

Remarks on probability theory and TMJ diagnosis

L.V. CHRISTENSEN and M.M. ASH* *Marquette University, School of Dentistry, Milwaukee, Wisconsin and *School of Dentistry, University of Michigan, Ann Arbor, Michigan, U.S.A.*

Summary

On the basis of the classic concepts of events and probability theory, this article analyzes some recently introduced diagnostic probability concepts as they pertain to temporomandibular joint (TMJ) diseases and disorders.

Probability theory

Events

Two different *simple* events, for example, $E_1 = \text{'Head' (H)}$ and $E_2 = \text{'Tail' (T)}$, cannot occur at the same time — they are mutually exclusive. By contrast, *compound* events can occur simultaneously; for example, $E_1 \geq 2$ and $E_2 \leq 4$ are not exclusive when the compound event is 3 or 3.5. Similarly, a patient may complain of both temporomandibular joint (TMJ) pain and TMJ crepitus; the two complaints are not mutually exclusive. On the other hand, the two different simple events of either presence of TMJ pain or absence of TMJ pain are mutually exclusive.

The intersection of two events E_1 and E_2 implies another event when E_1 and E_2 occur simultaneously:

$$E_1 \cap E_2$$

The special case of $E_1 \cap E_2 = \emptyset$ is an *empty set* where E_1 and E_2 cannot occur simultaneously: an impossible event where E_1 and E_2 are mutually exclusive.

When either E_1 or E_2 occurs — or when both E_1 and E_2 occur — we speak of the *union* of E_1 and E_2 :

$$E_1 \cup E_2$$

In general, the union of E_1 and E_2 means that either one or the other event occurs, but both events can also occur. For example, TMJ crepitus may exist in union with TMJ clicking, but either crepitus or clicking can also occur in cases of TMJ disease. The special case of the union of two events being equal to the outcome space (S) is:

$$E_1 \cup E_2 = S$$

In this case, either E_1 or E_2 must always occur. The *outcome space* S implies that a certain event occurs, and the two events are *exhaustive* because they exhaust all possible outcomes. For example, as an established clinical differential diagnosis we may state that TMJ clicking is associated with dysfunction and TMJ crepitus with

pathology (arthritis, arthrosis) of the mandibular locomotor system, so an outcome space S implies that either crepitus or clicking occurs and that there are no other possible outcomes. Note that we have not stated that the example is the 'clinical truth', only that mathematical logics demand certain formal propositions and relationships for their proper analysis — a fact often overlooked by clinical researchers. Other examples, analyzing purported and putative 'clinical truths', are the subject matter of this discussion.

Another special case is the *complementary* event: which event \bar{E} (complementary event) occurs whenever E *does not* occur:

$$E \cup \bar{E} = S \text{ and } E \cap \bar{E} = \emptyset$$

The event E and the complementary event \bar{E} are exhaustive and mutually exclusive events. For example, we may state that TMJ pathology occurs in the absence of TMJ crepitus (in this case there is an absence of mandibular movement, so-called closed lock).

Concept of probability

Let k be the *absolute* frequency of an event and let n be the total number of observations of the event, then by definition:

$$0 \leq k \leq n$$

Probability theory introduces the concept of the *relative* frequency (h) of an event:

$$h = k/n$$

The relative frequency of an event is the observed frequency of the event divided by the total number of observations (replications). The relative frequency (h) is also known as the rate of occurrence of the event. Whereas k ranges from 0 to n , the relative frequency h ranges from 0 to 1:

$$0 \leq h \leq 1$$

As n tends toward infinity, the relative frequency (h) approaches a certain limit: the *stability of h* . The stability of h is ill defined, but applied mathematics theorize that there is a fixed number P that is approached by h when a long run of observations are made of a particular event. The *hypothetical fixed number P is the probability of the event* under consideration.

In statistical mathematics, the relative frequency (h) of an event is used to estimate the probability (P) of the event, and there are mathematical rules for establishing the reliability of the estimate and for assigning probabilities to events — a fact often overlooked by clinical researchers (see Stockstill & Mohl, 1991). The Swiss mathematician E. Batschelet (1971) said, 'There are scientists who go far beyond the frequency interpretation of probability. They apply probability for the degree of belief'.

