Rapid Communication

$^{13}$C-Urea breath test for Helicobacter pylori: cut-off point determination by cluster analysis

Francois MION, Gaëlle ROSNER*, Marina ROUSSEAU* and Yves MINAIRE

UMR 5578, CNRS, and Fédération des Spécialités Digestives, Hôpital E. Herriot, 69437 Lyon, Cedex 03, France, and *INBIOMED, Passage du VERCORS, BP7312, 69365 Lyon, Cedex 07, France

(Received 1 April 1997; accepted 8 May 1997)

INTRODUCTION

1. This study was performed on a large set of $^{13}$C-urea breath test results to determine the optimal cut-off point of the test for the diagnosis of Helicobacter pylori (Hp) infection.

2. The following steps were applied to three sets of urea breath test results obtained in three groups of subjects (696 adults before anti-Hp treatment, 1056 adults after anti-Hp treatment and 173 children under 17 years of age): (1) demonstrate the distribution of urea breath test results as a mix of two normal populations (Hp negative and Hp positive) by logarithmic transformation of the results in each group of subjects; (2) apply statistical cluster analysis to determine the separation point between Hp-negative and -positive populations; (3) calculate the mean and SD of each population, and use these parameters in the equation of the normal distribution to establish the frequency curves of Hp-negative and -positive populations; and (4) determine the cut-off point of the urea breath test as the intersection of the two curves, and the risks of error related to it.

3. The optimal cut-off point was found at +3.00 $\delta$‰, with a risk of false-negative or -positive response of the urea breath test of less than 3%. From this, a cut-off point of +3.00 $\delta$‰ for the $^{13}$C-urea breath test is recommended, with an indetermined zone between +2.5 and +3.5 $\delta$‰ to account for the spontaneous variation of $^{13}$CO$_2$ in breath and the limits of GC–isotope ratio-MS analytical precision.

METHODS

The results of UBTs performed in 696 adults, to establish the diagnosis of Hp infection (diagnosis group), and in 1056 adults, to control the efficacy of Administration has cleared and approved a UBT application (Mereteck UBT™) for the diagnosis of Hp infection. UBT has been recognized as the best non-invasive test to control the efficacy of anti-Hp treatments [2–4]. This indirect diagnostic test relies on the urease activity of Hp: when present in the gastric cavity, Hp metabolizes $^{13}$C-urea given orally to ammoniac and $^{13}$CO$_2$, which is then readily detectable in breath. The UBT is performed, after an overnight fast, by obtaining two breath samples (through a plastic straw placed at the bottom of a 10 ml glass tube closed immediately at the end of a 20 s exhalation), one before ($T_0$) and the second 30 min ($T_{30}$) after oral administration of 75 mg of $^{13}$C-labelled urea, to measure $^{13}$C enrichment in CO$_2$. A citric acid solution (200 ml) is also given orally before the administration of labelled urea, to slow down gastric emptying and ensure a sufficient contact time between urea and the gastric mucosa [5]. $^{13}$C enrichment ($T_{30}–T_0$) is then calculated, yielding a negative or positive result for Hp infection. There is still some controversy about the best cut-off point; early reports used ($T_{30}–T_0$) = 5 $\delta$‰ $^{13}$C enrichment [5, 6]. However, using receiver operating characteristic curve analysis of UBT results compared with histological analysis of gastric biopsy specimens in 95 patients, we found the cut-off point ($T_{30}–T_0$) = 3 $\delta$‰ to improve the sensitivity of the test without decreasing its specificity [7]. The goal of the present study was to confirm this cut-off point by statistical analysis of a larger set of data.

Key words: adult, children, cut-off value, diagnosis value, Helicobacter pylori, urea breath test.

Abbreviations: Hp, Helicobacter pylori; IRMS, isotope ratio mass spectrometry; UBT, $^{13}$C-urea breath test.

Correspondence: Dr François Mion.
an anti-Hp treatment (post-treatment group), were used for this study. All these patients were included in therapeutic protocols approved by the French National Ethics Committee to evaluate the efficacy of triple therapy regimens (two antibiotics and a proton-pump inhibitor for 7–14 days), and the UBTs were performed as requested in these protocols. To control Hp eradication, UBTs were always done at least 4 weeks after the end of treatment. Similarly, the results of UBTs performed for the diagnosis of Hp infection in 173 children younger than 17 years old were used. For children, the same protocol as in adults was used, except for the dose of $^{13}$C-labelled urea: 50 mg was given orally instead of 75 mg. Furthermore, in children under the age of 4, a mask was used for breath sampling.

