



## Investigación original

# Concordancia diagnóstica del sistema de graduación Chisholm Mason y Focus Score en la sialoadenitis linfocítica focal como aproximación al síndrome de Sjögren

L.F. Peña Carvajalino<sup>a,\*</sup>, J. Correa Rodríguez<sup>a</sup>, J.F. Polo Nieto<sup>a</sup>, C.I. Ibáñez<sup>b</sup>, A. Escobar<sup>c</sup> y A. Rojas-Villarraga<sup>b</sup>

<sup>a</sup> Departamento de Patología, Fundación Universitaria de Ciencias de la Salud (FUCS), Hospital de San José, Hospital Infantil Universitario de San José, Bogotá, Colombia

<sup>b</sup> División de Investigaciones, Fundación Universitaria de Ciencias de la Salud (FUCS), Bogotá, Colombia

<sup>c</sup> Departamento de Reumatología, Fundación Universitaria de Ciencias de la Salud (FUCS), Hospital de San José, Bogotá, Colombia

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 2 de marzo de 2020

Aceptado el 9 de junio de 2020

On-line el 28 de julio de 2020

### Palabras clave:

Glándula salivar menor

Síndrome de Sjögren

Sialoadenitis linfocítica

## R E S U M E N

**Introducción:** Existen diferentes sistemas de graduación para interpretar la biopsia de glándula salivar menor (BGSM) en pacientes con síntomas secos. En la actualidad la metodología recomendada (Focus Score [FS]), aunque no se emplea de forma rutinaria, conlleva discrepancias diagnósticas.

**Objetivo:** Determinar la concordancia de la lectura histopatológica de BGSM entre las técnicas FS y Chisholm Mason (CM) para el diagnóstico de sialoadenitis linfocítica focal en el hospital de San José, entre enero de 2017 y junio de 2018.

**Métodos:** Estudio de concordancia. Se incluyeron BGSM de pacientes con síntomas secos. Se excluyó material no apto para el estudio, o con menos de 4 lóbulos o menos de 8 mm de área. Se clasificaron utilizando las técnicas FS y CM por 2 evaluadores independientes. Se evaluó el acuerdo interobservador, intraobservador y global entre las pruebas mediante el coeficiente de Kappa de Cohen.

**Resultados:** Se evaluaron 130 biopsias. Los valores K intraobservador entre FS y CM fueron 0,42 para el observador 1 y 0,23 para el observador 2. Los valores K del acuerdo interobservador fueron 0,47 para FS y 0,65 para CM. La concordancia entre los 2 sistemas de puntuación fue K de 0,13.

**Abreviaturas:** BGSM, biopsia de glándula salivar menor; CM, Chisholm Mason; FS, Focus Score; SLF, sialoadenitis linfocítica focal; SS, Síndrome de Sjögren.

\* Autor para correspondencia.

Correo electrónico: [lfpena@fucsalud.edu.co](mailto:lfpena@fucsalud.edu.co) (L.F. Peña Carvajalino).

<https://doi.org/10.1016/j.rcreu.2020.06.003>

0121-8123/© 2020 Asociación Colombiana de Reumatología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

**Conclusiones:** Se evidencia una probable sobreestimación de focos con CM, que resalta sus debilidades, y mayor precisión con FS. Por ende, la concordancia entre las 2 pruebas es débil. Es necesario un mayor esfuerzo para el uso del método estandarizado de lectura de BGSM con el FS para una correcta interpretación y beneficio en la clasificación de pacientes con síndrome seco.

© 2020 Asociación Colombiana de Reumatología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Diagnostic agreement of the Chisholm Mason and Focus Score grading system in focal lymphocytic sialadenitis in patients with dry symptoms

### A B S T R A C T

#### Keywords:

Salivary glands

Minor

Sjögren's syndrome

Sialadenitis

**Introduction:** There are different grading systems for interpreting the minor salivary gland biopsy (MSGB) in patients with dry symptoms (Sicca). The recommended methodology is currently the Focus Score (FS), which, although not performed routinely, results in diagnostic discrepancies.

**Objective:** To determine the concordance of the histopathological reading of MSGB between the FS and Chisholm Mason (CM) techniques for diagnosis of focal lymphocytic sialadenitis at the San José Hospital between January 2017 and June 2018.

**Methods:** Concordance study. MSGB of patients with sicca symptoms were included. Material not suitable for study and / or with less than 4 lobules and / or less than 8 mm of surface area was excluded. They were classified, using the FS and CM techniques, by two independent evaluators. The inter-observer, intra-observer, and overall agreement between the tests were evaluated using Cohen's Kappa coefficient.

**Results:** A total of 130 biopsies were evaluated. The intra-observer K values between FS and CM were 0.42 for observer 1, and 0.23 for observer 2. The K values of the inter-observer agreement were 0.47 for FS, and 0.65 for CM. The concordance between the two scoring systems was  $K = 0.13$ .

**Conclusions:** There is evidence of a probable over-estimation of foci with CM, highlighting its weaknesses and greater precision with FS technique, resulting in the weak concordance between the two tests. A greater effort is necessary for the use of the standardised MSGB FS reading method for a correct interpretation and benefit in the classification of patients with dry syndrome.

© 2020 Asociación Colombiana de Reumatología. Published by Elsevier España, S.L.U. All rights reserved.

