



## Artículo Original

# Concordancia intra e interobservador de la clasificación de Oestern y Tscherne para lesiones de tejidos blandos en fracturas cerradas periarticulares del miembro inferior

Intra- and interobserver agreement on the Oestern and Tscherne classification of soft tissue injury in periarticular lower-limb closed fractures

Carlos Oliver Valderrama-Molina<sup>1,2,3</sup>, Mauricio Estrada-Castrillón<sup>1</sup>, Jorge Andrés Hincapie<sup>2,3</sup>, Luz Helena Lugo-Agudelo<sup>2,3</sup>

<sup>1</sup> Hospital Pablo Tobón Uribe, Medellín, Colombia

<sup>2</sup> Universidad de Antioquia, Escuela de Medicina de la Rehabilitación en el Grupo Salud, Medellín, Colombia

<sup>3</sup> Universidad de Antioquia, Escuela de Medicina, Epidemiología Clínica Grupo Académico, Medellín, Colombia

Valderrama-Molina CO, Estrada-Castrillón M, Hincapie JA, Lugo-Agudelo LH. Intra- and interobserver agreement on the Oestern and Tscherne classification of soft tissue injury in periarticular lower-limb closed fractures. *Colomb Med.* 2014; 45(4): 173-78

© 2014 Universidad del Valle. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative Commons Attribution License, que permite el uso ilimitado, distribución y reproducción en cualquier medio, siempre que el autor original y la fuente se acrediten

### Historia

Recibido: 9 Diciembre 2013

Revisado: 30 Marzo 2014

Aceptado: 20 Noviembre 2014

### Palabras clave

Lesión de tejidos blandos, clasificación, fracturas tibiales, concordancia.

### Keywords

Soft tissue injury, classification, tibial fractures, reliability

### Resumen

**Antecedentes:** La lesión de los tejidos blandos en fracturas periarticulares del miembro inferior determina el momento adecuado para realizar la fijación ósea.

**Objetivo:** El propósito de este estudio fue determinar la concordancia intra e interobservador para la clasificación de Tscherne.

**Métodos:** Estudio descriptivo, prospectivo en pacientes ingresados al Hospital Pablo Tobón Uribe con fracturas de platillos tibiales o pilón tibial. Se realizó una evaluación estandarizada utilizando videos y fotografías al momento del ingreso y 24, 48 y 72 h después. Quince evaluadores que tenían diferentes niveles de entrenamiento produjeron 1200 observaciones. La concordancia entre e interobservador fue evaluada utilizando kappa ponderado para múltiples evaluadores y más de dos categorías.

**Resultados:** Veinte pacientes fueron admitidos con fracturas de los platillos o del pilón tibial. La concordancia intraobservador para los 15 evaluadores fue kappa 0.81 (95% IC 0.79-0.83), y la concordancia interobservador para los 15 evaluadores fue kappa 0.65 (95% IC 0.55-0.73). La concordancia interobservador a las 24 h del ingreso fue kappa 0.67 (95% IC 0.46-0.86).

**Conclusión:** Clasificar la gravedad de la lesión de los tejidos blandos es un paso fundamental en la planeación del manejo quirúrgico de las fracturas periarticulares del miembro inferior, con base en nuestros resultados podemos afirmar razonablemente que la clasificación de Oestern y Tscherne tiene un nivel de concordancia adecuado y podría utilizarse de forma sistemática para estandarizar la evaluación entre los ortopedistas para orientar el tratamiento y también con fines investigativos.

### Abstract

**Background:** The soft tissues injury in periarticular fractures of the lower extremities determines the proper time to perform bone fixation.

**Objective:** The aim of this study was to determine the intra and interobserver agreement in the Tscherne classification.

**Methods:** This is a descriptive, prospective study for patients admitted to the Pablo Tobón Uribe Hospital (PTUH) with tibial plateau or tibial pilon fractures. We performed a standardized evaluation using video photography at the time of admission and 24, 48, and 72 h after admission. Fifteen reviewers who had various levels of training produced a total of 1,200 observations. The intra- and interobserver agreement was assessed using a weighted kappa for multiple raters and more than two categories.

**Results:** Twenty patients were admitted with tibial plateau and tibial pilon fractures. The intraobserver agreement for all 15 raters was kappa 0.81 (95% CI 0.79-0.83), and the interobserver agreement for all 15 raters was kappa 0.65 (95% CI 0.55-0.73). The interobserver agreement at 24 h was kappa 0.67 (95% CI 0.46-0.86).

**Conclusions:** Classifying the severity of soft tissue injury is critical in planning the surgical management of fractures of the lower extremities. Based on our results, we can reasonably argue that the Tscherne classification produced an adequate level of agreement and could be used to standardize and to guide the treatment, and to conduct research studies.

