JOURNAL OF FORENSIC SCIENCES

*J Forensic Sci*, 2017
doi: 10.1111/1556-4029.13586
Available online at: onlinelibrary.wiley.com

**PAPER**

**ODONTOLOGY**

*Fernando Rivera-Mendoza,*[1,2] *D.D.S., M.Sc.(c); Stella Martín-de-las-Heras,*[3] *B.D.S., M.D., Ph.D.; Pablo Navarro-Cáceres,*[1,4] *M.Sc.; and Gabriel M. Fonseca,*[1,2] *D.D.S., Ph.D.*

# Bite Mark Analysis in Foodstuffs and Inanimate Objects and the Underlying Proofs for Validity and Judicial Acceptance

**ABSTRACT:** Even though one of the first bite mark cases was *Doyle v. State* in 1954 (a bitten cheese case), the research has focused on bite marks inflicted in human skin. As published Papers, Case Reports, or Technical Notes can constitute precedents which are relied upon in making the legal arguments and a considerable amount of case law exists in this area, we present a systematic review on bite mark analysis in foodstuffs and inanimate objects and their underlying proofs for validity and judicial acceptance according to *Daubert* rulings. Results showed that there is vulnerability in these procedures, and it is essential to demand for focus scrutiny on the known error rates when such evidence is presented in trials. These kinds of bite marks are well documented; however, there has been little research in this field knowing that the protocols of analysis and comparison are the responsibility of the forensic odontologists.

**KEYWORDS:** forensic science, forensic odontology, bite marks, identification, foodstuffs, *Daubert* rulings

Forensic odontology, "the application of the science of dentistry to the field of law," includes several different areas of knowledge; of these, between them, bite mark comparison has produced the most renowned successes and the most controversial cases of the discipline (1). Knowing that skin responds to stress in a nonlinear fashion, contributing to a distortion range, the most prominent research has been focused on bite marks inflicted in human skin, the inherent possibility of distortion, and how this can affect the identification of a suspect (1).

However, one of the first reported bite mark cases in modern legal history was *Doyle v. State* in 1954, in which the bite mark was not in skin but in a piece of cheese at the crime scene of a burglary (2). Bite marks in foodstuffs and inanimate objects are documented in the literature and in court records in different countries, using different exploration and analysis techniques (direct, photographic, microscopic, digital, etc.) (3–7). It has been stated that the most likely scenario for forensic dental analysis exists with bite marks in foodstuffs or other inanimate objects (8). Even though most contemporary cases cite *Doyle v. State* as the basis for rejecting arguments of unproven reliability and acceptability (9)—in fact, Dorion mentioned that its acceptance may have been premature and conducive to a misuse by forensic dentists and others (10)—the same Dorion stated that *Doyle* "...is significant for establishing admissibility in a court of law of bite marks in food" (10).

Although the methods of collecting bite mark evidence are relatively uncontroversial, and the majority of forensic odontologists are satisfied that bite marks can demonstrate sufficient detail for positive identification, bite mark testimony has been criticized on different grounds. Some of the key areas of controversy include the techniques and standards for comparison (1). Webster suggested a classification of bite marks in foodstuffs and inanimate objects to indicate which types of tooth features are likely to be recorded (11). Although this proposal is barely mentioned in the literature as a "commendable but too general classification" (12), there seems to be no other attempt to standardize terminology or procedures specifically for this type of evidence. The American Board of Forensic Odontology (ABFO) suggested that "...dental casts to life-sized photographs, casts of the bite patterns, reproductions of the pattern when in inanimate objects..." as the only mention of a comparison technique for these particular marks (13). The latest American Society of Forensic Odontology's (ASFO) Manual, in the section titled "Bite mark Pattern Recognition and Collection from Humans and Inanimate Objects: Non-invasive Analysis," provides an excellent review of different techniques for photographic documentation, three-dimensional evidence, and bite mark impressions (14). However, despite its comprehensive title, all procedures are focused on the analysis of bite marks in skin, which is then completed in invasive analysis. Dorion mentioned that there has been little research in the field, possibly because bite marks in perishables has the disadvantage of rapid and potentially extreme distortion of both the bite mark and the substrate, and other means of identification can be obtained from the material (12). Silver & Souviron highlighted that while the

[1]Faculty of Dentistry, Magister Program in Dentistry, Universidad de La Frontera, Temuco, Chile.
[2]Faculty of Dentistry, Laboratory of Forensic Dentistry, Universidad de La Frontera, Temuco, Chile.
[3]Department of Forensic Medicine and Forensic Odontology, University of Granada, Granada, Spain.
[4]Faculty of Dentistry, CICO Research Centre in Dental Sciences, Universidad de La Frontera, Temuco, Chile.

Received 10 Dec. 2016; and in revised form 29 April 2017, 25 May 2017; accepted 26 May 2017.

discovery of the material depends solely upon the expertise of crime scene investigators, the protocols for preservation, analysis, and comparison are the responsibility of odontologists (7).

Prior to 1993 in the United States, as the test for admission of novel scientific evidence, federal and most state courts adhered to "general acceptance" in the relevant scientific field. This changed with *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (15), in which the Supreme Court identified three main areas of inquiry when determining admissibility: the expert's qualifications; the dependability and validity of the theories, techniques, and methods used (referred to by the Court as "reliability"); and the relevance of the expected testimony to the issues in the case. Consequently, the admissibility of expert testimony is decided by the court, pursuant to Federal Rules of Evidence 104(a) and 702. Under the 104(a) rule, the proponent of the evidence has the burden of establishing that the pertinent admissibility requirements are met by a preponderance of the evidence; the Federal Rule of Evidence 702 sets out the factors for the court to consider when deciding admissibility questions. *Daubert* set forth a nonexclusive checklist for trial courts to use in assessing the reliability of scientific expert testimony. The specific factors explicated by the *Daubert* court are (i) whether the expert's technique or theory can be or has been tested: that is, whether the expert's theory can be challenged in some objective sense, or whether it is instead simply a subjective, conclusory approach that cannot reasonably be assessed for reliability; (ii) whether the technique or theory has been subject to peer review and publication; (iii) the known or potential error rate of the technique or theory when applied; (iv) the existence and maintenance of standards and controls; and (v) whether the technique or theory has been generally accepted in the scientific community. The *Daubert* ruling ensures that techniques, methodologies, and practices are not only commonly accepted, but that error rates, assessment of reliability, and validation studies are published to support their use. The *Daubert* philosophy has had an unavoidably international influence, even in European (16,17) and Latin American Courts (18).

Although court decisions have precedential value by principle of *Stare decisis*—"standing by that which is decided" (19), the underlying sources do not necessarily constitute precedents for other courts—such decisions can be based on certain published research to support them. As published studies are recommended for acceptance by courts because publications "increases the likelihood that substantive flaws in methodology will be detected" (15), we present a systematic review of published Papers, Case Reports, and Technical Notes focused on bite mark analysis in foodstuffs and inanimate objects, taking into account that a considerable amount of case law exists in the area of bite mark evidence (19). Their underlying proofs for validity and judicial acceptance in accordance with the *Daubert* factors were analyzed and discussed considering the scrutiny of this type of evidence and the uneven globalization of those cases and studies.

## Materials and Methods

We present a systematic systematic review of methodologies for forensic bite mark analysis in foodstuffs and inanimate objects. The search strategy and inclusion of the studied articles was based on the PRISMA® (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement published in 2009 (20). An electronic search was made of the EMBASE, MEDLINE, SCOPUS, SciELO, and LILACS databases up to March 2016 using the search terms "bite marks" OR "bitemarks" OR "bites" AND "foodstuff" OR "food" OR "object" AND "identification." Abstracts were reviewed for relevance to the defined review question. Full texts of case reports, case series, technical notes, and experimental studies on humans available in English, Spanish, and Portuguese were included. Newsgroup articles, reviews, experimental studies on animals, and letters to the editor were excluded. Results of interest examined included the description of analysis techniques of human bite marks in foodstuffs and inanimate objects for forensic purposes.