## Introducción

El síndrome de Sjögren (SS) es una enfermedad autoinmune cuya principal manifestación clínica es la sequedad a nivel glandular (xerostomía y xeroftalmia), aunque también puede presentar manifestaciones clínicas en otros órganos o sistemas. La mayoría de las veces se presenta como una enfermedad aislada o asociada con otra enfermedad autoinmune (poliautoinmunidad)<sup>1</sup>. Afecta preferiblemente a mujeres durante la cuarta y la quinta décadas de la vida<sup>2</sup>. Se considera que es la segunda enfermedad autoinmune más prevalente a escala global, con una incidencia anual de 6,92 casos/100.000 personas-año y una prevalencia de 60,82 casos/100.000 personas<sup>3</sup>. Su etiología es multifactorial<sup>4</sup> y se caracteriza por la presencia de infiltración mononuclear progresiva de las glándulas exocrinas. Es una de las enfermedades autoinmunes que generan mayor compromiso en la calidad de vida de quienes la padecen, incluyendo manifestaciones severas extraglandulares y malignas glandulares<sup>5</sup>.

Dentro de los diversos criterios de clasificación propuestos se cuenta con los criterios actualizados del *American College of Rheumatology/European League Against Rheumatism*<sup>6</sup>, que incluyen pruebas subjetivas y objetivas que demuestran la disfunción glandular, la presencia de autoanticuerpos específicos y el estudio histopatológico de la glándula salival menor. Sin embargo, las distintas clasificaciones a lo largo de la historia tienen en común la presencia de autoanticuerpos y la biopsia de glándula salival menor para su diagnóstico<sup>7,8</sup>. El papel diagnóstico de la histopatología se ha considerado un «estándar de oro». A pesar de tratarse de un procedimiento quirúrgico menor, es un criterio importante para la clasificación precisa de la enfermedad. La interpretación de la prueba es difícil y requiere un muestreo, un procesamiento y una interpretación que sean adecuados<sup>6</sup>.

Se han propuesto diferentes metodologías para interpretar los hallazgos en la biopsia de glándula salival menor (BGSM), y de esa manera poder categorizar la enfermedad. En la actualidad, al hallazgo de un *focus score* (FS)  $\geq 1$  en 4 mm<sup>2</sup> de tejido glandular se le denomina sialoadenitis linfocítica focal (SLF);

es la característica más distintiva, con sensibilidad y especificidad superiores al 80%<sup>9</sup>. La presencia de FS  $\geq 1$  se considera la manifestación histológica sugerente de SS<sup>10</sup>.

A pesar de los cambios a lo largo del tiempo en los sistemas de clasificación, todavía se usa ampliamente el sistema de calificación de Chisholm Mason (CM), el cual no distingue la SLF de la sialoadenitis crónica no específica, un trastorno inflamatorio que no implica autoinmunidad<sup>11</sup>. Hoy en día el FS no es llevado a cabo de forma rutinaria por todos los patólogos, a pesar de ser la metodología de evaluación recomendada<sup>11,12</sup>. Lo anterior da lugar a discrepancias diagnósticas<sup>13</sup> y, al no ser un método estandarizado, se ha observado que la concordancia entre patólogos es baja, tanto para calcular el número de focos promedio (FS) como para reconocer otras características histopatológicas presentes en la BGSM<sup>14</sup>.

El objetivo de este estudio es comparar los 2 sistemas de puntuación histológica más utilizados en los últimos tiempos para diagnóstico de SLF: Chisholm Mason y Focus Score, e identificar la concordancia a la hora de interpretar la biopsia de glándula salival menor en el diagnóstico de SLF, y de esa forma aportar datos en pro de la estandarización del reporte histopatológico en el contexto del SS.

## Materiales y métodos

### Material histológico de estudio (población)

Se llevó a cabo un estudio de concordancia diagnóstica entre 2 pruebas histológicas para SLF. En la revisión de la literatura solo se encontraron estudios de concordancia que evaluaron el acuerdo intra e interobservador de cada una de las pruebas de forma independiente, y en los cuales se evaluaban los criterios necesarios a la hora de interpretar una biopsia de glándula salival para el diagnóstico de SS<sup>11,15,16</sup>. Por ello, fue necesario realizar un consenso entre clínicos expertos (reumatólogos y patólogos) con el fin de establecer el grado de concordancia que se esperaba alcanzar en la lectura de las láminas entre las pruebas de CM y FS.

De acuerdo con el consenso clínico, el grado de concordancia esperada fue del 80%. Para evaluar el acuerdo entre las pruebas CM y FS mediante el coeficiente Kappa de Cohen, con un error tipo I del 0,05, error tipo II de 0,2, una hipótesis nula de 0,6, hipótesis alterna de 0,8 y una probabilidad de clasificación correcta del 50%, se requirió un tamaño de muestra de 126 láminas. El cálculo de este tamaño de muestra se realizó con el programa TAMAMU 1.1. Asimismo, se recolectaron 164 BGSM por medio de la base de datos del servicio de patología en el periodo de enero de 2017 a junio de 2018. Todas las BGSM durante este periodo fueron realizadas por un reumatólogo de la institución, experto en el procedimiento, siguiendo las indicaciones mundiales al respecto<sup>17</sup>.