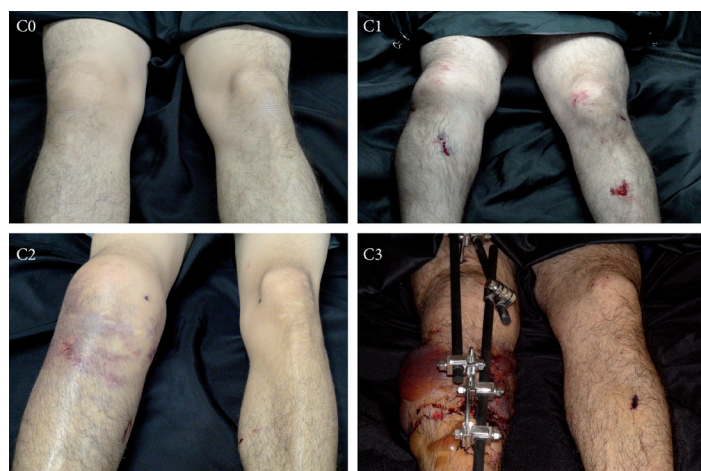
### Autor de correspondencia:

Carlos Oliver Valderrama Molina. Hospital Pablo Tobón Uribe, Medellín, Colombia, South America. Calle 78B Número 69-240 Consultorio 144. Tel: 00 57 444459144. Email: cvalderrama@hptu.org.co

## Introducción

El estado de los tejidos blandos juega un papel preponderante en el tratamiento de las lesiones óseas, principalmente de aquellas que ocurren en áreas con pobre cubrimiento tegumentario como lo es la tibia en sus porciones proximal, diafisaria y distal<sup>1,2</sup>. El tratamiento por etapas ha sido ampliamente favorecido para el manejo de las lesiones de tejidos blandos adyacente a las fracturas del pión o los platillos tibiales. En la primera etapa del tratamiento, un procedimiento de fijación externa temporal permite la resolución del proceso inflamatorio secundario al trauma, en una segunda etapa, se realizan los procedimientos de fijación ósea definitiva con técnicas abiertas o mínimamente invasivas<sup>3,4</sup>. Los resultados publicados con este manejo por etapas permitió la disminución de las complicaciones relacionadas con los tejidos blandos (ej. infección, dehiscencias de herida, necesidad de colgajos para cubrimiento); en el caso de las fracturas de platillos estas complicaciones disminuyeron de un 80% en la década de los ochenta, a cerca del 5% a principios de la década pasada<sup>4</sup> y para las fracturas del pión tibial de un 55.0% a un 3.4% a finales de la década de los noventa<sup>5</sup>.

Por lo anterior la determinación de la gravedad de la lesión inicial de los tejidos blandos, es hoy en día uno de los factores más relevantes para definir el manejo inicial y ayudar al cirujano a tomar decisiones sobre que técnicas de fijación ósea debe utilizar<sup>6</sup>. Sin embargo, es usual que la evaluación de la gravedad de estas lesiones sea subjetiva y varíe entre cirujanos, por esta razón se han propuesto diversas clasificaciones para tratar de estandarizar la medición de la magnitud de la lesión tegumentaria<sup>7</sup>. La clasificación más utilizada en la literatura para describir la lesión de los tejidos blandos en trauma cerrado es la clasificación de Oestern y Tschern<sup>8-10</sup>, publicada a comienzos de la década de los ochenta, esta clasificación determina la gravedad del trauma cerrado de tejidos blandos desde 0 hasta 3, siendo cero una lesión mínima o inexistente y tres la lesión más grave que incluye el síndrome compartimental y las lesiones vasculares (Fig. 1).



**Figura 1.** Clasificación de Oestern y Tschern. C0: Lesión mínima o inexistente de los tejidos blandos con fracturas de trazo simple producto de un trauma indirecto. C1: Encoriación superficial o contusión de la piel asociados a trazos de fracturas simples o mediana complejidad normalmente la lesión es producida por la presión que ejercen los fragmentos óseos sobre los tejidos blandos. C2: Excoriaciones profundas contaminadas y contusiones localizadas en piel y músculo resultante de trauma directo. El síndrome compartimental inminente cabe dentro de este grupo. Los trazos de fractura asociados son generalmente complejos. C3: Contusión extensa de la piel, destrucción del músculo o tejido celular subcutáneo, y las evulsiones cerradas. El síndrome compartimental y las lesiones vasculares están incluidos. Los trazos de fracturas son complejos.

