Measuring recovery from general anaesthesia using critical flicker frequency: a comparison of two methods

Y. Salib MB ChB FFARCSI, G. Plourde MSc MD,
K. Alloul BSc, A. Provost MD, A. Moore MD

Critical flicker frequency (CFF) is the frequency at which a flickering light appears steady. It is a sensitive measure for assessing recovery from anaesthesia. The CFF is almost always determined with the method of limits by which the flickering frequency is progressively decreased (or increased) until the patient reports a change from fusion to flicker (or flicker to fusion). This method has two disadvantages: it is influenced by the response bias (i.e., the subjective criterion used by the subject to decide that flicker is present or absent) and by the response delay (i.e., the interval between the perceptual change and the response). To avoid these problems, the method of forced choice is recommended. For each trial, the subject observes the light during two short successive periods. The light flickers during only one period, according to chance. The patient must indicate the period during which flickers occur. If uncertain, the patient has to make a guess. The aim of this study was to compare the two methods for assessing recovery from general anaesthesia. Two induction agents were used to obtain different recovery profiles. Twenty patients undergoing uncomplicated surgery lasting less than two hours were tested. They received either thiopentone or midazolam for induction, according to a randomized design. Vecuronium was used to facilitate tracheal intubation and anaesthesia was maintained with fentanyl, isoflurane and nitrous oxide. The CFF was measured before

induction and at 60, 120 and 180 minutes after arrival in the recovery room. The person measuring CFF was unaware of the induction agent used. The CFF determined by forced choice was higher than that by the method of limits (53.3 vs 49.3 Hz across induction agents, periods and patients, $P < 0.01$). Both methods yielded a similar profile of recovery, indicating that maximal CFF depression occurred at 60 min and that it was more pronounced for midazolam than for thiopentone. The forced choice method was easier to use than that of limits. We conclude that the method of forced choice and that of limits yield similar results. However, the method of forced choice is to be preferred because it eliminates response bias and response delay.

La fréquence critique de fusion (CFF) se définit comme la fréquence à laquelle la stimulation lumineuse intermittente apparaît comme continue. Elle est une mesure sensible de la récupération anesthésique. La CFF est la plupart du temps déterminée par la méthode des limites où la fréquence de la stimulation est progressivement diminuée (ou augmentée) jusqu'à ce que le sujet rapporte un changement (lumière fusionnée à intermittente ou vice versa). Cette méthode a deux désavantages : elle est influencée par le biais de la réponse (i.e. le critère subjectif utilisé par le sujet pour décider si le papillotement est présent ou absent) et par le délai de cette réponse (i.e. l'intervalle entre la perception et la réponse). Pour éviter ces difficultés, il est recommandé d'utiliser la méthode de la contrainte du choix. Pour chaque essai, le sujet observe la lumière pendant deux périodes successives. La lumière ne papillote que pendant une période, au hasard. Le patient doit indiquer la période au cours de laquelle on a introduit le papillotement. En cas d'incertitude, le patient doit deviner. L'objectif de cette étude consiste à comparer les deux méthodes pour évaluer la récupération de l'anesthésie générale. Deux agents d'induction sont administrés pour obtenir des profils différents de récupération. Le groupe étudié comprend vingt patients subissant une chirurgie non compliquée d'une durée de moins de deux heures. L'induction est réalisée aléatoirement avec soit du thiopentone soit du midazolam. Du vecuronium est

Key words

RECOVERY: assessment, critical flicker frequency.

From the Departments of Anaesthesia, Royal Victoria Hospital and McGill University, Montreal, Canada.

Address correspondence to: Dr. Gilles Plourde, Department of Anaesthesia, 687 Pine Avenue West, Suite S5.05, Montreal, Quebec H3A 1A1.