Theoretical mathematics make use of *theorems*: propositions that are provable on the basis of explicit assumptions. For example, when we make the explicit assumption that an unbiased (not doctored) coin has a 'Head' (H) and a 'Tail' (T) only, then we may *predict* that by tossing the coin we have a 50% *a priori* probability of obtaining either H or T:

$$P(H) = 0.5 \text{ and } P(T) = 0.5$$

Similarly, when we apply the two examination modalities of TMJ tomography and TMJ electrovibratography, we may explicitly assume — probably erroneously, that either modality yields a probability of 50% ($P = 0.50$) in arriving at a correct ('true') TMJ diagnosis. In this context, it is a fallacy to assume, either explicitly or implicitly, that one can conclusively establish a so-called *Gold Standard* (e.g. $P = 70\%$, $P = 95\%$, $P = 100\%$) for a clinical diagnosis. For example, TMJ imaging techniques, TMJ clinical examination techniques, etc. are by certain clinical researchers effectively assigned a probability of 100% in arriving at a correct ('true') TMJ diagnosis. In effect, these clinical researchers say, 'Trust me! Have I ever lied to you?' However there is cause for concern, as their probability models appear to be too optimistic and too oversimplified (Batschelet, 1971; Lund & Widmer, 1989; Greene, 1990; Goulet & Clark, 1990; Widmer, Lund & Feine, 1990; Mohl *et al.*, 1990a, b, c; Devore, 1991; Christensen, 1992a, b).

Axioms of probability theory

Applied mathematics make use of *axioms*: an undemonstrated ('self-evident or universally accepted principle') proposition — or postulate — concerning an undefined set of elements or properties.

Let an outcome space be S and let E_i ($i = 1, 2, 3, \dots$) be an event attached to S . Each event is a subset of S :

$$E_i \subset S \quad (i = 1, 2, 3, \dots)$$

The probability P_i associated with each event attached to S is:

$$P(E_i) = P_i \quad (i = 1, 2, 3, \dots)$$

Because each probability is an idealized relative frequency, it is postulated that:

$$0 \leq P_i \leq 1 \quad (i = 1, 2, 3, \dots)$$

In other words, the *first axiom* is: with each event attached to an outcome space there is an associated number known as the probability of the event. This number is restricted to the interval of zero (0) to one (1). Impossible events have $h = P = 0$, certain events have $h = P = 1$.

The impossible event has the empty set \emptyset , and the certain event has the outcome space S . Consequently, the *second axiom* is: the impossible event has attached the probability zero (0), and the certain event the probability one (1):

$$P(\emptyset) = 0 \text{ and } P(S) = 1$$

Let E_1 and E_2 be two mutually exclusive events attached to the outcome space S , and imagine that through n observations we determine the frequency of occurrence of E_1 and E_2 . When k_1 is the frequency of occurrence of E_1 and k_2 that of E_2 , then the relative frequencies are:

$$h_1 = k_1/n \text{ and } h_2 = k_2/n$$

and we can ask: how many times did either E_1 or E_2 occur? Because E_1 and E_2 are mutually exclusive, the total frequency is $k_1 + k_2$; the relative frequency of the new event ($E_1 \cup E_2$) is $h = k_1 + k_2/n$, so $h = h_1 + h_2$. Again, because we identify idealized relative frequencies with probabilities, we arrive at the *third axiom*: when the two mutually exclusive events E_1 and E_2 are attached to the outcome space S , we let $P_1 = P(E_1)$, $P_2 = P(E_2)$, and $P = P(E_1 \cup E_2)$, so that:

$$P = P_1 + P_2$$

The third axiom is known as the *addition rule*.

For example, let each of the mutually exclusive diagnostic modalities of TMJ tomography (TOM) and TMJ electrovibratography (EVG) have the probability of 50% in arriving at a correct ('true') TMJ diagnosis. The compound event (TOM, EVG) is the union of the two mutually exclusive events (TOM) and (EVG), so:

$$P(\text{TOM, EVG}) = 0.5 + 0.5 = 1$$

Let the third axiom be applied to event E and the complementary event \bar{E} :

$$P(E \cup \bar{E}) = P(E) + P(\bar{E})$$

Since $E \cup \bar{E} = S$ and $P(S) = 1$ (see second axiom), we obtain:

$$P(E) + P(\bar{E}) = 1$$

Traditionally, $P(E)$ is denoted by p and $P(\bar{E})$ by q (cf. binomial distribution), so:

$$p + q = 1$$

Finally, let E_1, E_2, \dots, E_n be n mutually exclusive and exhaustive events attached to the outcome space S with the probabilities P_1, P_2, \dots, P_n , then:

$$P_1 + P_2 + \dots + P_n = 1$$

In other words, we have partitioned the outcome space into several events — and properly so. Now, imagine that we have measured the so-called *validity* (a spurious measurement of 'truth') of a TMJ diagnosis arrived at through a TMJ clinical examination (EXA = arbitrary validity value of $P = 0.30$), a TMJ tomogram (TOM = arbitrary validity value of $P = 0.30$), and a TMJ electrovibratogram (EVG = arbitrary validity value of 0.40). (Note that from a logical standpoint, clinical validity measurements are practically always 'tail-biting arguments'). If the outcome space (EXA, TOM, EVG) is subdivided into three groups (a), we obtain: $G_1 = 0.30$, $G_2 = 0.30$, $G_3 = 0.40$. Because the three events (groups) are mutually exclusive and — per erroneous definition — exhaustive, we obtain the not uncommon example of:

$$P_1(G_1) + \dots + P_3(G_3) = 1$$

In effect, we have through spurious means arrived at a Gold Standard.

Conditional probability

When an observation (x_i = individual observation) is randomly determined, it implies that other individual observations have the same chance of being determined. For example, we have randomly determined that the probability of agreement between a clinical TMJ exam and the 'true' TMJ condition is $P(A) = 30/103 = 0.29$, that the probability of agreement between a TMJ tomographic exam and the 'true' TMJ condition is $P(B) = 35/103 = 0.34$, and that the probability of agreement between a TMJ electrovibratographic exam and the 'true' TMJ condition is $P(C) = 38/103 = 0.37$. The question can then be asked: what are the probabilities of $P(B)$ and $P(C)$ after exclusion of all clinical findings so that $P(A) = 0$. Although the clinical exam has taken on the probability of zero, it may still be considered as being attached to the outcome space. We consider the relative occurrence as a ratio: $0.34/0.37$, where we must find two probabilities which after addition yield the integer one. The adjustment is made by dividing the one ratio by the sum of the two ratios:

$$P(B) = 0.34/0.34 + 0.37 = 0.48,$$

$$P(C) = 0.37/0.34 + 0.37 = 0.52.$$

The adjusted values are the *conditional probabilities*. Traditionally, the conditional probability is denoted by a vertical bar ('given that'); for example:

$$P(B|A = 0) = 0.48$$

Consider also the following example concerning the imaginary relative incidence of absence and presence of TMJ pathology among 206 joints, and its rate of detection and rejection by TMJ tomography (TMJ TOM) and TMJ electrovibratography (TMJ EVG).

TMJ pathology	TMJ TOM	TMJ EVG	Total
Present ($n = 73$)	$35/206 = 17\%$	$38/206 = 18\%$	35%
Absent ($n = 133$)	$68/206 = 33\%$	$65/206 = 32\%$	65%
Total	$103/206 = 50\%$	$103/206 = 50\%$	100%

Regardless of detection method, the probability of the compound event of presence of TMJ pathology is 35%. Regardless of the absence and presence of pathology, the probability of the compound event of TMJ tomography is the rather irrelevant value of 50%.

On the other hand, a relevant question is: what is the incidence rate of detection of TMJ pathology through TMJ TOM — the answer is: $17/50 = 0.34 = 34\%$. Similarly, what is the incidence rate of detection of TMJ pathology through TMJ EVG — the answer is: $18/50 = 0.36 = 36\%$. In addition, a chi-square value can be calculated to determine whether column frequencies are independent of row frequencies; the null hypothesis might be: TMJ EVG is independent of the presence/absence of TMJ pathology. Finally, it is possible to test statistically for differences among proportions; the null hypothesis might be: the same proportion of cases with and without TMJ pathology have been subjected to TMJ TOM (Devore, 1991).

The literature is replete with more or less meaningful — occasionally absurd —

calculations of so-called positive and negative predictive values, false positive and negative values, and diagnostic sensitivity and specificity values. All of these values are ratios expressing probability. For example, for TMJ TOM the ratio known as the sensitivity value (accuracy) would be:

TMJ TOM	TMJ pathology
Positive	35/73 = 0.479
Negative	63/133 = 0.511
Total	0.990
Sensitivity = 0.479/0.990 = 48%	

While the ratio is mathematically correct, it appears incorrect to subjectively evaluate the ratio, or pass judgement on sensitivity, on the basis of an elusive Gold Standard, as is commonly done (Lund & Widmer, 1989; Greene, 1990; Goulet & Clark, 1990; Widmer, Lund & Feine, 1990; Mohl *et al.*, 1990a, b, c; Stockstill & Mohl, 1991). It appears that events have not been fully considered: the decision making matrix has become a framework for beliefs — often far removed from ‘clinical truth’ (Batschelet, 1971; Devore, 1991; Stockstill & Mohl, 1991; Christensen, 1992a, b).