Two researchers independently reviewed each title and abstract for potential relevance to the research question; articles included by either researcher underwent full-text screening. At the full-text screening stage, two researchers independently reviewed the full text of each article for inclusion. Disagreements were resolved through review and discussion among researchers. Further search methods included hand searching of selected journals to identify studies that may not have been located through electronic database searching. The references cited in all full-text articles and author searches conducted on the names of primary investigators of eligible studies were also searched to identify studies that had been conducted but not reported in an indexed source.

Selected studies were classified according to three of the categories established by the Journal of Forensic Sciences' Author Guidelines for published original material: (a) *Paper*, a full-length research report; (b) *Technical Note*, a description of a technical aspect of a field or issue, report on a procedure or method, or work on validation of techniques or methodologies (usually shorter than *Papers*); and (c) *Case Report*, usually a brief description or analysis of an unusual case or a small series of cases (21). These categories were evaluated according to their publication chronology and according to the affiliated countries.

The full texts were evaluated and agreed by three trained observers (S.M.D.H., F.R.M., G.M.F.) according to their compliance with each of the five *Daubert* factors: (a) *whether the theory or technique is falsifiable, refutable, and/or testable;* (b) *whether it has been subjected to peer review and publication;* (c) *the known or potential error rate;* (d) *the existence and maintenance of standards and controls concerning its operation;* and (e) *whether the technique or theory has been generally accepted in the scientific community.* These evaluations were quantified as 0 (unmentioned or insufficient), 1 (the reference is implicit or barely sufficient), or 2 (the reference is explicit or very sufficient), and the resulting rate was attributed to a specific category according to the relationship between a real/ideal rate as *strong* (80%–100%), *nonthreatened* (60%–80%), *threatened* (40%–60%), *vulnerable* (20%–40%), and *very vulnerable* (0%–20%). The outcome variables were presented with 95% confidence intervals (CIs) to estimate the true value of the population proportion responses.

Interrater agreement was measured for all pair combinations of observers/raters (Obs): Obs1, Obs2, and Obs3 (Cohen's $k$) (22); as well as for multiple raters (Fleiss' $k$) (23), with results categorized as the rating scale proposed by Landis and Koch (24): *poor agreement* (below 0.0); *slight agreement* (0.00–0.20); *fair agreement* (0.21–0.40); *moderate agreement* (0.41–0.60); *substantial agreement* (0.61–0.80); and *almost perfect agreement* (0.81–1.00).

## Results

The database search resulted in 57 articles, and the manual search added 16 articles. After removing duplicate records, as well as full texts not available or which did not meet the

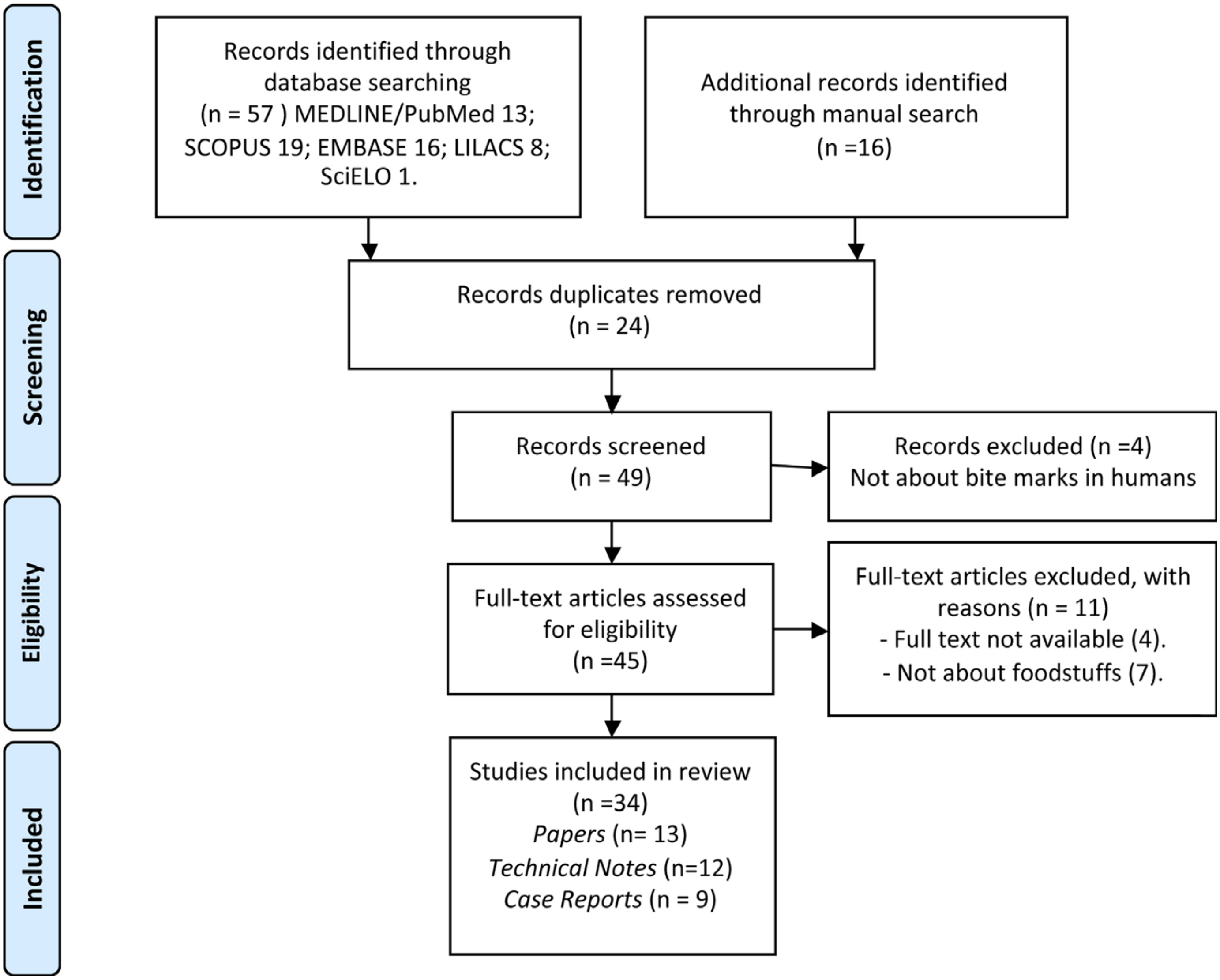


FIG. 1—*PRISMA® flowchart describing the search strategy and inclusion of the studied articles.*

inclusion criteria, the selection resulted finally in 34 articles included in this review (Fig. 1). According to the established categories, 13 *Papers* (38%), 12 *Technical Notes* (35%), and 9 *Case Reports* (27%) were recorded, all published between 1960 and 2015. While *Technical Notes* were published on an almost regular basis throughout the study period, *Case Reports* were published only until 2006. Except for a single publication in 1975 (25), *Papers* began to appear in 1990 and their publication frequency increased in the later years of the study period. Figure 2 shows the chronological relationship between categories of publications and continents belonging to affiliated countries. European countries (the U.K., Norway, Denmark) had a strong initial presence in *Case Reports* and *Technical Notes*, while the last 10 years of the study period showed a significant Asian (India) presence in *Technical Notes* and *Papers*. North America had a smaller presence between 1974 and 1990 (and an isolated *Technical Note* in 2010). *Papers* and *Case Reports* were detected from Africa (South Africa) and Oceania (Australia with Asian collaborations), mostly between 1995 and 2002. South America (Brazil) showed a scant presence of *Papers* and a *Technical Note* at the end of the period. It is important to highlight, in the second half of the study period, the remarkable partnerships between European countries (Portugal/Norway, 2009; Greece/the U.K., 2011) and between different continents (Malaysia/Australia, 1995 and 2001; South Africa/Norway, 2000; Brazil/Portugal, 2013; and Saudi Arabia/India/Australia, 2014), mostly devoted to publishing *Papers* (57%).