Accepted for publication 13th August, 1992.

utilisé pour faciliter l'intubation et l'anesthésie maintenue avec du fentanyl, de l'isoflurane et du protoxyde d'azote. On mesure la CFF avant l'induction et à 60, 120 et 180 min après l'arrivée en salle de réveil. La personne qui fait la mesure ignore la nature de l'agent d'induction. La CFF déterminée par contrainte de choix est plus élevée que celui déterminé par la méthode des limites (53,3 vs 49,3 Hz comparant agent d'induction, moments d'étude et patients $P < 0,01$). Les deux méthodes produisent un profil similaire de récupération, en montrant que la dépression maximale du CFF survient à 60 min, et est plus marquée pour le midazolam que pour le thiopentone. La méthode de contrainte de choix s'utilise plus facilement que la méthode des limites mais les résultats sont identiques. La méthode de contrainte de choix est préférable parce qu'elle élimine les biais et le délai de réponse.

Critical flicker frequency (CFF) is useful to assess the effects of psychotropic drugs.¹⁻⁴ Determination of the CFF is considered to be one of the best methods to assess the postoperative recovery of mental function because of its high sensitivity and reliability.

Critical flicker fusion is the frequency at which a flickering light source appears to be continuous.^{1,4} It measures a perceptual threshold. The most common method of determining the CFF is the method of limits in which the frequency of flicker is slowly decreased (or increased) until the subject reports onset of flicker (or loss of flicker). The CFF obtained with this method depends on two factors: the perceptual threshold, the factor to be measured, and the response bias, a confounding factor.^{4,6} The response bias is the set of internal rules used by the subject to decide whether or not flicker is perceived at any given frequency. The response bias depends on the perceived benefits and costs associated with correct and incorrect decisions. With the method of limits, it is impossible to determine whether changes in CFF arise from alterations in the perceptual threshold or in the response bias. In theory, changes in CFF measured with the method of limits could be caused solely by changes in the response bias. Another disadvantage of the method of limits is the requirement for the patient to report promptly the onset (or loss) of flicker sensation. If the response delay (i.e., the interval between perception and report) is long or variable, erratic results are obtained.¹⁻⁴

A simple way of eliminating the effect of response bias and of response delay is to use the method of forced-choice.⁴ For each trial, the subject observes the light during two short, successive periods. The light flickers during only one of the two periods, according to chance. The patient must indicate the period during which the light flickers. If uncertain, the patient must nevertheless respond with the best guess. The highest frequency at which the

patient correctly identifies the flickering light is the CFF. Because the subject must make a positive response on each trial, there is no response bias. Despite the superiority of the forced-choice method,⁴ the anaesthesia literature on CFF relies solely on the method of limits. With only one exception,⁷ the problem of response bias has been ignored.

We have compared the two methods of measuring CFF during recovery from general anaesthesia induced with either thiopentone or midazolam and maintained with isoflurane and N₂O. We have used induction agents with different recovery profiles to compare the two methods in their ability to identify differences in the speed of recovery.

Methods

The study was approved by the Hospital Ethics Committee, and written consent was obtained for all patients. Twenty healthy (ASA physical status I-II) patients aged 20-65 yr and having body surface surgery not exceeding 120 min were tested. Some patients had refractive errors but none had known ocular pathology. The patients were randomly allocated to either midazolam or thiopentone as induction agent (ten patients per group, open label).

No premedication was given. Fentanyl, 3 $\mu\text{g} \cdot \text{kg}^{-1}$, was administered two minutes before induction. The patient then received midazolam (0.2-0.3 $\text{mg} \cdot \text{kg}^{-1}$) or sodium thiopentone (5-7 $\text{mg} \cdot \text{kg}^{-1}$). The trachea was intubated with the aid of vecuronium 0.1 $\text{mg} \cdot \text{kg}^{-1}$ given after loss of the eyelash reflex. Anaesthesia was maintained with N₂O (60% end-tidal) and isoflurane (0.5-1% end-tidal) in O₂. The lungs were mechanically ventilated to end-tidal PCO₂ of 32-40 mmHg. Increments of fentanyl 0.5-1 $\mu\text{g} \cdot \text{kg}^{-1}$ were added as needed. The end-tidal isoflurane concentration was reduced to 0.3% for at least 15 min before the end of surgery. Residual paralysis was antagonized by neostigmine and glycopyrrolate. The trachea was extubated when the patient was awake.