Se hizo una primera revisión para verificar los criterios de elegibilidad, por parte de los investigadores evaluadores, mediante microscopía de luz, en la que se consideró el área de superficie con una regla milimétrica y el número de lóbulos (fig. 1). El estudio fue aprobado por el comité de ética en investigación en seres humanos HSJ-FUCS (CEISH) del Hospital San José, que es parte del macroproyecto en síndrome de Sjögren, con registro 0457-2018, y se desarrolló teniendo

en cuenta la Resolución 8430/1993 del Ministerio de Salud y demás normativas vigentes en ética de investigación en seres humanos.

### Criterios de inclusión y exclusión

Los criterios de inclusión fueron: 1) BGSM de paciente con síntomas secos realizadas en el Hospital San José; y 2) bloques de parafina o láminas de histología que estuvieran disponibles en la institución y no hubiesen sido retiradas por el paciente. Y los criterios de exclusión fueron: 1) material en bloque de parafina no apto para estudio (material insuficiente o agotado, alteraciones por archivo o manipulación); y 2) material histológico con estas características: glándula salival con menos de 4 lóbulos separados quirúrgicamente (o 6 lóbulos si estos eran pequeños) y área de superficie menor a 8 mm.

### Procedimiento para la lectura y el registro de datos

Se llevó a cabo lectura independiente por parte de 2 médicos residentes de tercer año de patología, previamente entrenados por un médico patólogo experto en la lectura de BGSM (experiencia durante años de residencia y entrenamiento durante 20 días antes del inicio del estudio). Cada uno de los lectores hizo la clasificación y determinación de presencia de SLF mediante la técnica de FS, y después las mismas láminas fueron codificadas nuevamente con código alfanumérico por una tercera persona, para ser leídas con una distancia en tiempo de una semana mediante la técnica de CM.

Se definió presencia de SLF usando la técnica CM, como aquellas con positividad para grado 3 o 4; a su vez, esta se definió empleando la técnica de lectura de FS como FS  $\geq 1$  ( $\geq 1$  foco por 4 mm<sup>2</sup>). Todas las lecturas se hicieron con el microscopio de luz, utilizando los objetivos desde  $\times 4$  hasta  $\times 40$ . Los desacuerdos entre los 2 lectores fueron resueltos por un tercer observador, médico patólogo experto en BGSM, para llegar a un consenso de diagnóstico definitivo.

Se utilizó un formato de recolección de datos que incluía variables sociodemográficas (sexo y edad), área total por evaluar, atrofia, fibrosis, dilatación ductal y centros germinales evaluados con las escalas de medición de FS por campo microscópico de 4 mm<sup>2</sup> (el número total de focos en el espécimen, dividido por el área de superficie glandular y multiplicado por 4) y CM (presencia de infiltración linfocítica difusa o focos de linfocitos con 5 grados, de 0 a 4).

### Análisis estadístico

Las pruebas estadísticas se llevaron a cabo utilizando el software STATA (versión 13). La descripción de las características sociodemográficas e histopatológicas se realizó por mediana y rango intercuartílico para la variable edad, en tanto que las variables cualitativas se describieron por medio de frecuencias absolutas y relativas. Se utilizó el coeficiente de Kappa para calcular los acuerdos interobservador e intraobservador y el acuerdo global entre ambas pruebas. Los valores de Kappa se interpretaron de la siguiente manera: 0,81-1,00, acuerdo casi perfecto; 0,61-0,80, acuerdo sustancial o alto; 0,41-0,60, acuerdo moderado; 0,21-0,40, acuerdo leve;

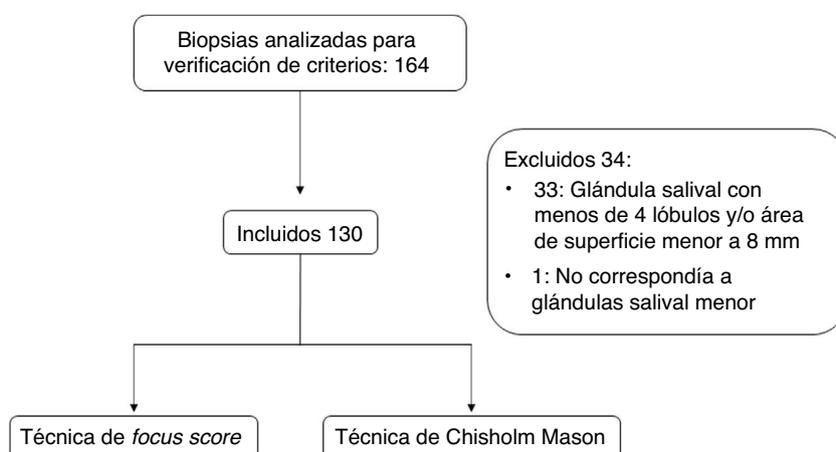


Figura 1 – Flujograma de la población evaluada.

**Tabla 1 – Características sociodemográficas de los pacientes de los cuales provenían las biopsias de glándula salival incluidas en el estudio y prevalencia de sialoadenitis linfocítica focal de acuerdo con evaluación por Chisholm Mason y Focus Score**

	Total 130 pacientes; n (%)	Sialoadenitis linfocítica focal por técnica de Focus Score; n (%)	Sialoadenitis linfocítica focal por técnica de Chisholm Mason; n (%)
Total		11 (8,46)	72 (55,38)
Sexo			
Femenino	122 (93,85)	11 (9,01)	68 (94,44)
Masculino	8 (6,15)	0 (0)	4 (5,56)
Edad: mediana (RIQ)	58 (49-65)		
20-39	14 (10,77)	0 (0)	6 (8,33)
40-59	57 (43,85)	6 (54,55)	32 (44,44)
>60	59 (45,38)	5 (45,45)	34 (47,22)

RIQ: rango intercuartílico.