Se ha descrito la relación entre la gravedad definida con este sistema y los desenlaces a largo plazo<sup>1</sup>; sin embargo, no conocemos estudios que hayan evaluado la concordancia inter o intraobservador de esta clasificación, y este punto ha sido considerado una de las razones por lo cual no se ha extendido su aplicación en la práctica clínica<sup>11</sup>. Otro aspecto importante es que la lesión de los tejidos blandos es evolutiva y la fase inflamatoria del trauma suele comenzar su resolución después de las 72 horas del trauma inicial<sup>6</sup>, por lo tanto la concordancia de esta clasificación podría verse influenciada por el momento en el que se aplique, en lo que respecta a la clasificación de Oestern y Tschern no se ha descrito cuál sería el momento ideal para clasificar la lesión de los tejidos blandos.

El propósito de este estudio fue determinar la concordancia inter e intraobservador de la clasificación de Oestern y Tschern en trauma de tejidos blandos asociados a fracturas periarticulares del miembro inferior (Platillos y Pión tibial) y describir como esta concordancia varía de acuerdo al nivel de experiencia de los evaluadores y el momento de la evaluación.

## Materiales y Métodos

### Diseño

Este es un estudio observacional, descriptivo, prospectivo, para evaluar la concordancia intra e interobservador de la clasificación de Oestern y Tschern, en pacientes consecutivos que ingresaron al departamento de urgencias del Hospital Pablo Tobón Uribe entre enero de 2011 y abril de 2012. Para la publicación de esta investigación nos hemos adherido a las recomendaciones de Guidelines for Reporting Reliability and Agreement Studies (GRRAS)<sup>12</sup>.

### Pacientes

Se incluyeron los pacientes entre 16 y 60 años residentes en el área metropolitana de la ciudad de Medellín (Colombia), con accidentes de tránsito y diagnóstico de fracturas cerradas de los platillos tibiales o el pión tibial (Clasificación AO 41A2 a 41C3 y 43A1 a 43C3) y que ingresaran al HPTU en las primeras seis horas después del trauma inicial. Se excluyeron los pacientes que no cumplían con los criterios de inclusión o que no desearan participar en la investigación.

### Evaluación clínica y videofotográfica

Al momento del ingreso a urgencias, se obtuvo el consentimiento informado del paciente y se llenó la ficha de ingreso que incluyó, entre otras, las siguientes variables: edad, sexo, mecanismo de trauma, antecedentes de hipertensión arterial, diabetes, tabaquismo. Se realizó la medición clínica de la circunferencia de la extremidad, en el caso de platillos tibiales a nivel de la tuberosidad tibial anterior, y en el caso del pión tibial a 3 cm de la punta del maléolo medial, comparadas ambas con la extremidad contralateral; además se hizo la medición de las excoriaciones, flictenas y equimosis, y el registro de su ubicación en la extremidad lesionada, se repitieron estas mediciones a las 24, 48 y 72 horas. Se obtuvieron imágenes videofotográficas de la extremidad del paciente al ingreso, a las 24, 48 y 72 horas. La obtención de las imágenes se hizo con una cámara Sony® de alta definición (1080i) modelo HDR-HR520 y para la toma de imágenes se utilizó un fondo negro estándar sin marcas, sobre el cual se posicionó la

extremidad comprometida, solo se obtuvieron imágenes de la extremidad y en ningún caso se incluyeron elementos que permitieran identificar al paciente. La cámara se puso en un trípode que permitió una visión perpendicular de la extremidad a una distancia de 50 centímetros, se obtuvieron videos de la cara anterior de ambas extremidades y de las caras medial y lateral de la extremidad lesionada, además se obtenían fotografías de 12 megapíxeles que incluían un marco de mínimo 50 cm de la cara anterior, medial y lateral de la extremidad lesionada. La persona encargada de realizar estas imágenes fue un médico general que recibió un entrenamiento básico por parte de un experto en fotografía, con el fin de estandarizar la cantidad de luz, el macro y otras características técnicas inherentes a las imágenes.

Este mismo procedimiento se repitió en la unidad de hospitalización del HPTU a las 24, 48 y 72 horas del trauma con el fin de observar la evolución de la lesión de los tejidos blandos.

Los videos y las fotografías se editaron para crear paquetes videofotográficos que tuvieran una duración de 1 minuto, de esta manera cada paciente tuvo 4 evaluaciones diferentes. Estos paquetes fueron almacenados en un computador portátil exclusivo para la investigación y se hizo una copia de respaldo en un disco duro externo de 1 terabyte.

Se hizo una difusión de la investigación y de la importancia de la clasificación Oestern y Tscherne, por medio de afiches entre los ortopedistas y residentes de la institución un mes previo al inicio de la recolección de los casos.