Critical flicker frequency

Critical flicker frequency (CFF) was measured with a device assembled in our laboratory, using a goggle with a built-in red light-emitting diode (4 mm diameter, peak wavelength: 660 nm, viewing distance: 1 cm). The person measuring the CFF was not aware of the induction agent used. The patients used their dominant eye, while the other eye was occluded with an opaque goggle. The mean luminous intensity was 2.6 millicandellae. The mean intensity was the same for all assessments. Flickering was produced by a sinusoidal wave form with 100% modulation (peak intensity 0.0 and 5.2 millicandellae). CFF was determined before induction (baseline) and 60, 120, 180 min after arrival in the recovery area.

PUPILLARY DIAMETER

Changes in pupillary diameter affect the CFF with changes in retinal illuminance¹ (i.e., the amount of light reaching the retina). To avoid the confounding effect of changes in pupillary diameter, the use of an artificial pupil has been recommended.¹ An artificial pupil is an opaque screen with a two millimetre circular orifice, through which the subjects look at the light source. Since the orifice is smaller than the smallest pupillary diameter that can be achieved, the amount of light reaching the retina remains constant. The use of an artificial pupil in heavily sedated patients who have poor control of their eye movements is not practical. The second-best method, which we have adopted, is to measure the pupillary diameter and use statistical procedures to determine if changes in pupillary diameter would account for CFF changes. The pupillary diameter was therefore measured with a pupil gauge before each determination of CFF.

THE METHOD OF LIMITS

The starting frequency was 60 Hz. Progressively the frequency was decreased at the rate of $1.3 \text{ Hz} \cdot \text{sec}^{-1}$ until the patient reported flicker. Frequency at which flicker was reported was recorded. The procedure was repeated until three consistent values (within 3 Hz) were obtained, and the mean value was used.

THE METHOD OF FORCED CHOICE

For each trial, the subject observed the light during two successive one-second periods. During one period, the light was flickering at a given frequency, while in the other, the light was steady. Whether flicker occurred in the first or second period was determined randomly ($P = 0.5$). The subject must identify the period during which flicker occurred, even if he or she had to guess. The pair of one-second periods (with the same order – flicker, non-flicker) may be presented up to four times within each trial, if the subject so requests. The starting frequency is 30 Hz and was initially increased by 2 Hz steps until the patients requested to see the stimuli more than once for a given trial. The step is afterward reduced to 1 Hz. The last frequency for which the patient correctly identifies the flicker period was recorded. The procedure was repeated twice, and the mean value was used for scoring. The method of forced choice with one Hz steps and two possible choices systematically overestimated the true flicker threshold by one Hz (see Appendix).

The method of limits was always used before that of forced-choice, because performing forced-choice trials immediately before limit trials can influence the response bias. Measuring the CFF with either method requires 2–3 minutes.

Analysis

The characteristics (age, etc.) of the midazolam and thiopentone groups were compared with *t* tests. The CFF results were analyzed with an analysis of variance (ANOVA) with three factors: time, CFF method and drug. The time factor consisted of four repeated levels – baseline and 60, 120, 180 min after arrival in the recovery area. The CFF method factor included two repeated levels: the method of limits and the method of forced choice. The drug factor consisted of two non-repeated levels: thiopentone and midazolam. Geisser-Greenhouse adjustment of the degrees of freedom was used for the time factor and interactions. Tukey's honestly significant difference test was used for post-hoc testing. *T* tests based on the pooled estimate of variance were used to compare midazolam versus thiopentone at each assessment time for each method. Least squares linear regression and the Pearson correlation coefficient (*r*) were used to examine the relationship between the two methods at each assessment time.⁸

The pupillary diameter results were analyzed with an ANOVA with two factors: time and drug. Because there were significant changes in pupillary diameter with time, and because changes in pupillary diameter influence CFF via changes in retinal luminance, an analysis of covariance (ANCOVA) with two factors (time and drug) was performed on the CFF results with pupillary diameter as a time-varying covariate. This was done separately for the limit and forced-choice methods.⁸ Unless otherwise specified, all statistical procedures were performed with the CSS STATISTICA programme (V 3.0 for PC) (Stat-Soft®, Tulsa, OK 74104, USA). A $P < 0.05$ was considered significant except for the pooled *t* tests for which $P < 0.017$ was required (Bonferroni, three tests).