0,00-0,20, acuerdo leve, según la escala descrita por Landis y Koch<sup>18</sup>.

## Resultados

### Características de la población

De los 164 pacientes a los cuales se les realizó biopsia de glándula salival entre enero de 2017 y junio de 2018 en el Hospital San José, 130 fueron incluidos en el presente estudio (fig. 1 y tabla 1). De estos 122 (93,8%) fueron mujeres y 8 (6,1%) hombres, con una mediana de edad de 58 años (RIC 49-65).

De las 130 biopsias llevadas a cabo mediante la técnica de FS se diagnosticaron 11 (8,46%) con SLF, en tanto que por CM se diagnosticaron 72 (55,38%). De las 122 mujeres a las cuales se les hizo la biopsia con el método CM, 68 fueron diagnosticadas con SLF y 11 con FS, mientras que de los 8 hombres presentes en el estudio 4 tenían SLF con CM y ninguno con FS (tabla 1).

### Acuerdo intraobservador

El observador 1 mostró un acuerdo de 73% (Kappa moderado) y el observador 2 un acuerdo de 58,4% (Kappa leve). La tabla 2

**Tabla 2 – Clasificación de SLF por cada observador según sistema de puntuación**

	SLF positiva	SLF negativa
FS		
Observador 1	23 (17,69)	107 (82,31)
Observador 2	23 (17,69)	107 (82,31)
CM		
Observador 1	58 (44,62)	72 (55,38)
Observador 2	73 (56,15)	57 (43,85)

CM: Chisholm Mason; FS: Focus Score; SLF: sialoadenitis linfocítica focal.

reporta la clasificación de SLF por cada observador según cada sistema de puntuación. En la tabla 3 se muestran los valores de kappa y acuerdo para SLF.

### Acuerdo interobservador

El acuerdo interobservador (tabla 3) fue moderado para FS y alto para CM (Kappa 0,47 y Kappa 0,65, respectivamente). De las 130 biopsias de glándula salival examinadas, 111 fueron concordantes por el método de FS (98 con focus score < 1 y

**Tabla 3 – Acuerdo y valores kappa en la evaluación interobservador, intraobservador y global**

Evaluadores/técnicas de evaluación histológica comparadas	Intraobservador						Interobservador (1 vs. 2)			Global		
	Observador 1			Observador 2			Acuerdo (%)	K	p	Acuerdo (%)	K	p
	Acuerdo (%)	K	p	Acuerdo (%)	K	p						
FS vs. CM	73	0,42	<0,0001	58,4	0,23	0,0001	-	-	-	53	0,13	0,0009
FS vs. FS	-	-	-	-	-	-	84,62	0,47	<0,0001	-	-	-
CM vs. CM	-	-	-	-	-	-	82,31	0,65	<0,0001	-	-	-

CM: Chisholm Mason; FS: Focus Score.

**Tabla 4 – Principales características histopatológicas de la glándula salival menor**

Características histopatológicas	Consenso interobservador (n: 130) %
Atrofia	125 (96,15)
Fibrosis	125 (96,15)
Dilatación ductal	112 (86,15)
Adiposis	45 (34,62)
Centros germinales	6 (4,62)

13 con *focus score*  $\geq 1$ ), y empleando el método de CM hubo concordancia en 107 (53 grado 0-2 y 54 grado 3-4). Cuando hubo discrepancia sobre los hallazgos histológicos adicionales (atrofia, fibrosis, dilatación ductal, adiposis y centros germinales) se hizo un consenso entre los observadores. La frecuencia de dichas características se expresa en la [tabla 4](#), en la que se pone en evidencia que las más frecuentes son atrofia y fibrosis, 96,15% de las biopsias, mientras que el 4,62% de las biopsias presentaba centros germinales.

#### Concordancia entre los 2 sistemas de puntuación histológica de Chisholm Mason y Focus Score

Se presentaron desacuerdos en 19 de las 130 biopsias examinadas por FS y en 23 de las examinadas por CM. De acuerdo con lo planteado se realizó un consenso con el patólogo experto, y de los 19 desacuerdos por FS 18 (94,7%) fueron definidas como SLF negativa y una (5,26%) como SLF positiva. Con CM de los 23 desacuerdos 17 (73,91%) se definieron como SLF positiva y 6 (26,08%) como SLF negativa. El acuerdo global entre CM y FS fue débil, con un Kappa 0,13 (porcentaje de acuerdo 53%) ([tabla 3](#)).

## Discusión

En el presente trabajo se ha realizado por primera vez, en el conocimiento de los investigadores, un análisis de concordancia entre FS y CM para SLF, con uno de los mayores tamaños de muestras reportados en la literatura, y se encontró que el acuerdo entre las 2 pruebas diagnósticas era débil (Kappa 0,13).