### *Daubert Rulings*

Ruling *(b) whether it has been subjected to peer review and publication* was evaluated as *strong* with a score of 100% (68/68) as all articles were recorded as published in journals with an Editorial Board. Ruling *(a) whether the theory or technique is falsifiable, refutable, and/or testable* and ruling *(e) whether the technique or theory has been generally accepted in the scientific community* had similar scores, 64.71% (44/68; 95% CI: 53.347–76.065) and 61.76% (42/68; 95% CI: 50.214–73.316), respectively, and were evaluated as *nonthreatened.* Ruling *(d) the existence and maintenance of standards and controls concerning its operation* was rated 29.41% (20/68; 95% CI: 18.582–40.240) and evaluated as *vulnerable.* Ruling *(c) the known or potential error rate* had the weakest score of 11.77% (8/68; 95% CI: 4.107–19.422) and was evaluated as *very vulnerable* (Table 1).

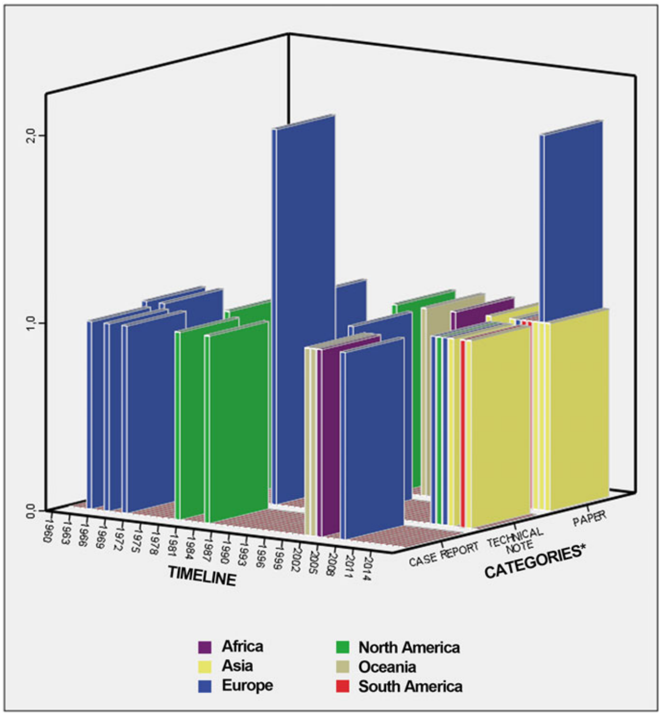


FIG. 2—*Continental geographical distribution of Case Reports, Technical Notes, and Papers on the timeline. *Categories were established according to the Journal of Forensic Sciences' Author Guidelines for the published original material (21). The X axis represents the year of publications of the articles, while the Y axis represents the number of articles accepted in the study.*

Cohen's $k$ values for each pair of observers (Obs) for sets of reading are summarized in Table 2. The values revealed *fair agreement* ($k$ = 0.21–0.40) for 42% of measures, and the highest agreements (fair to moderate) were recorded between Obs1 and Obs3 ($k$ = 0.300–0.555). The lowest agreements involved Obs2 in all measures excepting ruling *(c)*. This rule showed the highest agreements, whereas rulings *(d)* and *(e)* showed the lowest agreements. Poor agreement was detected between Obs1 and Obs2 in ruling *(d)* ($k$ = −0.119).

Fleiss' $k$ values, a measure of agreement among the observers as a group, are summarized in Table 3. According to the rating scale proposed by Landis and Koch, the highest overall agreement (fair agreement) was found for ruling *(c)* and *(a)* ($k$ = 0.3751 and $k$ = 0.3429, respectively), whereas the lowest agreements were slight for *(d)* ($k$ = 0.1024) and poor for *(e)* ($k$ = −0.0447).

## Discussion

The *Vulnerability* concept has been used with different meanings in different contexts; however, they all seem to indicate the possibility that a system (and all its parts) may have a weak or limited ability to endure internal or external threats. Among the several concepts related to vulnerability are robustness, resilience, and damage tolerance, which are all "opposites" to vulnerability. *Robustness* could be defined as the system's ability to resist the event, and *Resilience* as the system's ability to return to the stable situation after the event. Although unusual, these "words" have been used in forensic arenas to a greater or lesser extent. Citing Ostrow's article from Los Angeles Times (April 9, 1977, pg. A1), Hale affirmed that "…scientific evidence is viewed as *hard* evidence by law enforcement officers and the public and, therefore, as less *vulnerable* to constitutional attack" (26). Paul C. Gianelli, a law professor at Case Western Reserve University in Cleveland, used the word "*vulnerable*" to qualify many previously accepted forensic techniques (including bite mark evidence), which today are coming under deeper attention from courts. Criminal defense lawyers have become more aggressive in challenging the admissibility of forensic evidence; they have exposed the lack of empirical support for these techniques, the main subject required today under *Daubert*, and according to Gianelli, the nature itself of this vulnerability (27). In fact, "*Strengthening* Forensic Science in the United States: A Path Forward" is a suggestive title. *Strengths* and *vulnerabilities* of forensic science evidence during trials represent the main focus of the National Academy of Sciences' (NAS) report which recommended better training and establishing better connections among experts and legal scholars and practitioners (1).

### *Case Reports*

The literature documents a long history of teeth marks or bites left in foodstuffs or inanimate objects. Our review detected nine case reports, but when we expanded the search to books and other sources not evaluated in this review, we found the most varied range of bitten substances, all of cases sentenced and many of them potentially treatable as legal precedents (Table 4). In the Common law legal system (and in several countries with a Roman law system), the principle of *Stare decisis* establishes a precedent, a point of law, that is not afterward to be departed from unless a court finds it necessary to overrule a prior case that may have been hastily decided or was decided contrary to principle. As seen, a considerable amount of case law exists in the area of bite mark evidence in different countries; beyond the results analyzed, it is interesting to note the development of some of these cases. One of the first bite mark cases reported in U.S.A. was *Doyle v. State*, which was an appellate case. Doyle was convicted on the basis of a bitten piece of cheese, and the conviction was appealed the same year on the grounds that Doyle's constitutional rights were violated. The appeals court denied Doyle relief, and this case resulted in a handful of citations but also began a list of more than 360 bite mark cases through the end of 2009 (10,14). The two experts were a dentist and a ballistics expert—Gianelli said that neither had any experience in bite mark analysis (28). In 1996, in a robbery scene, a block of cheese with a bite was taken into evidence because the victim reported that no one in her family was responsible for the bite. An odontologist conducted an analysis of the cheese, concluding that the defendant's dental impression matched the dental impression taken from the substrate found in the home. The expert was not certified by ABFO, but the court accepted the testimony, finding that there was no abuse of discretion in the district court's refusal to hold an evidentiary hearing to analyze the expert's testimony (29). In 1997 in Mississippi (U.S.A.), *Banks v. State* (30) caused much controversy when a forensic dental expert testified, as a prosecution witness, that the defendant's teeth were consistent with the marks in a piece of bologna sandwich found at the crime scene. The material was photographed and impressions taken, but the evidence was destroyed. Consequently, the defense expert was unable to reach definite conclusions. Reversing the conviction, the Mississippi Supreme Court wrote that "…the prejudicial impact of the State's destruction of the sandwich on the persuasive value of Banks' case is plainly apparent, and Dr. (…)'s destruction of