Results

The thiopentone and midazolam groups were comparable with regard to sex ratio (six men in each group), age (35.3 ± 13.8 yr (mean $\pm$ SD) for thiopentone and 35.4 ± 14.0 yr for midazolam), duration of surgery (86 ± 25 min for thiopentone and 90 ± 23 min for midazolam), time between end of surgery and arrival to the recovery room (9.8 ± 3.9 min for thiopentone and 13.6 ± 6.2 min for midazolam), total dose of morphine received in the recovery room (4.1 ± 3.4 mg for thiopentone and 3.0 ± 3.1 mg for midazolam) and total dose of droperidol received for nausea and vomiting in the recovery room (0.22 ± 0.36 mg for thiopentone and 0.24 ± 0.25 mg for midazolam).

Three patients from the midazolam group missed one assessment because of nausea, vomiting or poor cooperation. This occurred at 60, 120 and 180 min. Two patients from the thiopentone group also missed one assessment.

FIGURE 1 CFF values for thiopentone at baseline (time 0), and at times 60, 120 and 180 min after arrival in the recovery room, with the forced choice and the limit methods. Time 0 is pre-induction baseline in this and all other figures. Values at 60 min were lower than all other values for both methods ($P < 0.01$). Empty rectangles: limit. Filled rectangles: forced choice. Bars: one standard deviation.

FIGURE 2 CFF values for midazolam at baseline (time 0), and at times 60, 120 and 180 min after arrival in the recovery room, with the forced choice and the limit methods. Values at 60 min were significantly lower than any other values for both methods ($P < 0.01$). Empty rectangles: limit. Filled rectangles: forced choice. Bars: one standard deviation.

This occurred at 60 and 120 min. The ten missing values (five patients $\times$ one missed assessment $\times$ 2 CFF methods) were replaced by values obtained by maximum likelihood estimation (BMDP 5V programme)⁹ for all statistical tests, except regression analysis.

The results are illustrated in Figures 1 and 2. The ANOVA yielded two significant effects: time ($P < 0.001$) and method ($P < 0.01$). There was no significant effect for drug and no interactions. The CFF at baseline was higher than all other values ($P < 0.01$), while the CFF at 60 min was lower than all other values ($P < 0.01$). The means collapsed across methods and drugs were: baseline minute: 57.4 Hz; 60 min: 44.6 Hz; 120 min: 50.7 Hz and 180 min: 52.4 Hz.

The CFF with forced choice was higher than the limit method. The means collapsed across time and drug were 53.2 versus 49.3 Hz.

The pooled *t* test indicated that the CFF for midazolam was lower than for thiopentone at 60 min with the forced-choice method (44.2 vs 51.1 Hz, $P < 0.01$), and the limit method (38.3 vs 44.6 Hz, $P < 0.01$). No other significant differences between the two drugs were found.

The ANOVA for pupillary diameter yielded an effect ($P < 0.01$) for time. The pupillary diameter at baseline (3.2 ± 0.4 min) for thiopentone and midazolam groups combined was ($P < 0.01$) larger than during all other periods (60 min: 2.7 ± 0.4 min; 120 min: 2.8 ± 0.4 min and 180 min: 2.9 ± 0.4 min). The ANCOVAs revealed no relationships between pupillary size and CFF changes ($P = \text{NS}$).

The relationship between the two methods is shown in Figure 3. There was a correlation between the method of limits and the method of forced choice only at 60 min ($r = 0.6$, $P < 0.005$) and at 120 min ($r = 0.8$, $P < 0.001$). At baseline and 180 min, the *r* was respectively 0.2 and 0.3, ($P = \text{NS}$).