La BGSM es de primordial importancia para el diagnóstico, la clasificación y el pronóstico de SS. La clasificación validada, el conjunto de criterios y los hallazgos histopatológicos reflejan la definición de SS como una enfermedad autoinmune que involucra a los epitelios exocrinos<sup>19</sup>. Se trata de una enfermedad prevalente a escala mundial, con un espectro clínico amplio y múltiples facetas que lo hacen difícil de diagnosticar. En consecuencia, suele permanecer no diagnosticado, o ser diagnosticado años después del inicio de los síntomas (se dice que desde el inicio de los síntomas secos el tiempo promedio al diagnóstico tarda 6 años)<sup>20</sup>. De ahí la importancia del reconocimiento temprano, ya que este puede evitar retrasos en el diagnóstico, así como permitir una evaluación diagnóstica adecuada y optimizar la intervención terapéutica<sup>21</sup>. Por ello, se necesita llevar a cabo una evaluación de la BGSM en la cual la SLF es la característica distintiva de aproximación al SS.

Hasta hace pocos años los métodos utilizados para leer y reportar las BGSM ([tabla 5](#)) no se encontraban estandarizados. Daniels et al.<sup>22</sup> fueron los primeros en elaborar un protocolo, y recientemente, en el 2016, Fisher et al.<sup>9</sup> (en representación del Sjögren's Histopathology Workshop Group y del EULAR Sjögren's Syndrome Study Group) proporcionaron un consenso más detallado. Sin embargo, su uso no ha sido ampliamente aceptado, por lo que hay poca información disponible sobre la fiabilidad de estas directrices.

Un informe elaborado en Italia en 2012 señaló que la clasificación de CM todavía se utilizaba en la mayoría de los centros de reumatología experimentados. En dicho estudio Tavoni et al.<sup>15</sup> demostraron un acuerdo interobservador en general aceptable cuando se clasificaron las BGSM mediante la evaluación de CM como consistente o no, con un diagnóstico de SS ( $K=0,70$ ). No obstante, surgieron discrepancias significativas en el momento de una clasificación definitiva en los grados individuales (0-4). Los autores resaltaron las limitaciones que genera la clasificación de CM, por tratarse de un método más antiguo y menos específico y reproducible, en lugar de usar el método recomendado de FS<sup>15</sup>.

Luego de comparar nuestro estudio con la literatura, resaltan las debilidades de la clasificación de CM, definida como una escala categórica que no evalúa la gravedad en pacientes con valores de FS  $> 1$  y sobreestima el conteo de focos<sup>11</sup>,

**Tabla 5 – Principales clasificaciones utilizadas para evaluar las biopsias de glándula salival menor**

Chisholm-Mason		Greenspan		Tarpley	
Grado	Linfocitos 4 mm <sup>2</sup>	Puntaje por foco	Linfocitos 4 mm <sup>2</sup>	Clase	Linfocitos 4 mm <sup>2</sup>
0	Ninguno	1	Un foco	0	Normal
1	Infiltrado escaso	2	Dos focos (hasta 10 focos)	1	Infiltrado mínimo uno o dos agregados celulares
2	Moderado	12	Infiltrado confluyente	2	>2 agregados
3	Un foco			3	Infiltrado difuso y destrucción parcial de los acinos
4	> 2 focos			4	Infiltrado difuso con o sin fibrosis con destrucción completa del acino

Adaptado de Parra-Medina et al.<sup>31</sup>.

ya que todos estos se tienen en cuenta con independencia del estado del parénquima glandular, lo que determina un probable sesgo sistemático. En el presente estudio se puede ver claramente que con la prueba de CM se clasificó el 55,38% de SLF, mientras que con FS solamente el 8,46%, lo cual deja en evidencia la probable sobreestimación de focos con CM y una mayor precisión con FS. Este estudio proporciona la evidencia de que el poder diagnóstico en BGSM podría ser mejorado por métodos estandarizados a escala nacional e internacional.

Otros autores, como Costa et al.<sup>11</sup>, han hecho la comparación entre CM vs. CM y FS vs. FS, con resultados en los cuales, al confrontarlos con los presentes, se reconocen valores Kappa interobservador muy similares cuando se evalúa por CM 0,64 vs. 0,65 y diferentes para FS 0,71 vs. 0,47<sup>11</sup>. Sin embargo, si bien la fiabilidad del FS en el estudio de Costa et al. fue satisfactoria, se observaron disparidades en los métodos de evaluación utilizados por los patólogos locales de los 14 centros de los que provenían las muestras. El protocolo para la determinación del FS no se siguió de manera rutinaria en todos ellos, con el resultado de que a menudo se sobreestimó el FS. En contraste, en nuestro estudio hubo un diagnóstico predominantemente mayor hacia un FS menor a 1, lo que explicaría la diferencia de los valores de Kappa al contrastar los resultados.

Desde el punto de vista estadístico se han descrito algunas paradojas relacionadas con la interpretación del cálculo del coeficiente de Kappa de Cohen<sup>23</sup>. Se ha resaltado que las dificultades surgen porque el Kappa no solamente mide el acuerdo, sino que se ve afectado en diferentes vías por la presencia de sesgo entre los observadores, así como por las distribuciones de los datos en las categorías implementadas, lo cual está relacionado directamente con la prevalencia. Dentro de las paradojas se encuentra que si la proporción de acuerdos esperados es grande, el proceso estadístico de corrección puede convertir valores relativamente altos de acuerdos observados en valores relativamente bajos de Kappa<sup>24,25</sup>.