#### Evaluadores

Para evaluar la concordancia se invitaron 5 residentes de tercer y cuarto año del programa de Ortopedia de la Universidad Pontificia Bolivariana, además se invitaron cinco ortopedistas del HPTU con experiencia en el tratamiento de las lesiones de platinos y pilon tibial, quienes hubieran hecho entre 5 y 10 casos de estas fracturas en el último año o que trataran estas lesiones en urgencias, que no fueran los médicos tratantes y no hubieran tenido un contacto directo con los casos clínicos del estudio. Se invitaron también cinco ortopedistas cuya principal área de trabajo fuera el trauma y realizaran más de 20 procedimientos quirúrgicos de este tipo de fracturas al año. Con lo anterior se obtuvo un espectro de evaluadores quienes podrían utilizar la clasificación de Oestern y Tscherne y que incluyeron 5 inexpertos (Residentes), 5 usuarios habituales (Ortopedistas) y 5 expertos (Ortopedistas de Trauma). Se determinó el número de evaluadores por nivel de experiencia siguiendo la recomendación empírica de Audige y colaboradores<sup>13</sup>.

#### Sesiones de evaluación

Se hicieron sesiones de evaluación con 4 pacientes, de tal manera que cada sesión contaba con 16 paquetes videofotográficos, con una duración de 16 minutos por sesión, esto con el fin de evitar el cansancio de los evaluadores. Cuando se tenía un grupo de 4 pacientes con todas las evaluaciones videofotográficas se hacía el montaje de una presentación de powerpoint 2007<sup>®</sup>, esto constituía una sesión, cada uno de los paquetes se organizaban de forma aleatoria, así ningún evaluador revisó la misma secuencia de casos. Cada sesión se hizo en un salón del área académica del hospital con un evaluador por sesión y sin otra persona en la sala, excepto por quien organizaba el paso de imágenes, con esto se garantizó la independencia de las evaluaciones. Las imágenes se presentaron

en un televisor Sony HD de 42", el evaluador se encontraba frente al televisor a una distancia máxima de 1.5 metros, la conexión se realizó con un cable HDMI desde el computador portátil al televisor. No se le entregó información clínica al evaluador acerca del caso a observar, el evaluador tenía un tiempo máximo de 1 minuto para registrar su resultado en la ficha diseñada con este fin. Para evaluar la concordancia intraobservador, se repitió la sesión entre 10 y 14 días, bajo las mismas condiciones antes descritas, con la misma matriz de casos pero con una nueva asignación aleatoria de los paquetes para evitar el sesgo de memoria.

#### Control de sesgos

El sujeto del presente estudio fue el paciente que ingresó a urgencias del HPTU con fractura periarticular del miembro inferior, los sesgos a este nivel fueron controlados con los criterios de inclusión y exclusión. La estandarización de la técnica videofotográfica permitió controlar los sesgos relacionados con el objeto, sin embargo se reconoce que esta evaluación puede diferir de la evaluación clínica en el escenario de urgencias, por lo que se recolectó la evaluación inicial del ortopedista de turno en urgencias que ingresaba al paciente como una variable adicional para comparar la concordancia con las evaluaciones videofotográficas. Para evitar que los evaluadores fueran influenciados por las respuestas de otro evaluador las sesiones se diseñaron para ser individuales. Para evitar el sesgo de memoria en la evaluación de la concordancia intraobservador se dejó un intervalo de 10 a 14 días entre evaluación, con el mismo fin se hizo una asignación aleatoria de la matriz de videos para cada sesión y para cada evaluador.

#### Análisis estadístico

En cada sesión de evaluación se registraron los resultados en un documento diseñado para tal fin. El análisis de las variables cuantitativas se hizo con los estadísticos descriptivos, medias y medianas con desviaciones estándar y rangos intercuartílicos, respectivamente, y proporciones para las variables cualitativas. En el análisis de concordancia intraobservador se utilizó el estadístico Kappa ponderado por pesos cuadráticos y para el interobservador se utilizó el estadístico Kappa ponderado por pesos cuadráticos con más de 2 evaluadores y más de 2 categorías siguiendo el método propuesto por Fleiss<sup>14</sup>. Los resultados se presentan con Kappa por categoría de la clasificación, por momento en la evolución (Ingreso, 24, 48 y 72 horas) y por nivel de entrenamiento, cada uno con sus respectivos intervalos de confianza. Se clasificó el índice de Kappa encontrado utilizando el sistema propuesto por Landis y Koch<sup>15</sup>: 0=Pobre, 0.01-0.2= Leve, 0.21-0.40=Regular, 0.41-0.60=Moderado, 0.61-0.80= Sustancial, 0.81-1= Casi perfecto.