TABLE 1—*Evaluation of the 34 articles included in this review: 0 (unmentioned or insufficient); 1 (the reference is implicit or barely sufficient); and 2 (the reference is explicit or very sufficient).*

| | | | *Daubert* rulings | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | (a) | | | | (b) | | | | (c) | | | | (d) | | | | (e) | | | |
| Articles (n=34) | Country | Categories | Obs1 | Obs2 | Obs3 | Agr* | Obs1 | Obs2 | Obs3 | Agr* | Obs1 | Obs2 | Obs3 | Agr* | Obs1 | Obs2 | Obs3 | Agr* | Obs1 | Obs2 | Obs3 | Agr* |
| Fearnhead (1960) | U.K. | (II) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| Ström (1963) | Norway | (II) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| Taylor (1963) | U.K. | (III) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 |
| Layton (1966) | U.K. | (III) | 1 | 0 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 1 | 1 |
| Keiser-Nielsen (1969) | Denmark | (III) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Dinkel (1974) | U.S.A. | (II) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Solheim & Leidal (1975) | Norway | (I) | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 1 | 1 |
| Sperber (1978) | U.S.A. | (III) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| Rudland (1982) | U.K. | (II) | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 2 | 1 | 1 |
| Webster (1982) | U.K. | (II) | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 1 | 1 | 2 | 2 | 2 |
| McCullough (1983) | U.S.A. | (III) | 0 | 1 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| Kogon et al. (1990) | Canada | (I) | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 1 | 1 |
| Nambiar et al. (1995) | Malaysia/ Australia | (I) | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 0 | 1 | 1 | 1 | 2 | 1 | 1 |
| Whittaker (1995) | U.K. | (II) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| McKenna et al. (2000) | Australia | (III) | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 2 |
| Bernitz et al. (2000) | South Africa, Norway | (I) | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 2 | 1 | 1 |
| Nambiar et al. (2001) | Malaysia/ Australia | (III) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Bernitz & Kloppers (2002) | South Africa | (III) | 1 | 0 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 2 |
| Lessig et al. (2006) | Germany | (III) | 1 | 0 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 2 | 2 | 2 |
| Rai et al. (2006) | India | (I) | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Pereira et al. (2009) | Portugal/ Norway | (II) | 2 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 |
| Gorea & Jasuja (2010) | India | (I) | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 0 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| Siderits et al. (2010) | U.S.A. | (II) | 2 | 1 | 0 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| Stavrianos et al. (2010) | Greece/U.K. | (II) | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 |
| Naether et al. (2011) | Switzerland | (I) | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Mahajan et al. (2012) | India | (II) | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Nascimento et al. (2012) | Brazil | (I) | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Marques et al. (2013) | Brazil, Portugal | (I) | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Gorea et al. (2014) | S. Arabia/India/ Australia | (I) | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Pinheiro Flores et al. (2014) | Brazil | (II) | 2 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Przystańska et al. (2015) | Poland | (I) | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 2 | 1 | 1 |
| Parimala et al. (2015) | India | (II) | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 2 |
| Rathore & Sood (2015) | India | (I) | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 1 | 1 |
| Daniel et al. (2015) | India | (I) | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Rate | | | | | | 44/68 | | | | 68/68 | | | | 8/68 | | | | 20/68 | | | | 42/68 |
| Score (vulnerability) | | | | | | 64.706 (nonthreatened) (CI: 53.347 to 76.065) | | | | 100* (strong) | | | | 11.765 (very vulnerable) (CI: 4.107 to 19.422) | | | | 29.412 (vulnerable) (CI: 18.582 to 40.240) | | | | 61.765 (non-threatened) (CI: 50.214 to 73.316) |

Categories were established according to the Journal of Forensic Sciences' Author Guidelines: (I) Paper; (II) Technical Note; and (III) Case Report (21). Daubert factors: (a) whether the theory or technique is falsifiable, refutable, and/or testable; (b) whether it has been subjected to peer review and publication; (c) the known or potential error rate; (d) the existence and maintenance of standards and controls concerning its operation; and (e) whether the technique or theory has been generally accepted in the scientific community.

Obs1: Observer 1; Obs2: Observer 2; Obs3: Observer 3; Agr.*: agreement between observers (disagreements were resolved through review and discussion).

TABLE 2—*Interrater Cohen's* k *scores for pairs of raters: (a) whether the theory or technique is falsifiable, refutable, and/or testable; (b) whether it has been subjected to peer review and publication; (c) the known or potential error rate; (d) the existence and maintenance of standards and controls concerning its operation; and (e) whether the technique or theory has been generally accepted in the scientific community.*

| Ruling Criterion | Observations | | *k* value | Quality of agreement | Standard error* | Approximate T† |
|---|---|---|---|---|---|---|
| *(a)* | Obs1 | Obs2 | 0.276 | Fair | 0.127 | 2.267 |
| | | Obs3 | 0.555 | Moderate | 0.122 | 4.225 |
| | Obs2 | Obs3 | 0.211 | Fair | 0.142 | 1.627 |
| *(b)* | Obs1 | Obs2 | 0.000§ | | | |
| | | Obs3 | 0.000§ | | | |
| | Obs2 | Obs3 | 0.000§ | | | |
| *(c)* | Obs1 | Obs2 | 0.421 | Moderate | 0.187 | 2.770 |
| | | Obs3 | 0.421 | Moderate | 0.156 | 2.770 |
| | Obs2 | Obs3 | 0.288 | Fair | 0.168 | 1.881 |
| *(d)* | Obs1 | Obs2 | −0.119 | Poor agreement | 0.144 | −0.842 |
| | | Obs3 | 0.348 | Fair | 0.137 | 2.415 |
| | Obs2 | Obs3 | 0.131 | Slight | 0.150 | 0.868 |
| *(e)* | Obs1 | Obs2 | 0.038 | Slight | 0.029 | 0.809 |
| | | Obs3 | 0.300 | Fair | 0.169 | 2.576 |
| | Obs2 | Obs3 | 0.136 | Slight | 0.075 | 1.429 |
| Number of valid cases | | | 34 | | | |

*Not assuming the null hypothesis.
†Using the asymptotic standard error assuming the null hypothesis.
§Not calculated because (b) was a constant.