Esto puede explicarse por los valores de acuerdo esperados por azar en el análisis de concordancia interobservador. En el caso de CM este valor fue del 49%, mientras que en el cálculo a través de FS fue de 70,8% (datos no mostrados). Dichos valores se ven influenciados por la prevalencia de la enfermedad, de modo tal que a medida que la prevalencia se aleja de 0,5, sea más baja o más alta, el porcentaje de acuerdo esperado aumenta, y por consiguiente el coeficiente de Kappa disminuye (tabla 3). Si se tiene en cuenta la baja prevalencia descrita para la SLF, la cual se ha demostrado que es menor al 15% en individuos sanos y ausente en muestras *post mortem*<sup>10,26</sup>, se podría considerar entonces, con los resultados del presente estudio, que la lectura de la biopsia utilizando la técnica de FS refleja mejor la verdadera prevalencia de esta condición. Todo lo anterior se encuentra en concordancia con esta paradoja ya discutida en la literatura científica a escala global<sup>27,28</sup>.

Dado que el acuerdo interobservador mediante FS fue mayor (84,62%) que el observado con CM (82,31%), se llevó un número menor de biopsias a consenso con el experto después de la lectura de los 2 observadores por medio de la primera técnica, en comparación con la segunda (FS 19 vs. CM 23). Mediante dicho consenso el 94,7% (18/19) de las biopsias por FS se definieron como negativas para SLF, mientras que solo el 23% se definieron como negativas mediante CM (6/23). Nue-

vamente, esto evidencia una posible sobreestimación de la frecuencia de SLF cuando se utiliza la técnica del CM.

En otro estudio, que incluyó la revisión independiente de 56 BGSM por 2 patólogos, el acuerdo fue alto para el número de focos (ICC 0,97 [IC 95%: 0,96 a 0,99]) así como para el FS (ICC 0,96 [IC 95%: 0,94 a 0,99])<sup>29</sup>. Se argumentan discrepancias en el hallazgo de SLF cuando se realiza empleando FS en comparación con CM, dado que esta última no distingue entre los patrones de inflamación de la SLF frente al hallazgo de sialoadenitis crónica no específica (no relacionada con SS) basándose en diferentes características que se consideran en el momento de realizar la técnica de FS:

- a) Un foco de gran tamaño puede contarse como varios focos casi confluyentes o como un solo foco.
- b) A su vez puede ser difícil determinar si un agregado de linfocitos es lo suficientemente denso como para constituir un foco.
- c) La medición del área de la superficie puede ser difícil, pero se beneficia considerablemente del uso de una retícula ocular o herramientas basadas en computadoras.
- d) Definir si se está frente a un verdadero foco, teniendo en cuenta las características del tejido adyacente, puede generar conflicto entre los distintos observadores; la atrofia y la fibrosis desempeñan un papel importante en la biopsia y son hallazgos que se podrían considerar de rutina. En nuestro estudio el 96,1% de las biopsias mostró estas características.

Manganelli et al.<sup>28</sup> analizaron los cambios estructurales de las glándulas salivales menores mediante un análisis morfométrico en 20 pacientes con SS primario, en comparación con 20 sujetos normales pareados por edad y sexo. No se encontraron diferencias estadísticamente significativas en el tejido acinar, los conductos intralobulares y el tejido fibroso entre los 2 grupos, lo que sugiere que la atrofia acinar no se relaciona con el proceso inflamatorio, sino que es un fenómeno dependiente de la edad.

Los hallazgos de Leroy et al.<sup>29</sup> también apuntan a que la atrofia acinar no tiene relación con el proceso inflamatorio. De hecho, los cambios atróficos en muestras de BGSM son extremadamente comunes en especímenes de sujetos en necropsia que no tienen historia de enfermedad del tejido conectivo. Así lo muestran Scott et al.<sup>30</sup> en un estudio de glándulas salivales labiales aparentemente no enfermas de 70 necropsias, en el que se evidencia que la atrofia acinar, la dilatación y la hiperplasia ductal fueron frecuentes, tendieron a aumentar con la edad y afectaron a las mujeres antes que a los hombres.

A medida que aumentaba la edad había una reducción progresiva del volumen acinar, proporcional a los aumentos correspondientes en las proporciones de los conductos y el tejido fibroso. Ocasionalmente, los focos linforreticulares estaban presentes en el parénquima no alterado, lo que sugiere, por lo tanto, que dichos focos pueden no desarrollarse exclusivamente en enfermedades autoinmunes<sup>30</sup>. Lo anterior pone en evidencia la importancia de la atrofia y la fibrosis, porque pueden surgir desafíos cuando ambas características están presentes en la misma glándula. De tal manera, si la SLF predomina, entonces el FS debe ser determinado. De todos modos esto no es claro y genera discrepancias.

Es de resaltar que en nuestro estudio el 89% de las BGSM provenía de individuos mayores de 40 años y el 45,38% de mayores de 60 años, de tal forma que la presencia de atrofia, tal como lo ha resaltado el grupo EULAR para la estandarización de la lectura de BGSM, conlleva disparidades en su análisis, ya que los focos adyacentes a áreas de atrofia no se deben tener en cuenta para la lectura mediante FS.

Dentro de las debilidades del presente estudio se resalta la ausencia del uso de herramientas computarizadas para la medición más precisa del área glandular, como las utilizadas por Costa et al.<sup>11</sup>. Así mismo, la ausencia de una evaluación por duplicado de la lectura por cada técnica (CM y FS), con el fin de calcular la concordancia intraobservador de cada una, podría ser otra limitación del estudio; no obstante, resaltan los análisis novedosos de concordancia intraobservador de CM vs. FS.