Para el cálculo del tamaño de la muestra se utilizó la fórmula revisada por Cicchetti<sup>16</sup> para la utilización del índice de Kappa ponderado en una escala ordinal de 4 categorías como la de Oestern y Tscherne. En esta fórmula, K representa el número de categorías en una escala nominal, para el caso de la clasificación de Oestern y Tscherne que tiene 4 categorías encontramos: de tal manera que se requería de 32 casos para la evaluación videofotográfica planeada.

El software utilizado para los estadísticos descriptivos fue SPSS v17<sup>®</sup>. Para el cálculo de Kappa intra e interobservador y sus respectivos intervalos de confianza del 95% se utilizó AGREESTAT 2011<sup>17</sup>; para el cálculo de Kappa intraobservador consolidado y sus respectivos intervalos de confianza del 95% se utilizó EPIDAT V4.0<sup>18</sup>.



El estudio tuvo la aprobación del comité de investigaciones y ética del HPTU y del Instituto de Investigaciones Médicas de la Universidad de Antioquia, cada paciente firmó un consentimiento informado para la realización de las imágenes.

### Resultados

Entre febrero de 2011 y febrero de 2012 ingresaron al HPTU 67 pacientes con fracturas de platinos y pilón tibial, 40 de estos pacientes no cumplieron criterios de inclusión (21 pacientes ingresaron después de 6 horas del trauma, 17 pacientes tenían fracturas no desplazadas que fueron dados de alta inmediatamente, un paciente era sordomudo y no tenía acompañante para consentir el ingreso al estudio y un paciente era un politraumatizado en estado crítico); Siete pacientes que cumplían criterios de inclusión no ingresaron al estudio porque en cinco de ellos no se avisó a los autores para realizar las imágenes y en dos pacientes no hubo disponibilidad para realizar los videos (ingreso en horas de la madrugada). Ingresaron al estudio 20 pacientes (14 hombres, edad media de 33.9 años), 18 casos fueron fracturas de platinos tibiales, de acuerdo con la clasificación AO ocho casos 41C y 10 casos 41B3. Los dos casos de fracturas de pilón tibial fueron clasificados como AO 43C1; 16 casos fueron inmovilizados con métodos no invasivos (Férula de yeso 3, Brace 13) y cuatro casos con fijación externa. Todos los pacientes requirieron reducción abierta y fijación interna para el tratamiento de sus lesiones óseas y ninguno presentó complicaciones relacionadas con los tejidos blandos en el seguimiento postquirúrgico (Media del seguimiento 11.7 meses (DS 3.4))

Se obtuvieron 80 paquetes videofotográficos (4 por cada paciente). Se realizaron 150 sesiones de evaluación, 75 interobservador y 75 intraobservador, con lo cual se registraron 1,200 mediciones de la clasificación, 300 mediciones por cada momento evaluado (Ingreso, 24, 48 y 72 horas). Cada grupo de evaluadores (Inexpertos, usuarios habituales, expertos) hicieron 50 sesiones de evaluación, 25 intraobservador y 25 interobservador, con lo cual se obtuvieron 400 evaluaciones por grupo. La concordancia entre la evaluación presencial en urgencias y las evaluaciones videofotográficas fue kappa 0.63 (IC 95% 0.55–0.70).

La concordancia intraobservador global para los 15 evaluadores fue kappa 0.81 (IC 95% 0.79- 0.83), lo cual corresponde a una concordancia intraobservador “casi perfecta” según el sistema propuesto por Landis y Koch para la clasificación de kappa (Tabla 1).

La concordancia interobservador global para los 15 observadores fue kappa 0.65 (IC 95% 0.55-0.73), lo cual corresponde a una

concordancia interobservador sustancial; la concordancia interobservador por grupo de evaluadores y por momento de la evaluación pueden observarse en la Tabla 2.

### Discusión

La gravedad de las lesiones de tejidos blandos en pacientes con fracturas periarticulares del miembro inferior es considerada el principal factor para definir el mejor momento del tratamiento quirúrgico. Se han descrito varios sistemas de clasificación para determinar esta gravedad, siendo la clasificación de Oestern y Tschern e el sistema más sencillo y utilizado en la literatura ortopédica, se ha descrito su asociación con las complicaciones postquirúrgicas<sup>19,20</sup>. Sin embargo, no conocemos estudios previos que hayan evaluado la concordancia intra e interobservador de esta clasificación. En este estudio encontramos una concordancia intraobservador “casi perfecta” y una interobservador “sustancial” utilizando el sistema de Landis y Koch para la clasificación de la concordancia cuando se utiliza el estadístico kappa.