Discussion

We found that the two methods of measuring the CFF yielded similar recovery profiles. For both the thiopentone and midazolam groups, maximal depression of CFF occurred at 60 min postoperatively, and was followed by gradual recovery. The CFF at 180 min was still significantly lower than the baseline value. The CFF for the midazolam group was lower than that of the thiopentone group at 60 min. These findings confirm that recovery following induction with midazolam is slower than with thiopentone^{7,10} and that CFF is reduced following thiopentone⁷ or midazolam.^{11,12} The analysis of covariance showed that the CFF changes could not be solely ascribed to changes in pupillary diameter.

As expected, the forced-choice method yielded higher CFF values than the method of limits (53.2 vs 49.3 Hz, i.e., a difference of 3.9 Hz). Two factors account for the difference: (1) the forced-choice method overestimates the true threshold by one Hz, when two alternatives are used (see Appendix); and (2) even if one corrected the systematic, i.e., one Hz overestimation, the forced-choice method yields a perceptual threshold which is undoubtedly higher than that required for perceiving flicker with certainty.

There was a good correlation between the CFF measured with the forced choice and limit methods at time 60 and 120 min. There was, however, no correlation between the two methods, at baseline and at 180 min, when the patients were more alert. We have no explanation for the absence of correlation. The findings suggest that response bias occurs mostly when the patient is alert.

Should the method of forced choice replace the more popular method of limit? We believe so, for three reasons. The method of forced choice is easy to use, it eliminates the problem of response bias and removes the need for prompt notification of perceptual changes by the patient. This last advantage was most appreciated because, with the method of limits, there is always concern about patients taking too long to respond because of drowsiness, inattention or poor motivation. Without prompt notification of changes in perception, the method of limits yields erratic results, explaining perhaps why CFF is sometimes considered unreliable.¹³

Rather than using only one light source and having the patient indicate the period during which flickers occur, as we have done, it would be simpler, as suggested by Wesnes *et al.*,⁴ to use two (or more) light sources (only one flickering for any trial) and have the patient identify the flickering light. Care should be taken to ensure that the lights are identical in all respects. This should be established beforehand, by showing that the detectability of flicker is the same for both lights.

We conclude that the method of forced choice is superior, and should be preferred over that of the more commonly-used method of limit for measuring CFF in anaesthesia. Another suggestion is to take into account the changes in the pupillary diameter. If these suggestions are followed, we feel that the usefulness of CFF for assessing the effects of drugs on the CNS will be enhanced.

Acknowledgements

Supported by a grant from Hoffmann-LaRoche. Many thanks to Peter April for assembling the CFF machine; to Susan Caney for word processing; to Jocelyn Faubert for sharing his knowledge of CFF; and to Frederick Salevsky for formulating the Appendix.

FIGURE 3 A comparison between CFF values with the limit and forced choice at time baseline (time 0), and at times 60, 120 and 180 min after arrival in the recovery room. A correlation was only found at 60 ($r = 0.6$, $P < 0.005$) and 120 ($r = 0.8$, $P < 0.001$) min. At 60 min, only seven empty squares (midazolam) are seen. This is because there are four patients which are represented by two squares only because of overlap. These squares are crossed by the regression line. Empty squares: midazolam. Squares with an X: thiopentone.

Appendix

Consider a flicker discrimination task that requires the subject to determine the period during which the light is flickering. It is assumed that at any flicker frequency below the critical flicker frequency, CFF, the subject is always able to make the correct choice. Above CFF the choice of flickering light is considered random, with the probability of a correct guess being 1/2. The frequency at which the first incorrect flicker determination occurs is recorded. Because of the assumptions, the estimated CFF (CFF_{est}) is never less than the CFF, but may be greater due to a series of correct guesses at frequencies above CFF. It is assumed that the actual CFF is close to an integer value; this is not necessary to the arguments, but simplifies understanding. A closer estimate of CFF can be made by reducing the frequency step.

At the first frequency step above CFF, $CFF + 1$ in this case, the test subject is guessing. Half of the time the guess will be incorrect and so, for half of all trials, CFF_{est} will equal CFF. For the other half of all trials, CFF_{est} will be greater than CFF. Amongst this subset of trials, half again will have a miss at $CFF + 2$, the next frequency step. The other half have a hit at $CFF + 2$. In the first case, the CFF_{est} is $CFF + 1$, in the second CFF_{est} is greater than $CFF + 1$. At this point in the argument, the CFF_{est} equals CFF in half of all trials and equals $CFF + 1$ in another 1/4 of all trials. CFF_{est} is greater than $CFF + 1$ in the remaining 1/4 of all trials. Now consider this last subset of trials. Using the same reasoning as above, half of these trials yield a CFF_{est} of $CFF + 2$ and half yield CFF_{est} greater than $CFF + 2$.