In each case, $^{13}$CO$_2$ enrichment in $T_0$ and $T_{30}$ breath samples was measured by GC–isotope ratio–MS (IRMS) at a central facility, and $(T_{30} - T_0)$ values were calculated and expressed in $\delta^{13}$C. In the three groups of patients, visual analysis of the histograms of $(T_{30} - T_0)$ values showed a mix of two positively skewed populations: the logarithmic transformation of $(T_{30} - T_0)$, $\ln(T_{30} - T_0)$, transformed these distributions into symmetrical (normal) ones (Fig. 1). These two populations were presumed to be representative of Hp-negative and Hp-positive patients.

The normal distribution of $\ln(T_{30} - T_0)$ allowed cluster analysis to be performed on these data, in order to determine the minimal intra-class variance [9], and thus the $\ln(T_{30} - T_0)$ value best separating the presumed Hp-negative and Hp-positive populations. Thereafter, the parameters (mean and SD) of these Hp-negative and Hp-positive populations were established. These parameters were then used in the equation of normal distribution to draw the frequency curves for both Hp-negative and Hp-positive populations. The intersection between the two curves corresponds to the $\ln(T_{30} - T_0)$ value at which the probability tilts over: below this value, it is more likely for an individual to belong to the Hp-negative group; above this value, it is most likely for an individual to be Hp positive. The optimal $(T_{30} - T_0)$ cut-off point is then obtained from the $\ln(T_{30} - T_0)$ intersection value.

Finally, in the three groups of UBT results, the $\ln(T_{30} - T_0)$ cut-off value was standardized for each presumed Hp-negative and Hp-positive distribution, according to the formula $Z = (X_i - \mu)/\sigma$ (where $X_i = \text{cut-off point}$, $\mu = \text{mean of the population}$ and $\sigma = \text{SD}$). The $Z$ value (called the normal deviate) indicates how many SDs from the mean the cut-off point is located. The $Z$ values obtained were then reported to the table of the proportions of the normal curve, which gives the proportion of the normal curve that is more extreme than a given normal deviate $Z$ [10]. Thus, the probability for a given individual Hp negative to produce a UBT result corresponding to an $\ln(T_{30} - T_0)$ value greater than the cut-off point (i.e. a false-positive result), as well as the probability for a patient belonging to the Hp-positive population producing a UBT result corresponding to an $\ln(T_{30} - T_0)$ value smaller than the cut-off point (i.e. a false-negative result), was determined.

![Fig. 1. Histogram of the logarithmically transformed UBT values, $\ln(T_{30} - T_0)$, obtained in 1056 adults tested after anti-Hp treatment. Note the coexistence of two normally distributed populations (presumed Hp-negative and Hp-positive patients).](image-url)
RESULTS

In order to increase the symmetry of the distributions, 2.5% of the highest and lowest values of $\ln(T_{30} - T_0)$ were removed. The minimal intra-class variance was calculated by cluster analysis, indicating the $\ln(T_{30} - T_0)$ value of separation between the two normal distributions of logarithmically transformed UBT results: this $\ln(T_{30} - T_0)$ value of separation was 0.95 in the adult diagnosis group, 0.74 in the adult post-treatment group and 0.68 in the paediatric group. From this, the mean ($\mu$) and SD ($\sigma$) of $\ln(T_{30} - T_0)$ for the presumed Hp-negative and Hp-positive populations were calculated:

- For Hp negative: $\mu = -0.936 \pm 1.041$ and $\sigma = 2.728 \pm 0.696$ for Hp positive in the adult diagnosis group;
- For Hp positive: $\mu = -0.986 \pm 0.994$ for Hp negative, and $2.552 \pm 0.746$ for Hp positive, in the post-treatment group; and $\mu = -0.998 \pm 1.003$ for Hp negative and $2.637 \pm 0.767$ for Hp positive in the paediatric population.

These parameters were then used in the equation of the normal distribution $f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $x$ representing the value of the variable $\ln(T_{30} - T_0)$, to draw the frequency curves of Hp-negative and Hp-positive populations (Fig. 2). Obtain the intersection point of these two curves, and thus the $(T_{30} - T_0)$ cut-off point. The intersection point was found to be $\ln(T_{30} - T_0) = 1.18$ in the adult diagnosis group, 1.01 in the adult post-treatment group, and 1.00 in the paediatric group. Thus, the optimal UBT cut-off point was $T_{30} - T_0 = 3.25\%_\alpha$ in the adult diagnosis group, $2.75\%_\alpha$ in the adult post-treatment group and $2.72\%_\alpha$ in the paediatric group.