TABLE 3—*Interrater Fleiss'* k *scores for multiple raters: (a) whether the theory or technique is falsifiable, refutable, and/or testable; (b) whether it has been subjected to peer review and publication; (c) the known or potential error rate; (d) the existence and maintenance of standards and controls concerning its operation; and (e) whether the technique or theory has been generally accepted in the scientific community.*

| Ruling criterion | Response | *k* value | Quality of agreement | 95% CI* | | *z*-stat | *p* (vs > 0) |
|---|---|---|---|---|---|---|---|
| *(a)* | 0 | 0.2066 | | −0.2590 | 0.6699 | 2.0862 | 0.0370 |
| | 1 | 0.2897 | | 0.0485 | 0.5303 | 2.9263 | 0.0034 |
| | 2 | 0.4706 | | 0.2288 | 0.7117 | 4.7527 | 0.0000 |
| | Overall | 0.3429 | Fair | 0.1332 | 0.5518 | 4.4933 | 0.0000 |
| *(b)* | 0 | 0.000† | | | | | |
| | 1 | 0.000† | | | | | |
| | 2 | 0.000† | | | | | |
| | Overall | 0.000† | | | | | |
| *(c)* | 0 | 0.4386 | | 0.1408 | 0.7354 | 4.4300 | 0.0000 |
| | 1 | 0.3159 | | −0.0207 | 0.6510 | 3.1900 | 0.0014 |
| | 2 | 0.3131 | | −0.4280 | 1.0444 | 3.1625 | 0.0016 |
| | Overall | 0.3751 | Fair | 0.0971 | 0.6520 | 4.2507 | 0.0000 |
| *(d)* | 0 | 0.1289 | | −0.1041 | 0.3612 | 1.3016 | 0.1930 |
| | 1 | 0.0977 | | −0.1273 | 0.3220 | 0.9866 | 0.3238 |
| | 2 | −0.0408 | | −0.817 | 0.0000 | −0.4122 | 0.6802 |
| | Overall | 0.1024 | Slight | −0.1097 | 0.3138 | 1.1480 | 0.2510 |
| *(e)* | 0 | −0.0099 | | | | −0.1000 | 0.9203 |
| | 1 | −0.0603 | | −0.2854 | 0.1642 | −0.6089 | 0.5426 |
| | 2 | −0.0303 | | -0.2606 | 0.1993 | −0.3060 | 0.7596 |
| | Overall | −0.0447 | Poor | | | −0.4652 | 0.6418 |
| Number of valid cases: | | 34 | | | | | |
| Number of responses: | | 3 | | | | | |
| Number of observers/raters: | | 3 | | | | | |

*Jackknife confidence interval.
†Not calculated because (b) was a constant.

the sandwich was unnecessary and inexcusable" (7,28,30). Although we agree with Gianelli that the statement "*Daubert* is likely to have little effect on bite mark admissibility" is "highly debatable" (28), the fact is that the courts still continue to admit this kind of evidence (31). The impact of *Daubert* has not yet been fully assessed; little attention has been paid to satisfying the legal requirements, and published evidence of the reliability and validity of bite mark analysis is required (9,17). The lack of strong, well-designed studies on bite marks is frequently alluded to in the courts (especially in the U.S.A.), and "…case reports, reviews and anecdotal commentaries cannot address this deficit" (32).

*Technical Notes and Papers*

Several methods of bite mark analysis have been reported, all involving three steps: (i) reproductions of both the bite mark and the suspect's dentition through a variety of methods; (ii) direct or indirect comparison of the dentition and bite mark; and (iii) evaluation of the points of similarity or dissimilarity (although there is no minimum number of points of identity, some experts have accepted minimum ranges from 8 to 52 points) (31). Although at first glance bite marks in foodstuffs and objects can provide better analysis than bite marks in skin (8), specific problems may arise depending on the type of material (33,34). A bitten apple

TABLE 4—*Historical and current reports of bitten foodstuffs and inanimate objects, all of them prosecuted, sentenced, and potentially legal precedent cases.*

| Year of Case | Reference | Country | Substrate |
|---|---|---|---|
| 1872 | Senn (2011) (10) | U.K. | Apple* |
| † | Senn (2011) (10) | Russia | Cigar holder |
| 1905 | Senn (2011) (10) | Germany | Cheese |
| 1906 | Senn (2011) (10) | U.K. | Cheese |
| 1924 | Sen (2011) (10) | Canada | Apple |
| 1966 | Layton (1966) (3) | U.K. | Cheese |
| 1954 | Dinkel (1974) (33); Rai et al. (2006) (49); Senn (2011) (10); Committe… (2011) (31) | U.S.A. | Cheese |
| † | Keiser-Nielsen (1969) (35) | Denmark | Apple |
| 1970 | Dinkel (1974) (33) | Japan | Apple |
| 1976 | Sperber (1978) (50) | U.S.A. | Chewing gum |
| 1979 | Silver & Souviron (2009) (7) Lipton et al. (2013) (51) | U.S.A. | Bologna |
| 1982 | McCullough (1983) (52) | U.S.A. | Cheese |
| 1985 | Committe… (2011) (31) | U.S.A. | Apple |
| 1996 | McKenna et al. (2000) (5) | Australia | Chocolate |
| 1996 | Committe… (2011) (31) | U.S.A. | Cheese |
| 1990 | Nambiar et al. (2001) (6) | Australia | Chewing gum |
| 1997 | Committe… (2011) (31) Gianelli (2007) (28) | U.S.A. | Bologna sandwich |
| 1998 | Bernitz et al. (2000) (4); Senn (2011) (10) | South Africa | Cheese** |
| 1999 | Bernitz & Kloppers (2002) (30); Senn (2011) (10) | South Africa | Cheese |
| 2000 | Committe… (2011) (31) | U.S.A. | Cheese |
| 2002 | Senn (2011) (10) | South Africa | Apple |
| † | Lessig et al. (2006) (53) | Germany | Apple |
| 2007 | Pereira et al. (2009) (54) | Portugal | Cheese |
| † | Bernstein (2011) (34) | U.S.A. | Chewing gum |
| † | Bernstein (2011) (34) | U.S.A. | Hardwood |
| † | Dorion (2011) (12) | U.S.A. | Block of hashish |
| † | Dorion (2011) (12) | U.S.A. | Pacifier |
| † | Dorion (2011) (12) | U.S.A. | Pencil |

*The evidence was poorly managed and the accused had strong political influence. He was acquitted (10).

†Unreported date.

**The evidence did not convince the High Court of South Africa (4,10).

may show class characteristics of the biter and some individual tooth arrangements, while a piece of cheese can not only show those class characteristics but can also individualize tooth characteristics (7). Dorion stated that semihard chocolate is probably one of the best bite mark impression perishables because of its three-dimensional stability (12). However, most of the references are subjective or case-dependent by nature (7), with very few exceptions that analyze not only the quality of marks but also the dimensional changes and preservation methods under variable conditions (12). In early 1969, Keiser-Nielsen stated “bite marks in foodstuffs should be photographed and duplicated in plaster of Paris as soon as possible. The dentist has a wide variety of impression materials from which to choose and he is skilled in handling them” (35). The classical methods of classifying bite marks in foodstuffs (11), although very general (12), have not been improved and are still effective. *The existence and maintenance of standards and controls concerning its operation* seems to be a *Daubert* rule with little updating, at least as regards by ABFO or ASFO guidelines (13,14). Pretty & Sweet affirmed that the lack of direction in European and American forensic dental organizations complicates this matter (32).

Although the increase of *Papers* in the last 5 years is clear (8/13), this has not been reflected in the improvement of their quality, according to the rules evaluated (Table 1). As seen above, *Case Reports* often continue to represent the gold standard on how to proceed or what technique to use in those situations. In 2011, Dorion remembered the 1977 bite mark survey of the American Academy of Forensic Sciences, Odontology Section; the 45 contributors listed only 13 cases of bite marks in inanimate objects. According to Dorion, the report concluded “the amount of information received was insufficient to form reliable conclusions for evaluating procedural and technical methods employed” (12). Undoubtedly the scarcity of experimental or empirical work cannot be balanced only with (early) case reports. However, there are excellent references on how criminal cases can affect/modify/improve a procedure from an experimental point of view. In *State v. Shabangu*, a murder case in 1998, a piece of cheese with very clear teeth marks analyzed by the experts Dr. Piet van Niekerk, Professor Vince Phillips, and Professor Herman Bernitz, was rejected by the High Court of South Africa because the court was only prepared to accept the match between the cheese bite marks and the suspect as substantive evidence. The points questioned were as follows: which dental features are common, uncommon, and very uncommon within the relevant populations; it is necessary to refine the pattern association analysis of bite marks, explaining how warping, shrinkage and distortion do not affect it; and the need of use of metric analysis and of microscopic analysis in individualizing features within the bite marks (4,10). Back in South Africa one year later, *State v. Nxele* was heard in the Pietermaritzburg High Court; this time, the inclusion of a ballistics expert, Senior Superintendent Burgert Kloppers, in the bite mark team allowed a piece of cheese with clear teeth marks to be analyzed through a DMC comparison microscope (used for analyzing firearm barrels and bullets) enhancing the ability of the forensic odontologists to present evidence correctly (10,36). In 1974, Dinkel concluded that criminal cases could force reassessments of both protocols and analysis methodologies (33). We believe that *Technical Notes* and *Papers* should draw on *Case Reports* and lead a healthy synergy with each other. The contribution of Bernitz et al. (2000) (4) represents exactly that point and is one of the best scored in this review.