La evaluación de los sistemas de puntuación por parte de personal en formación (residentes de patología), si bien constituye una debilidad, se manejó con un entrenamiento específico durante sus años de formación. Dicho entrenamiento lo llevaron a cabo estudiantes del último año y se extendió un año más, para aumentar tiempo antes de iniciar las lecturas, así como la resolución de los conflictos con el patólogo experto. Por último, la ausencia de contraste con las demás variables clínicas para confirmar o descartar SS podría considerarse una debilidad. Esto, aun cuando no era el objetivo del estudio, puede ser tenido en cuenta en una futura investigación.

La BGSM es un apartado importante en el diagnóstico del SS; nuestro estudio demuestra las dificultades que conlleva su interpretación correcta y la importancia de empezar a estandarizar su reporte, comenzando por utilizar el método sugerido por el consenso propuesto por Fisher et al.<sup>9</sup>. En condiciones de vida real el FS no siempre se usa; algunos patólogos, así como reumatólogos, siguen prefiriendo el uso de la clasificación mediante CM<sup>15</sup>. Queda puesta en evidencia en nuestros resultados la posible sobreestimación del número de focos con el método de CM. Es necesario un mayor esfuerzo para la consecución de la estandarización, la definición de infiltrados linfocíticos (foco) y el cálculo de FS.

## Financiación

Los autores declaran que este trabajo no recibió financiación alguna.

## Conflictos de intereses

Los autores declaran que no existe conflicto de intereses con respecto al tema y los resultados de este estudio.

## Agradecimientos

Agradecemos al personal de apoyo y a los técnicos en histología del Hospital de San José, de la Fundación Universitaria de Ciencias de la Salud, quienes ayudaron a recopilar los materiales necesarios para los procesos realizados en este estudio.

## BIBLIOGRAFÍA

- Anaya JM, Rojas-Villarraga A, Mantilla RD, Arcos-Burgos M, Sarmiento-Monroy JC. Polyautoimmunity in Sjögren syndrome. *Rheum Dis Clin North Am*. 2016;42:457-72, <http://dx.doi.org/10.1016/j.rdc.2016.03.005>.
- Billings M, Dye BA, Lafolla T, Baer AN, Grisius M, Alevizos I. Significance and implications of patient-reported xerostomia in Sjögren's syndrome: Findings from the National Institutes of Health Cohort. *EBioMedicine*. 2016;12:270-9, <http://dx.doi.org/10.1016/j.ebiom.2016.09.005>.
- Qin B, Wang J, Yang Z, Yang M, Ma N, Huang F, et al. Epidemiology of primary Sjögren's syndrome: A systematic review and meta-analysis. *Ann Rheum Dis*. 2015;74:1983-9, <http://dx.doi.org/10.1136/annrheumdis-2014-205375>.
- Barone F, Colafrancesco S. Sjögren's syndrome: From pathogenesis to novel therapeutic targets. *Clin Exp Rheumatol*. 2016;34 4 Suppl 98:58-62.
- Lackner A, Ficjan A, Stradner MH, Hermann J, Unger J, Stamm T, et al. It's more than dryness and fatigue: The patient perspective on health-related quality of life in primary Sjögren's syndrome-A qualitative study. *PLoS One*. 2017;12:e0172056, <http://dx.doi.org/10.1371/journal.pone.0172056>.
- Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: A consensus and data-driven methodology involving three international patient cohorts. *Ann Rheum Dis*. 2017;76:9-16, <http://dx.doi.org/10.1136/annrheumdis-2016-210571>.
- Riega-Torres JC, Villarreal-Gonzalez AJ, Ceceñas-Falcon L, Salas-Alanis JC. Sjögren's syndrome (SS), a review of the subject and saliva as a diagnostic method. *Gac Med Mex*. 2016;152:371-80.
- Jonsson R, Kroneld U, Bäckman K, Magnusson B, Tarkowski A. Progression of sialadenitis in Sjögren's syndrome. *Br J Rheumatol*. 1993;32:578-81, <http://dx.doi.org/10.1093/rheumatology/32.7.578>.
- Fisher BA, Jonsson R, Daniels T, Bombardieri M, Brown RM, Morgan P, et al. Standardisation of labial salivary gland histopathology in clinical trials in primary Sjögren's syndrome. *Ann Rheum Dis*. 2017;76:1161-8, <http://dx.doi.org/10.1136/annrheumdis-2016-210448>.
- Radfar L, Kleiner DE, Fox PC, Pillemer SR. Prevalence and clinical significance of lymphocytic foci in minor salivary glands of healthy volunteers. *Arthritis Rheum*. 2002;47:520-4, <http://dx.doi.org/10.1002/art.10668>.
- Costa S, Quintin-Roué I, Lesourd A, Jousse-Joulin S, Berthelot JM, Hachulla E, et al. Reliability of histopathological salivary gland biopsy assessment in Sjögren's syndrome: A multicentre cohort study. *Rheumatology (Oxford)*. 2015;54:1056-64, <http://dx.doi.org/10.1093/rheumatology/keu453>.
- Shiboski SC, Shiboski CH, Criswell L, Baer A, Challacombe S, Lanfranchi H, et al. American College of Rheumatology classification criteria for Sjögren's syndrome: A data-driven, expert consensus approach in the Sjögren's International Collaborative Clinical Alliance cohort. *Arthritis Care Res (Hoboken)*. 2012;64:475-87, <http://dx.doi.org/10.1002/acr.21591>.
- Vivino FB, Gala I, Hermann GA. Change in final diagnosis on second evaluation of labial minor salivary gland biopsies. *J Rheumatol*. 2002;29:938-44.
- Stewart CM, Bhattacharyya I, Berg K, Cohen DM, Orlando C, Drew P, et al. Labial salivary gland biopsies in Sjögren's syndrome: Still the gold standard? *Oral Surg Oral Med Oral*