La concordancia intraobservador fue excelente en este estudio. El grupo de usuarios comunes tuvo el nivel de concordancia más bajo (kappa 0.67 IC 95% 0.61-0.72), sin embargo aún para ellos esta concordancia fue “sustancial”. Este grupo incluía un cirujano de miembro superior, un cirujano de tumores, y 3 ortopedistas con menos de 5 años de experiencia. Para el grupo de expertos la concordancia intraobservador fue la más alta (kappa 0.86 IC 95% 0.83–0.89), este grupo estaba constituido por 2 cirujanos de rodilla con más de 10 años de experiencia, 1 cirujano de pie y tobillo con más de 5 años de experiencia y 2 cirujanos de trauma con más de 5 años de experiencia. El grupo de los residentes mostró una concordancia intraobservador más allá de lo esperado, superior a la de los ortopedistas y muy cercana a la de los cirujanos de trauma. Las diferencias entre las concordancias puede ser explicada por el nivel de familiaridad con la clasificación y la necesidad de usarla. Para el grupo de cirujanos de trauma ortopédico es fundamental determinar la gravedad de los tejidos blandos ya que esto define el tipo de tratamiento. Para el grupo de ortopedistas pudiera no ser tan relevante y su interés por usar esta, o cualquier, clasificación puede ser bajo. El grupo de residentes escogido ha estado expuesto a la clasificación durante toda su formación y se ha entrenado específicamente en su utilización, esto podría explicar los resultados obtenidos en este grupo.

El comportamiento de la concordancia intra e interobservador durante los cuatro momentos de las evaluaciones, muestra que al ingreso (<6 horas del trauma) fue casi siempre inferior al compararla con los otros tres momentos. En las evaluaciones

**Tabla 1.** Concordancia intraobservador (Kappa (IC 95%)).

	Todos los evaluadores*	Inexpertos (Residentes)*	Usuarios regulares (Ortopedistas)*	Expertos (Ortopedistas)*
Global	0.81 (0.79 - 0.83)	0.79 (0.75 - 0.83)	0.67 (0.61 - 0.72)	0.86 (0.83 - 0.89)
Ingreso	0.70 (0.64 - 0.77)	0.70 (0.58 - 0.81)	0.65 (0.53 - 0.77)	0.74 (0.64 - 0.84)
24 h	0.85 (0.81 - 0.88)	0.81 (0.73 - 0.88)	0.67 (0.57 - 0.78)	0.91 (0.86 - 0.96)
48 h	0.86 (0.82 - 0.89)	0.82 (0.75 - 0.89)	0.68 (0.58 - 0.79)	0.91 (0.86 - 0.97)
72 h	0.89 (0.85 - 0.92)	0.91 (0.86 - 0.96)	0.67 (0.57 - 0.78)	0.90 (0.86 - 0.96)

\* Kappa fue clasificado como 0= Pobre, 0.01 – 0.2= Leve, 0.21- 0.40=Regular, 0.41 - 0.60=Moderado, 0.61 a 0.80=Sustancial, 0.81 – 1 Casi perfecto

**Tabla 2.** Concordancia Interobservador (Kappa (IC 95%))

	Todos los evaluadores*	Inexpertos (Residentes)*	Usuarios Regulares (Ortopedistas)*	Expertos (Ortopedistas)*
Global	0.65 (0.55 - 0.73)	0.69 (0.59 - 0.80)	0.57 (0.48 - 0.66)	0.72 (0.61 - 0.81)
Ingreso	0.45 (0.25 - 0.65)	0.48 (0.15 - 0.8)	0.46 (0.19 - 0.74)	0.61 (0.40 - 0.82)
24 h	0.67 (0.46 - 0.86)	0.69 (0.43 - 0.94)	0.58 (0.37 - 0.79)	0.73 (0.51 - 0.95)
48 h	0.67 (0.51 - 0.83)	0.73 (0.55 - 0.91)	0.55 (0.39 - 0.72)	0.73 (0.55 - 0.90)
72 h	0.68 (0.51 - 0.85)	0.75 (0.60 - 0.92)	0.61 (0.44 - 0.77)	0.71 (0.49 - 0.93)

\* Kappa fue clasificado como 0= Pobre, 0.01 - 0.2= Leve, 0.21- 0.40=Regular, 0.41- 0.60=Moderado, 0.61 a 0.80=Sustancial, 0.81 - 1 Casi perfecto.

después del ingreso la concordancia intraobservador se mantuvo en el nivel “casi perfecto” de Landis y Koch, y la concordancia interobservador se mantuvo en el nivel “sustancial”. Teniendo en cuenta que la lesión de tejidos blandos es progresiva, pareciera que determinar la gravedad de esta lesión al ingreso no es fácil. Solo la aparición de signos más definidos después de las 24 horas del trauma tales como equimosis, flictenas, aumento del edema y otros, permitiría que la clasificación fuera más precisa y concordante.