## *Error Rates*

Avon affirmed that even under carefully controlled conditions, albeit in a forced-decision model, errors in interpretation occur even among the most experienced observers (37). As there are advantages and disadvantages for each method used in bite mark comparison, the current standard in forensic odontology is to use two methods of comparison, and forensic odontologists should be familiar with all methods of bite mark comparisons and appropriately use those methods that are indicated for the study case (14). This recommendation is part of an accepted convention implicitly or explicitly mentioned in almost all guidelines: “…it is better to be conservative in both the approach to an analysis and final opinion rendered concerning the value of the bite mark evidence and the correlation with any suspected biter(s)” (14). However, even though the examiner may have great experience, the pool of possible biters may be small or the bite mark pattern may demonstrate sufficient characteristics (leading to an obvious, logical, and understandable analysis) (14), the expert opinion must be based on scientifically derived techniques in which comparisons have been used to calculate error rates where possible (37).

The definition of the error in an expert opinion has become relevant after *Daubert,* and the NAS report exposed the problems of accuracy, reliability and validity of some forensic disciplines (1). Although the concept of error is often vague and subject to a variety of interpretations, the known error rate to which *Daubert* refers can include a number of things such as the confidence interval, the statistical significance of a result, or the probability that a reported conclusion is incorrect (38). Christensen et al. (2014) (38) emphasized that the measures to minimize and account for error and limitations in the procedures should now be apparent, and must be appropriately communicated. The authors added that in the absence of known error rates for a procedure, it is not acceptable to derive error rates from practitioner proficiency tests, professional exercises, or studies that were not designed to estimate method error rates (38).

Although considerable efforts have been devoted to determining the level of expert training, the subjective nature of evaluations and the statistical analysis developed (including error rates) in bite marks in experimental models, these studies have focused on the skin injuries (37,39). Our review detected only one *Paper* with the highest score on the topic: In 1995, Nambiar et al. evaluated quantitatively the suitability of using shape analysis methodology on bite marks produced in selected foodstuffs. Their conclusion reported success rates of 33% in apples, 38% in cheddar cheese and 50% in chewing gum. Chocolate-coated confectionery bars could not be read usefully. The bite marks were compared with the dental casts of the perpetrators ("self" matches) and with dental casts of three other subjects ("nonself" matches) (40). Incidentally, this is another remarkable case where a forensic dental team has managed to link a *Paper* with *Case Reports* (5,6). Of the other *Papers*, only 6 referenced an implicit error rate, which does not seem to relate to the timeline. For such studies the known error rate, or an experimental design that may expose it, do not appear to be a priority. On the other hand, neither *Technical Notes* nor *Case Reports* have mentioned any known error rates (Table 1). It is extremely important to emphasize that there is no doubt about the fallacy of giving the same weight to case reports, technical notes and papers considering the hierarchy that each has as scientific evidence (41). However, and beyond the recognized weakness of case reports and technical notes (as opposed to the more robust papers), we do not consider this a limitation of our study. As we seen, judicial arenas has raised other ways of seeing this "weak" evidence: In particular, the case studies may help to identify changes in criminal profiles, rare types of committing criminal acts, and unusual wound patterns, or to establish new criteria for giving expert evidence (42). We agree with Madea in that even rare cases are not unique cases, they cannot be simulated experimentally for ethical reasons and their interdisciplinary reconstruction can reveal peculiarities. The author assess that "the role of case histories is mainly the augmentation of experience based knowledge" (42). Added to this, and perhaps even more important, we consider that this particular scenario poses new rules of game where delineating boundaries between real science and junk science is sometimes left to judges or evaluators without skills or without objective instruments to do this. We sincerely hope that this study will at least give rise to call to attention on the way in which the scientific literature can be invoked in trial, and as will be seen later, to propose objective elements of good judgment to evaluate it.

Beecher-Monas (43), citing several authors, stated that in each step of a bite mark analysis (making overlays, photographing, and all tracings), errors could be introduced in various ways. She added that even if more objective techniques are attempted, comparison methods are ultimately subjective processes, and she concluded that although error rates appear to be high, they have never been rigorously quantified (42). Christensen et al. (38) concluded that it is imperative for researchers and practitioners to have a thorough understanding of the various concepts of error, and recommended strongly that educational programs in forensic sciences as well as training programs for practitioners should address error and error analysis. Very pessimistically, Beecher-Monas stressed that the courts do not seem to take error into account even when experts declare high error rates (42). Avon, after calculating error rates using an animal model for human dermal bite marks, noted that attributing the wrong dentition to a bite mark constitutes a critical error as it would be analogous to inculpating an "innocent" person (37).

It is important to mention some unavoidable limitations of this study: The authors are neither lawyers nor North Americans, considering the origin and application of the *Daubert* rulings. This could explain the differences in the overall observers' appreciation (Fleiss' $k$ scores) rating the existence or not of standards and controls of the techniques used (ruling *d*) or even the acceptance of the technique in the scientific community (ruling *e*). This last point in particular could be easily interpreted (and resolved, in fact) by the publication itself; however, and as mentioned above, there are still innumerable prejudices and unresolved elements that cast doubt on this assessment. Although the agreement has acceptable margins, this is further emphasized by the remarkable difference in forensic experience reflected in the differences of agreement between Observer 2 and the others (Cohen's $k$ scores). Even though U.S. standards have proven to be a strong reference not only at the level of forensic dental research (it is impossible to deny the influence of ABFO guidelines in international protocols) (44) but also in general legal terms (16–18), when the underlying proofs for validity and judicial acceptance in accordance with the *Daubert* factors are analyzed and discussed, the evident uneven globalization of cases, technical notes and studies (see Fig. 2) could represent an incorrect, or at least insufficient, training and scientific basis for this type of evidence, both for non-U.S. researchers and even for non-U.S. courts. Standardization is undoubtedly a precious treasure when seeking quality assurance of the expert testimony; however, assuming the responsibility "for standards and especially to avoid large variations from one country to another," the International Organization for Forensic Odonto-Stomatology (IOFOS is the only worldwide forensic odontology organization) stated that quality assurance in forensic dentistry is difficult to define and implement, the practices vary significantly between countries, and defining its steps according to specific protocols seems to be an impossible task to carry out at an international level (45). IOFOS affirms that "quality assurance should never be static but be reconsidered and if possible always improved," and suggested that work and recommendations be reconsidered from time to time (45). This is probably the greatest scope of our study: to address a specific forensic topic, to review almost everything published about it, and to try to analyze it, detect inconsistencies, or validate it according to the most widely recognized current measure at the global level. Beyond its explicit value in the American courts, under *Daubert*, the recognized guidelines for admitting scientific expert testimony—that the methods should have a reliable application to the facts and the methodology should be a product derived from the scientific method—are universally recognized as requirements for scientific statements made by experts. *Daubert* represents a keystone of advice to the courts that any technical theory, such as bite mark evidence, can be and has been tested and that the science underlying such techniques has been published the peer-reviewed literature. However, it emphasizes that courts must be careful to balance skepticism and flexibility; publication is no guarantee of reliability, and the sufficiency of early decisions should be retested for admissibility (17). It has been stated that researchers should take more vigorous steps to establish the scientific basis of procedures, and the court should act as true sentinel of the evidence presented at trial, as strongly recommended by the NAS (17,38).