Multiplication rule

Let A and B be two events attached to the same outcome space according to:

$$P(A \cap B) = P(A) \times P(B|A)$$

The formula expresses the *multiplication rule* for probabilities: the probability of the simultaneous occurrence of the two events A and B is the product of the probability of event A and the conditional probability of event B (given A).

In a given sample of TMJ tomograms and TMJ electrovibratograms, imagine that we have determined there is agreement between the two examination modalities in 94% of cases (A). Imagine further that — in another similar investigation, there is agreement in 90% of cases (B). What is the probability of simultaneous occurrence of A and B given that A has been determined:

Project	Probability of agreement
1	0.94
2	0.90
$P(B A) = 0.94 \times 0.90 = 85\%$.	

In other words, the application of multiple conditions (research projects) decreases the chances of success.

Imagine that a given clinical researcher has not been correct in his few published (even peer reviewed) papers and that we grant him five successive opportunities to be correct in his next ten papers. The question is now: what is the probability that he is correct all five times — the answer is:

$$\begin{aligned} & (5/10)(4/9)(3/8)(2/7)(1/6) = \\ & (0.50)(0.44)(0.38)(0.29)(0.17) = \\ & 0.0041 = 0.4\% \end{aligned}$$

Whether the researcher's opinions are published in, say, papers 3 through 7 or papers 6 through 10, he should, probably, consider his options (e.g. an in depth study of signal detection theory).

References

- BATSCHLET, E. (1971) *Introduction to Mathematics for Life Scientists*, 342–402. Springer Verlag, Berlin.
- CHRISTENSEN, L.V. (1992a) Physics and the sounds produced by the temporomandibular joints. Part I. *Journal of Oral Rehabilitation*, **19**, in press.
- CHRISTENSEN, L.V. (1992b) Physics and the sounds produced by the temporomandibular joints. Part II. *Journal of Oral Rehabilitation*, **19**, in press.
- DEVORE, J.L. (1991) *Probability and Statistics for Engineering and the Sciences*, 1–79, 558–593. Brooks/Cole Publishing Company, Pacific Grove, California.
- GOULET, J.P. & CLARK, G.T. (1990) Clinical TMJ examination methods. *California Dental Association Journal*, **18**, 25.
- GREENE, C.S. (1990) Can technology enhance TMJ disorder diagnosis? *California Dental Association Journal*, **18**, 21.
- LUND, J.P. & WIDMER, C.G. (1989) An evaluation of the use of surface electromyography in the diagnosis, documentation, and treatment of dental patients. *Journal of Craniomandibular Disorders Facial & Oral Pain*, **3**, 125.
- MOHL, N.D., MCCALL, W.D., LUND, J.P. & PLESH, O. (1990a) Devices for the diagnosis and treatment of temporomandibular disorders. Part I: Introduction, scientific evidence, and jaw tracking. *Journal of Prosthetic Dentistry*, **63**, 198.
- MOHL, N.D., LUND, J.P., WIDMER, C.G. & MCCALL, W.D. (1990b) Devices for the diagnosis and treatment of temporomandibular disorders. Part II: Electromyography and sonography. *Journal of Prosthetic Dentistry*, **63**, 332.
- MOHL, N.D., OHRBACH, R.K., CROW, H.C. & GROSS, A.J. (1990c) Devices for the diagnosis and treatment of temporomandibular disorders. Part III: Thermography, ultrasound, electrical stimulation, and electromyographic biofeedback. *Journal of Prosthetic Dentistry*, **63**, 472.
- WIDMER, C.G., LUND, J.P. & FEINE, J.S. (1990) Evaluation of diagnostic tests for TMD. *California Dental Association Journal*, **18**, 53.
- STOCKSTILL, J.W. & MOHL, N.D. (1991) Evaluation of temporomandibular joint sounds. Diagnostic analysis and clinical implications. *Dental Clinics of North America*, **35**, 75.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.