander G, Ranklev-Twetman E: Evaluation of anaesthesias in malignant hyperthermia negative patients. Acta Anesthesiol Scand 1995; 39:819-21

30. Britt B: The North American protocol for the caffeine halothane contracture test in the diagnosis of malignant hyperthermia,

Malignant Hyperthermia—A Genetic Membrane Disease. Edited by S Ohnishi, T Ohnishi. Boca Raton, FL, CRC Press, 1994, pp 171–98

31. Halsall PJ, Cain PA, Ellis FR: Retrospective analysis of anaesthetics received by patients before susceptibility to malignant hyperpyrexia was recognized. Br J Anaesth 1979; 51:949–54

32. Allen GC, Rosenberg H, Fletcher JE: Safety of general anesthesia in patients previously tested negative for malignant hyperthermia susceptibility. ANESTHESIOLOGY 1990; 72:619–22

33. Larach MG: Should we use muscle biopsy to diagnose malignant hyperthermia susceptibility? ANESTHESIOLOGY 1993; 79:1–4

34. Wedel DJ, Nelson TE: Malignant hyperthermia—diagnostic dilemma: False-negative contracture responses with halothane and caffeine alone. Anesth Analg 1994; 78:787–92

35. Ørding H: Investigation of malignant hyperthermia susceptibility in Denmark. Dan Med Bull 1996; 43:111–25

36. Bendixen D, Skovgaard LT, Ørding H: Analysis of anaesthesia in patients suspected to be susceptible to malignant hyperthermia before diagnostic in vitro contracture test. Acta Anesthesiol Scand 1997; 41:480–4

37. Larach MG, Landis JR, Bunn JS, Diaz M: Prediction of malignant hyperthermia susceptibility in low-risk subjects. An epidemiologic investigation of caffeine halothane contracture responses. The North American Malignant Hyperthermia Registry. ANESTHESIOLOGY 1992; 76: 16–27

38. Allen GC, Larach MG, Kunselman AR: The sensitivity and specificity of the caffeine halothane contracture test: A report form the North American Malignant Hyperthermia Registry. ANESTHESIOLOGY 1997 (In press)

39. Ørding H: Diagnosis of susceptibility to malignant hyperthermia in man. Br J Anaesth 1988; 60:287–302

40. Ørding H, for the European Malignant Hyperthermia Group: *In vitro* contracture test for diagnosis of malignant hyperthermia following the protocol of the European MH group: Results of testing patients surviving fulminant MH and unrelated low-risk subjects. Acta Anesthesiol Scand 1997; 41:955–66

41. Larach MG, Localio AR, Allen GC, Denborough MA, Ellis FR, Gronert GA, Kaplan RF, Muldoon SM, Nelson TE, Ørding H, Rosenberg H, Waud BE, Wedel DJ: A clinical grading scale to predict malignant hyperthermia susceptibility. ANESTHESIOLOGY 1994; 80:771–9

42. Hopkins PM, Ellis FR, Halsall PJ, Stewart AD: An analysis of the predictive probability of the in vitro contracture test for determining susceptibility to malignant hyperthermia. Anesth Analg 1997; 84:648–56

43. von Richthofen V, Wappler F, Scholz J, Fiege M, Köchling A: Prediction of malignant hyperthermia susceptibility with the clinical grading scale. ANESTHESIOLOGY 1997; 87:A997

44. Fricker R, Knolle E, Kress HG: Clinical grading (CG) in the diagnosis of malignant hyperthermia (MH). ANESTHESIOLOGY 1997; 87:A1000

45. Kalow W, Sharer S, Britt B: Pharmacogenetics of caffeine and caffeine-halothane contractures in biopsies of human skeletal muscle. Pharmacogenetics 1991; 1:126–35

46. Peng F, Hall WJ: Bayesian analysis of ROC curves using Markov-chain Monte Carlo methods. Med Decis Making 1996; 16:404–11

47. Islander G, Bendixen D, Ranklev-Twetman E, Ørding H: Results of in vitro contracture testing of both parents of malignant hyperthermia susceptible probands. Acta Anesthesiol Scand 1996; 40:579–84

48. Ellis FR, Halsall PJ, Christian AS: Clinical presentation of suspected malignant hyperthermia during anaesthesia in 402 probands. Anaesthesia 1990; 45:838–41

49. Urwyler A, Censier K, Kaufmann MA, Drewe J: Genetic effects on the variability of the halothane and caffeine muscle contracture tests. ANESTHESIOLOGY 1994; 80:1287–95

50. Nelson TE, Flewellen EH, Gloyna DF: Spectrum of susceptibility to malignant hyperthermia—diagnostic dilemma. Anesth Analg 1983; 62:545–52

51. Deufel T, Sudbrak R, Feist Y, Rubsam B, Du Chesne I, Schafer KL, Roewer N, Grimm T, Lehmann-Horn F, Hartung EJ, Muller CR: Discordance, in a malignant hyperthermia pedigree, between in vitro contracture-test phenotypes and haplotypes for the MHS1 region on chromosome 19q12-13.2, comprising the C1840T transition in the RYR1 gene. Am J Human Genet 1995; 56:1334–42

52. Serfas KD, Bose D, Patel L, Wrogemann K, Phillips MS, MacLennan DH, Greenberg CR: Comparison of the segregation of the RYR1 C1840T mutation with segregation of the caffeine/halothane contracture test results for malignant hyperthermia susceptibility in a large Manitoba Mennonite family. ANESTHESIOLOGY 1996; 84:322–9