It would be very pretentious to propose a "correct way" to analyze the documents included in this study, taking into account that each of *Daubert* factors still represents assignments in controversy and dependent on the observers. In particular, although the presence of an explicit known or potential error rate—the ruling *(c)*—may be simple to verify in a published document, the NAS stated that bite mark testimony has been "introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing to explain the limits of the discipline" (1). As far beyond that the concept of "error" can be confusing or even misused both in courtrooms and in forensic research, we agree with Christensen et al. (38) in that "as forensic scientists, we must concerned with the clarity, reliability, and validity of our methods." This statement is emphatically reaffirmed by Saks et al., who added in a very recent article (46) that "careful research would need to be designed in order to isolate the various possible causes of the errors and to try to develop ways to reduce errors stemming from those causes." We agree with the authors that this should be applicable to both false positives and false negatives, so that forensic scientists (beyond their expected depositions at trial) should make explicit in their research results not only the error rates but also the most demanded claim made by the scientific community: the validation of a field's technique, still absent in bite mark identification (46). Specifically in the field of bite mark analysis in foodstuffs and inanimate objects, we consider that publications (the obvious result of scientific research), still lack error rates and validated methodologies, and we believe that they are keys factors in solving the equation.

Although the ruling *(b) "whether it has been subjected to peer review and publication"* may be the easiest to assess (the publication supposes the required peer review—assuming the journal has a panel of reviewers), the other criteria may be dependent on the qualification of the observer (as shown in Tables 2 and 3). Beyond *Daubert*, the U.S. Federal Rule of Evidence 702 governs the admission of expert testimony, where a witness may qualify as an expert by knowledge, skill, experience, training or education (the rule does not distinguish between "scientific," "technical," or "other specialized" knowledge). Unfortunately, not all countries have such standards of evidence admissibility (at least they do not exist in our countries), so that the assessment must be delegated to the judge's discretion or common sense. This is undoubtedly another great weakness of our system.

As far as the ruling *(a)*, falsifiability, refutability, and/or testability of the theory should be positively valued, in our opinion, only if the document explicitly—as a cooking recipe—each of the steps to reproduce the technique and be able to contrast it. The rulings *(d)*—the standardization—and *(e)*—the acceptance by the scientific community—were, in our opinion, the most difficult criterions to evaluate. We believe that the reason for this is simple to understand: Science evolves, paradigms change, the thresholds of appropriate standardization are dynamic (and as seen, depending on the geographical and historical context), and obviously the observer qualification can result fundamental to correctly evaluate the state of art and the know-how of a specific technique. As scientists, it was not easy for us to evaluate these last two rules—even considering all the literature within our reach; we are sure that asking this task to the judges may not be sufficiently consistent or accurate. We were unable to find a useful forensic tool to assess the methodological quality of the identified articles, so PRISMA® was only used to guide the search and inclusion strategy (another obvious limitation of this study). Although it is clear that because of the nature and objectives of this study, *Daubert* has been used as a tool for this assessment, it is necessary to propose some kind of tool that allows the objective evaluation of the methodological quality of a technique or procedure proposed in courts. We agree with the NAS report (1) in that the minimum set of information for properly specifying the process of any new analytical method, included in the guidelines presented in ISO/IEC 17025 (47), can provide an excellent basis for attempting to standardize this matter.

Bite mark evidence has represented the greatest successes and the most resounding failures of forensic dentistry (44). The reasons in both cases have been the possibility/impossibility of proving the uniqueness of dentition, and the possibility that it may have transferred to the bitten substrate. However, the concept of uniqueness has been severely questioned, claiming that this it is "largely irrelevant" to forensic practice and to the legal system (48). Page et al. (48) stated that mistakes and misidentifications "are made because of guesswork, poor performance, lack of standards, bias and observer error".

We do not agree with the pessimistic position of Saks et al. (46) when declaring the "impeding fall of bite mark evidence," as they themselves limit the criminal context only to the finding of bite marks in the substrate skin. As we have seen, there are a large number of studies and cases of other substrates that offer the possibility of recording the dental pattern more reliably. However, we have also seen that there is minimal (or even null) attention in trying to give specific "step-by-step" protocols to analyze these perishables (ABFO, ASFO, BAFO, and IOFOS do not have them). Nor do we defend simplistic positions; it is evident that the paradigm has taken the bow to the objective, quantitative valuation of the expert testimony. We consider this an invaluable opportunity to learn from what has historically been done to strengthen the bite mark investigation, at least on these substrates. Setting the height of the bar for objectively establishing the critical limit separating "real science" from "junk science" in legal arenas can be extremely difficult. Science evolves, paradigms fall and rise again, and the concept of jurisprudence may be hiding a bias in itself. From that point of view, the only constant is the current contribution of scientific method.

## Conclusions

Bite mark analysis in inanimate objects and foodstuffs seems to offer more reliability than in the skin, and numerous *Papers*, *Technical Notes* and *Case Reports* have addressed different procedures to identify a perpetrator in the forensic arenas. However, in light of current standards of the admissibility of evidence, there is great vulnerability in all steps of these procedures. As controversy over bite mark analysis still continues to cause skepticism in court, it is essential to increase the demand for a methodological scientific basis when such evidence is presented in trials, focusing scrutiny on the explicit presentation of known error rates. Without falling into extreme positions, forensic dental researchers should update their methodologies recognizing these vulnerabilities. Although bite mark analyses may be strongly disputed, there is no certainty that their usefulness will be rebutted. The robustness of a method is achieved not only by exposing its success rates but also by strengthening its vulnerabilities, and especially, recognizing its potential error rates.

*Acknowledgments*

The authors wish to thank Dr. Carlos Zaror S. (CIEESPO, Faculty of Dentistry, Universidad de La Frontera, Temuco, Chile) for his helpful suggestions and assistance with statistical methods.The authors would also like to thank to Prof. Herman Bernitz (University of Pretoria, South Africa), to Prof. Stanley Kogon (University of Western Ontario, Canada) and especially to Prof. Helen James (University of Adelaide, Australia) for their invaluable contributions and submission of many of the articles reviewed in this research.

The authors are also thankful to all reviewers for their comments, assistance, and support reviewing comprehensively the Westlaw database.