- Pathol Oral Radiol Endod. 2008;106:392-402, <http://dx.doi.org/10.1016/j.tripleo.2008.04.018>.
15. Tavoni AG, Baldini C, Bencivelli W, Cavazzini L, Covelli M, de Vita S, et al. Minor salivary gland biopsy and Sjögren's syndrome: Comparative analysis of biopsies among different Italian rheumatologic centers. *Clin Exp Rheumatol*. 2012;30:929-33.
  16. Carubbi F, Alunno A, Cipriani P, Bartoloni E, Baldini C, Quartuccio L, et al. A retrospective, multicenter study evaluating the prognostic value of minor salivary gland histology in a large cohort of patients with primary Sjögren's syndrome. *Lupus*. 2015;24:315-20, <http://dx.doi.org/10.1177/0961203314554251>.
  17. Colella G, Cannavale R, Vicidomini A, Itró A. Salivary gland biopsy: A comprehensive review of techniques and related complications. *Rheumatology (Oxford)*. 2010;49:2117-21, <http://dx.doi.org/10.1093/rheumatology/keq225>.
  18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-74, <http://dx.doi.org/10.2307/2529310>.
  19. Altman DG, Bland JM. Measurement in medicine: The analysis of method comparison studies. *J R Stat Soc Ser D Statistic*. 1983;32:307-17, <http://dx.doi.org/10.2307/2987937>.
  20. Mariette X, Criswell LA. Primary Sjögren's syndrome. *N Engl J Med*. 2018;378:931-9, <http://dx.doi.org/10.1056/NEJMcp1702514>.
  21. Novljan MP, Rozman B, Jerse M, Rotar Z, Vidmar G, Kveder T, et al. Comparison of the different classification criteria sets for primary Sjögren's syndrome. *Scand J Rheumatol*. 2006;35:463-7, <http://dx.doi.org/10.1080/03009740600759860>.
  22. Daniels TE, Silverman S, Michalski JP, Greenspan JS, Sylvester RA, Talal N. The oral component of Sjögren's syndrome. *Oral Surg Oral Med Oral Pathol*. 1975;39:875-85, [http://dx.doi.org/10.1016/0030-4220\(75\)90108-5](http://dx.doi.org/10.1016/0030-4220(75)90108-5).
  23. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46:423-9, [http://dx.doi.org/10.1016/0895-4356\(93\)90018-v](http://dx.doi.org/10.1016/0895-4356(93)90018-v).
  24. Lantz CA, Nebenzahl E. Behavior and interpretation of the kappa statistic: Resolution of the two paradoxes. *J Clin Epidemiol*. 1996;49:431-4, [http://dx.doi.org/10.1016/0895-4356\(95\)00571-4](http://dx.doi.org/10.1016/0895-4356(95)00571-4).
  25. Vach W. The dependence of Cohen's kappa on the prevalence does not matter. *J Clin Epidemiol*. 2005;58:655-61, <http://dx.doi.org/10.1016/j.jclinepi.2004.02.021>.
  26. Chisholm DM, Waterhouse JP, Mason DK. Lymphocytic sialadenitis in the major and minor glands: A correlation in postmortem subjects. *J Clin Pathol*. 1970;23:690-4, <http://dx.doi.org/10.1136/jcp.23.8.690>.
  27. Daniels TE, Cox D, Shiboski CH, Schiødt M, Wu A, Lanfranchi H, et al. Associations between salivary gland histopathologic diagnoses and phenotypic features of Sjögren's syndrome among 1,726 registry participants. *Arthritis Rheum*. 2011;63:2021-30, <http://dx.doi.org/10.1002/art.30381>.
  28. Manganelli P, Salaffi F, Nervetti A, Raffaini N, Olivetti G. Morphometric study of salivary glands in primary Sjögren's syndrome. *Clin Exp Rheumatol*. 1989;7:273-6.
  29. Leroy JP, Soulier PYC. Follow up study of labial salivary gland lesions in primary Sjögren's syndrome. *Ann Rheum Dis*. 1992;51:777-80, <http://dx.doi.org/10.1136/ard.51.6.777>.
  30. Scott J. Qualitative and quantitative observations on the histology of human labial salivary glands obtained post mortem. *J Biol Buccale*. 1980;8:187-200.
  31. Parra-Medina R, Molina C, Gonzalez MJ. Histopatología de la glándula salival menor. En: Anaya JM, Sarmiento-Monroy JC, García-Carrasco M, editores. *Síndrome de Sjögren segunda edición*. Bogotá: Editorial Universidad del Rosario; 2017. p. 105-13, <http://dx.doi.org/10.12804/tm9789587388411>.