Continuing the argument, a series is generated for the average estimate CFF:

$$CFF_{est} = 1/2 CFF + 1/4 (CFF + 1) + 1/8 (CFF + 2) + \dots + 1/2^{n+1} (CFF + n) + \dots \quad (\text{Eq. 1})$$

CFF_{est} becomes exact as $n \rightarrow \infty$; for $n \approx 20$, the values of $1/2^{n+1}$ and $n/2^{n+1}$ are small. This expression for CFF_{est} can then be evaluated using the binomial expansion with a very small error. Rearranging Eq. 1 yields:

$$CFF_{est} = CFF (1/2 + 1/4 + 1/8 + \dots 1/2^{n+1} + \dots) + 1/4 (1 + 2(1/2) + 3(1/4) + \dots n/2^{n-1} + \dots) \quad (\text{Eq. 2})$$

The expansion of $(1-x)^{-1}$ with $x = 1/2$ gives:

$$1/2 + 1/4 + 1/8 + \dots + 1/2^n + \dots = 1,$$

while the expansion of $(1-x)^{-2}$ with $x = 1/2$ gives:

$$1 + 2(1/2) + 3(1/4) + \dots + n(1/2^{n-1}) + \dots = 4$$

Substituting these into Eq. 2 above yields:

$$CFF_{est} = CFF + 1$$

CFF_{est} , averaged over a large number of trials, overestimates the actual CFF by one Hz.

References

- 1 Brown JL. Flicker and intermittent stimulation. In: Graham CH. (Ed.). Vision and Visual Perception, New York: Wiley, 1965: 251-320.
- 2 Smith JM. Critical flicker frequency (CFF) and psychotropic drugs in normal human subjects - a review. Psychopharmacol 1976; 47: 175-82.
- 3 Hindmarch I. Psychomotor function and psychoactive drugs. Br J Clin Pharmacol 1980; 10: 189-209.
- 4 Wesnes K, Simpson P, Christmas L. The assessment of human information - processing abilities in psychopharmacology. Hum Psychopharmacol 1987; 1: 79-92.
- 5 Drummond GB. The assessment of postoperative mental function. Br J Anaesth 1975; 47: 130-42.
- 6 Simonson E, Brozek J. Flicker fusion frequency. Background and applications. Physiol Rev 1952; 32: 349-78.
- 7 Vickers MD. The measurement of recovery from anaesthesia. Br J Anaesth 1965; 37: 296-302.
- 8 Kirk RE. Experimental Design. Procedures for the Behavioral Sciences. 2nd ed. California: Belmont, 1982.
- 9 Schluchter MD. Unbalanced repeated measures models with structured covariance matrices. In: Dixon WJ, Brown MB, Engelman L, Jennrich RI (Eds.). BMDP Statistical Software Manual, Berkeley: University of California Press, 1990: 1207-44.
- 10 Reves JG, Vinik R, Hirschfield AM, Holcomb C, Strong S. Midazolam compared with thiopentone as a hypnotic component in balanced anaesthesia: a randomized, double-blind study. Can Anaesth Soc J 1979; 26: 42-9.
- 11 Kestin IG, Harvey PB, Nixon C. Psychomotor recovery after three methods of sedation during spinal anaesthesia. Br J Anaesth 1990; 64: 675-81.
- 12 Wilson E, David A, MacKenzie N, Grant IS. Sedation during spinal anaesthesia: comparison of propofol and midazolam. Br J Anaesth 1990; 64: 48-52.
- 13 Cashman JN, Power SJ. An evaluation of tests of psychomotor function in assessing recovery following a brief anaesthetic. Acta Anaesthesiol Scand 1989; 33: 693-7.