Este es el primer estudio que evalúa la concordancia de la clasificación de Oestern y Tscherné con unos resultados satisfactorios que difiere de otros estudios en los que se evaluó la concordancia de otros sistemas de clasificación. Un estudio sobre la clasificación de Gustilo y Anderson para fracturas abiertas obtuvo un porcentaje de acuerdo de apenas 60%<sup>21</sup>, y otros estudios de concordancia de sistemas de clasificación de fracturas por radiología, como la clasificación de Young Burgess con kappa 0.72 o la de Tile con kappa 0.40<sup>22</sup>. La mayoría de los estudios que evalúan la concordancia interobservador tienen problemas metodológicos comunes<sup>13</sup>. Para solucionar este inconveniente recientemente se ha publicado la guía GRRAS<sup>12</sup> para la publicación de los resultados de estudios de concordancia, nosotros nos adherimos a este guía. Consideramos que logramos controlar la mayoría de dificultades metodológicas descritas en el trabajo de Kottner *et al*<sup>12</sup>. Se controló la independencia de las evaluaciones porque se hicieron de forma individual, evitamos el sesgo de memoria al organizar las presentaciones con secuencias aleatorias que garantizaron que nunca un evaluador viera las mismas imágenes en el mismo orden, además dejamos un periodo mínimo de 10 días para las evaluaciones intraobservador. Utilizamos el estadístico de kappa ponderado por pesos cuadráticos, calculado por el método de Fleiss, para una clasificación ordinal como la de Oestern y Tscherné.

La utilización de videos y fotografías como elemento válido para la evaluación de concordancias ha sido ampliamente descrito en la literatura dermatológica, con resultados “casi perfectos” al comparar las evaluaciones en vivo con las evaluaciones videofotográficas<sup>23,24</sup>. Sin embargo, existen pocos estudios en la literatura ortopédica que utilicen la evaluación videofotográfica para determinar concordancia entre evaluadores en una clasificación específica. Uno de estos estudios evaluó el porcentaje de acuerdo en la clasificación de Gustilo Anderson para fracturas abiertas<sup>21</sup>. En nuestro estudio la concordancia entre la evaluación en vivo y las evaluaciones videofotográficas fue kappa 0.63, similar a la concordancia interobservador encontrada entre los 15 evaluadores.

Reconocemos algunas limitaciones en este estudio. Primero, el cálculo de tamaño de muestra se hizo para el número de casos necesarios, sin embargo no logramos el número de 32 casos y esto impactó en la precisión de los resultados, principalmente en las evaluaciones en cada momento, sin embargo al reunir las 80 evaluaciones en un ítem “global”, el número de casos fue suficiente y obtuvimos unos resultados precisos e informativos. Segundo, la diferencia entre la evaluación clínica y la evaluación con videos puede implicar un sesgo de información. Sin embargo, las características del diseño de nuestro estudio garantizó que todos los evaluadores estuvieran en la misma condición cuando hacían la valoración videofotográfica; adicionalmente, dadas las limitaciones de conducir un estudio como este en el escenario real, parece que la estrategia que nosotros utilizamos es la mejor opción para determinar la concordancia para este tipo de clasificación. Finalmente, como cualquier estudio de concordancia no podemos generalizar estos resultados ya que están influenciados por factores culturales, sociales, educativos que no se puede garantizar que sean replicados en otros entornos.

## Conclusión

Clasificar la gravedad de la lesión de los tejidos blandos es un paso fundamental en la planeación del manejo quirúrgico de las fracturas periarticulares del miembro inferior. Basados en nuestros resultados podemos afirmar razonablemente que la clasificación de Oestern y Tscherné tiene un nivel de concordancia adecuado y puede utilizarse de forma sistemática para estandarizar la evaluación entre los ortopedistas que tratan habitualmente este tipo de lesiones y también con fines investigativos. La asociación entre la gravedad del trauma de tejidos blandos según la clasificación de Oestern y Tscherné y las complicaciones postquirúrgicas debe probarse en un estudio prospectivo multicéntrico con un tamaño de muestra adecuado que permita controlar otros factores asociados a estas complicaciones.

## Financiación:

Este estudio fue financiado con recursos del Comité para el Desarrollo de la Investigación (CODI) de la Universidad de Antioquia del año 2010 y de la estrategia de sostenibilidad para el Grupo de Rehabilitación en Salud 2013 - 2014 de la Universidad de Antioquia.

## Conflicto de interés:

Ninguno de los autores tiene conflicto de interés.