## References

1. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. Strengthening forensic science in the United States: a path forward. Washington, DC: National Academies Press, 2009.
2. 159 Tex.Crim. 310, 263 S.W.2d 779 (1954).
3. Layton JJ. Identification from a bite mark in cheese. J Forensic Sci Soc 1966;6(2):76–80.
4. Bernitz H, Piper S, Solheim T, Van Niekerk PJ, Swart TJ. Comparison of bitemarks left in foodstuffs with models of the suspects' dentitions as a means of identifying a perpetrator. J Forensic Odontostomatol 2000;18(2):27–31.
5. McKenna CJ, Haron MI, Brown KA, Jones AJ. Bitemarks in chocolate: a case report. J Forensic Odontostomatol 2000;18(1):10–4.
6. Nambiar P, Carson G, Taylor JA, Brown KA. Identification from a bitemark in a wad of chewing gum. J Forensic Odontostomatol 2001;19(1):5–8.
7. Silver WE, Sourivon RR. Dental autopsy. Boca Raton, FL: CRC Press, 2009.
8. Dailey JC. Methods of comparison. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text, 2nd edn. Boca Raton, FL: CRC Press, 2011;469–89.
9. Pretty IA, Sweet D. The judicial view of bitemarks within the United States criminal justice system. J Forensic Odontostomatol 2006;24(1):1–11.
10. Senn DR. History of bitemark evidence. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text. 2nd edn. Boca Raton, FL: CRC Press, 2011;3–22.
11. Webster G. A suggested classification of bite marks in foodstuffs in forensic dental analysis. Forensic Sci Int 1982;20(1):45–52.
12. Dorion RBJ. Nonperishables and perishables. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text. 2nd edn. Boca Raton, FL: CRC Press, 2011;159–63.
13. American Board of Forensic Odontology. ABFO diplomates reference manual. March 2016 edn., http://www.abfo.org (accessed June 14, 2016).
14. Dailey JC, Golden GS, Senn DR, Wright FD. Bitemarks. In: Senn DR, Weems RA, editors. ASFO manual of forensic odontology, 5th edn. Boca Raton, FL: CRC Press, 2013;257–324.
15. 509 U.S. 579 (1993).
16. De Luca S, Navarro F, Cameriere R. La prueba pericial y su valoración en el ámbito judicial español. RECPC 2013;19:e1–14; http://criminet.ugr.es/recpc/15/recpc15-19.pdf (accessed September 18, 2016).
17. Pretty IA. Reliability of bitemark analysis. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text, 2nd edn. Boca Raton, FL: CRC Press, 2011;587–99.
18. Bórquez P. Forma y fondo del peritaje médico legal. Rev Med Chile 2009;137:852–7.
19. Barsley RE. Case law. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text, 2nd edn. Boca Raton, FL: CRC Press, 2011;517–41.
20. Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336–41.
21. Journal of Forensic Sciences. Author guidelines; http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1556-4029/homepage/ForAuthors.html (accessed December 10, 2016).
22. Cohen JA. Coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20(1):37–46.
23. Fleiss JL, Nee JCM, Landis JR. Large sample variance of kappa in the case of different sets of raters. Psychol Bull 1979;86(5):974–7.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–74.
25. Solheim T, Leidal TI. Scanning electron microscopy in the investigation of bite marks in foodstuffs. Forensic Sci 1975;6(3):205–15.
26. Hale A. The admissibility of bite mark evidence. South Calif Law Rev 1978;51:309–34.
27. Hansen M. The uncertain science of evidence. ABA J 2005;9(7):48–53; http://www.abajournal.com/magazine/article/the_uncertain_science_of_evidence (accessed September 18, 2016).
28. Giannelli PC. Bite mark analysis. School of Law, Case Western Reserve University. Faculty Publications. Paper 153, 2007; http://scholarlycommons.law.case.edu/faculty_publications/153 (accessed September 18, 2016).
29. *Seivewright, III v. State*, 7 P.3d 24 (Wyo. 2000).
30. 725 So.2d 711 (Miss. 1997).
31. Committee on the Development of the Third Edition of the Reference Manual on Scientific Evidence; Committee on Science, Technology, and Law; Policy and Global Affairs; Federal Judicial Center; National Research Council. Reference manual on scientific evidence. 3rd edn. Washington DC: National Academies Press, 2011.
32. Pretty IA, Sweet D. The scientific basis for human bitemark analyses – a critical review. Sci Justice 2001;41(2):85–92.
33. Dinkel EH Jr. The use of bite mark evidence as an investigative aid. J Forensic Sci 1974;19(3):535–47.
34. Bernstein ML. The nature of bitemarks. In: Dorion RBJ, editor. Bitemark evidence. A color atlas and text, 2nd edn. Boca Raton, FL: CRC Press, 2011;53–65.
35. Keiser-Nielsen S. Forensic odontology. U Tol Law Rev 1969;3:633–51.
36. Bernitz H, Kloppers BA. Comparison microscope identification of a cheese bitemark: a case report. J Forensic Odontostomatol 2002;20(1):13–6.
37. Avon SL, Victor C, Mayhall JT, Wood RE. Error rates in bite mark analysis in an in vivo animal model. Forensic Sci Int 2010;201(1–3):45–55.
38. Christensen AM, Crowder CM, Ousley SD, Houck MM. Error and its meaning in forensic science. J Forensic Sci 2014;59(1):123–6.
39. Molina A, Martin-de-las-Heras S. Accuracy of 3D scanners in tooth mark analysis. J Forensic Sci 2015;60(Suppl 1):S222–6.
40. Nambiar P, Bridges TE, Brown KA. Quantitative forensic evaluation of bite marks with the aid of a shape analysis computer program: Part 2; "SCIP" and bite marks in skin and foodstuffs. J Forensic Odontostomatol 1995;13(2):26–32.
41. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. J Clin Nurs 2003;12(1):77–84.
42. Madea B. Case histories in forensic medicine. Forensic Sci Int 2007;165(2–3):111–4.
43. Beecher-Monas E. Reality bites: the illusion of science in bite-mark evidence. Cardozo Law Rev 2009;30(4):1369–410.
44. Fonseca GM, Briem-Stamm AD, Cantín M, Lucena J, Bentkovski A. Odontología forense I: Huellas de Mordedura. Int J Odontostomat 2013;7(1):149–57.
45. Solheim T, Kvaal S, Grusd L, Stene-Johansen W. Quality assurance in forensic odonto-stomatology. General; http://www.iofos.eu/Quality-Ass/general.htm (accessed April 26, 2017).
46. Saks MJ, Albright T, Bohan TL, Bierer BE, Bowers CM, Bush MA, et al. Forensic bitemark identification: weak foundations, exaggerated claims. J Law Biosci 2016;3(3):538–75.
47. International Organization for Standardization (ISO). ISO/IEC 17025:2005. General requirements for the competence of testing and calibration laboratories, http://www.iso.org/iso/catalogue_detail?csnumber=39883 (accessed December 10, 2016).
48. Page M, Taylor J, Blenkin M. Uniqueness in the forensic identification sciences – fact or fiction? Forensic Sci Int 2011;206(1–3):12–8.
49. Rai B, Anand S, Madan M, Dhattarwal S. Bite marks: a new identification technique. Int J Forensic Sci 2006;2(1):e1–5.
50. Sperber ND. Chewing gum - an unusual clue in a recent homicide investigation. J Forensic Sci 1978;23(4):792–6.
51. Lipton BE, Murmann DC, Pavlik EJ. History of forensic odontology. In: Senn DR, Weems RA, editors. ASFO manual of forensic odontology, 5th edn. Boca Raton: CRC Press, 2013;1–40.
52. McCullough DC. Rapid comparison of bite marks by xerography. Am J Forensic Med Pathol 1983;4(4):355–8.

53. Lessig R, Wenzel V, Weber M. Bite mark analysis in forensic routine case work. EXCLI J 2006;5:93–102.
54. Pereira C, Santos JC, Solheim T. Evidence collection of a tooth mark in a crime scene: importance of the dental materials in forensic dentistry. Rev Port Estomatol Cir Maxilofac 2009;50:141–4.

Additional information and reprint requests:
Gabriel M. Fonseca, D.D.S., Ph.D.
Laboratory of Forensic Dentistry, Faculty of Dentistry,
Universidad de La Frontera
Francisco Salazar 01145, Building L
4811230 Temuco
Chile
E-mail: gabriel.fonseca@ufrontera.cl