The formula $[Z = (X_i - \mu)/\sigma]$ gives the distance (called normal deviate) between the cut-off point $(X_i)$ and the mean ($\mu$) of each normal distribution, expressed in SDs ($\sigma$). The normal deviates were found to be greater than two SDs ($>2\sigma$) in all cases, indicating, by reporting these values to the table of the proportions of the normal curve, that the proportions of HP-negative and HP-positive populations respectively producing a UBT result greater or smaller than the cut-off point were always lower than 2.5%. Thus, the risks of false positivity or false negativity of the UBT for the diagnosis of Hp infection are lower than 2.5% when using a $(T_{30} - T_0)$ cut-off point of: (a) 3.25 $\delta%_\alpha$ in the adult population before any anti-Hp treatment; (b) 2.75 $\delta%_\alpha$ in the adult population after treatment; and (c) 2.72 $\delta%_\alpha$ in the paediatric population.

Replacing the cut-off points obtained for the three different groups of patients by the same UBT $(T_{30} - T_0)$ cut-off point of 3.00 $\delta%_\alpha$, as determined in our previous study [7], the corresponding normal deviates, $Z$, for both Hp-negative and -positive distributions were greater in all cases than 1.94 SDs (1.94$\sigma$). Thus, according to the table of the proportions of the normal curve, the risks or error of the UBT with this cut-off point were still lower than 3%, a more than adequate efficacy for a non-invasive test.

DISCUSSION

Using a statistical cluster analysis of a large set of UBT data, we were able to confirm the validity of the cut-off point of $(T_{30} - T_0) = 3.00\%_\alpha$ for the diagnosis of Hp infection, determined by receiver operative characteristic curve analysis compared with histological analysis of gastric biopsy specimens used as the gold standard in a previous study [7]. The theoretical interest of the statistical method used here is to allow the analysis of a large set of data without the need of a gold standard to be used as reference. Interestingly, the UBT cut-off point as determined by statistical cluster analysis was very similar to that previously obtained by comparison with a gold standard.

With regard to the slight variations in the precision of the UBT when modifying its cut-off point around 3 $\delta%_\alpha$, it does not appear to be useful to use different cut-off points depending on the clinical situation (before or after treatment) or the population tested (adults or children): a cut-off point of $(T_{30} - T_0) = 3.00\%_\alpha$ seems applicable in all cases. In fact, rather than set a very strict cut-off point for the UBT, it would make more sense to determine a ‘grey zone’ for this test, in which the risk of error of the test is maximal. For individuals having a UBT result falling in this ‘grey zone’, a second test or a different diagnostic method (such as a direct test on gastric biopsy specimens) would be recommended to assess more accurately their Hp status. Such a ‘grey zone’ would also take into account the spontaneous
fasting individual variation of respiratory $^{13}$CO$_2$ (around $\pm 0.7\%\delta$, [11]), as well as the analytical precision of the GC-IRMS $^{13}$CO$_2$ measurements ($1 SD = 0.2 \%\delta$, [8]). If a 'grey zone' set for $(T_30 - T_0)$ values between 2.5 and 3.5 $\%\delta$ was applied to the population tested, 2.3% of the patients tested in the adult diagnosis group (16 of 696 patients), only 0.7% of the adults tested after treatment (seven of 1056 patients) and 1.7% of the children tested (three of 173) would have fallen in it. The UBT would thus be inconclusive with regard to Hp infection in approximately 1.4% of the patients. Finally, it must be stressed that the 3.00 $\%\delta$ cut-off point may not be valid if the UBT is performed according to a different protocol. For example, some studies have shown that the UBT may be less accurate if performed in the non-fasting state [12] or with a test meal other than citric acid [13]. In these cases, the optimal cut-off point of the test may be different.

In conclusion, this study shows that it is possible to estimate the cut-off point of UBT by cluster analysis of a large set of data, without comparison with a gold standard. It confirms the validity of the $(T_30 - T_0) = 3.00 \%\delta$ cut-off point for the diagnosis of Hp infection. The study also underlines the necessity of a 'grey zone' in which the accuracy of the UBT does not allow a precise determination of the Hp status for a given individual.

REFERENCES