## References

1. Dirschl DCL. Classification of fractures. In: Rockwood CA, Green DP, Bucholz RW, eds. *Rockwood and Green's fractures in adults*. 7th ed. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2010. pp. 39–52.
2. Volgas D. Preoperative assessment and classification of soft-tissue injuries. In: Volgas DA, Harder Y, eds. *Manual of soft tissue management in orthopaedic trauma*. Stuttgart: Thieme Verlag; 2011. pp. 54–77.
3. Sirkin M, Sanders R, DiPasquale T, Herscovici DJr A staged protocol for soft tissue management in the treatment of complex pilon fractures. *J Orthop Trauma*. 2004; 18(8) Suppl: S32–S38.
4. Egol KA, Tejwani NC, Capla EL, Wolinsky PL, Koval KJ. Staged management of high-energy proximal tibia fractures (OTA types 41): the results of a prospective, standardized protocol. *J Orthop Trauma*. 2005; 19(7): 448–55.
5. Thordarson DB. Complications after treatment of tibial pilon fractures: prevention and management strategies. *J Am Acad Orthop Surg*. 2000; 8(4): 253–65.
6. Tull F, Borrelli J Jr Soft-tissue injury associated with closed fractures: evaluation and management. *J Am Acad Orthop Surg*. 2003; 11(6): 431–8.
7. Südkamp NP. Soft-tissue injury: In: Rüedi PT, Buckley RE, Moran GC. *AO Principles of Fracture Management*. 2 ed. Stuttgart: Thieme Verlag; 2007. pp. 86–113.
8. Oestern HJ, Tscherner H. Physiopathology and classification of soft tissue lesion. *Hefte Unfallheilkd*. 1983; 162: 1–10.
9. Oestern HJ, Tscherner H. Pathophysiology and classification of soft tissue damage in fractures. *Orthopade*. 1983; 12(1): 2–8.
10. Tscherner H, Oestern HJ. A new classification of soft-tissue damage in open and closed fractures (author's transl) *Unfallheilkunde*. 1982; 85(3): 111–5.
11. Sirkin SM, Liporace M. Chapter 14: Fractures with soft tissue injuries. In: Browner BD, Levine MA, Jupiter JB, Trafton GP, Kretter C. *Skeletal Trauma: Basic Science, Management and Reconstruction*. 4 ed. Philadelphia: Saunders Elsevier; 2009. pp. 367–396.
12. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, *et al*. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011; 64(1): 96–106.
13. Audige L, Bhandari M, Kellam J. How reliable are reliability studies of fracture classifications? A systematic review of their methodologies. *Acta Orthop Scand*. 2004; 75(2): 184–94.
14. Fleiss JL, Levin BA, Paik MC. *Statistical methods for rates and proportions*. 3. Hoboken, NJ: J Wiley; 2003.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1): 159–74.
16. Cicchetti DV. Testing the normal approximation and minimal sample size requirements of weighted Kappa when the number of categories is large. *Appl Psychol Measur*. 1981; 5(1): 101–4.
17. Gwet KL. *AgreeStat* 2011. Accessed: 1 November 2013; Available from: <http://advancedanalyticsllc.com/irrhbk/agreestat.html>.
18. Epidat. V4.0 2012. Accessed: 1 November 2013; Available from: [http://www.sergas.es/MostrarContidos\\_N3\\_T01.aspx?IdPaxina=62713](http://www.sergas.es/MostrarContidos_N3_T01.aspx?IdPaxina=62713).
19. Hoiness P, Engebretsen L, Stromsoe K. Soft tissue problems in ankle fractures treated surgically. A prospective study of 154 consecutive closed ankle fractures. *Injury*. 2003; 34(12): 928–31.
20. Gaston P, Will E, Elton RA, McQueen MM, Court-Brown CM. Fractures of the tibia. Can their outcome be predicted? *J Bone Joint Surg Br*. 1999; 81(1): 71–6.
21. Brumback RJ, Jones AL. Interobserver agreement in the classification of open fractures of the tibia. The results of a survey of two hundred and forty-five orthopaedic surgeons. *J Bone Joint Surg Am*. 1994; 76(8): 1162–6.
22. Koo H, Leveridge M, Thompson C, Zdero R, Bhandari M, Kreder HJ, *et al*. Interobserver reliability of the young-burgess and tile classification systems for fractures of the pelvic ring. *J Orthop Trauma*. 2008; 22(6): 379–84.
23. Rimner T, Blozik E, Fischer Casagrande B, Von Overbeck J. Digital skin images submitted by patients: an evaluation of feasibility in store-and-forward teledermatology. *Eur J Dermatol*. 2010; 20(5): 606–10.
24. Ribas J, Cunha Mda G, Schettini AP, Ribas CB. Agreement between dermatological diagnoses made by live examination compared to analysis of digital images. *An Bras Dermatol*. 2010; 85(4): 441–7.