



ORIGINAL

# VALIDADE E FIABILIDADE DO TESTE STRESS MANUAL E GRAVITACIONAL NAS FRATURAS MALEOLARES EQUIVALENTES

*Isabel Rosa, Joaquim Rodeia, Pedro Xavier Fernandes, Raquel Teixeira, Hugo Ribeiro,  
Luís Sobral*

*Serviço de Ortopedia e Traumatologia, Centro Hospitalar de Lisboa Ocidental, Hospital de S. Francisco  
Xavier*

*Nova Medical School – Faculdade de Ciências Médicas de Lisboa*

## **Isabel Rosa**

Assistente Graduado de Ortopedia e Traumatologia e Assistente Convidado de Ortopedia – Especialidades  
Médicas e Cirúrgicas I; Aluna de doutoramento

## **Joaquim Rodeia, Pedro Xavier Fernandes, Hugo Ribeiro**

Interno de Formação Específica de Ortopedia

## **Raquel Teixeira**

Interna de Formação Específica de Ortopedia

## **Luís Sobral**

Assistente de Ortopedia e Traumatologia e Assistente voluntário de Ortopedia – Especialidades Médicas e  
Cirúrgicas I

**Submetido em** 27 dezembro 2018

**Revisto em** 22 maio 2019

**Aceite em** 22 maio de 2019

**Tipo de Estudo:** Estudo de Diagnóstico

**Nível de Evidência:** II

**Declaração de conflito de interesses:** Nada a declarar.

## **Correspondência**

Isabel Rosa

Serviço de Ortopedia e Traumatologia

Centro Hospitalar de Lisboa Ocidental, Hospital de S.  
Francisco Xavier

Estrada do Forte do Alto do Duque

1449-005 Lisboa

Telefone: +351 965 758 709

[isabelpiresrosa@gmail.com](mailto:isabelpiresrosa@gmail.com)

## RESUMO

Objetivo: Terá o teste *stress* manual e gravitacional qualidade para separar populações de indivíduos com fratura maleolar com ou sem lesão do ligamento deltoide?

O objetivo primário consiste em determinar a validade e fiabilidade do teste *stress* manual, comparando-o com a radiografia *stress* gravitacional. Secundariamente avaliar se o espaço túbio talar medial  $\geq 4$  ou  $\geq 5$  mm tem valor preditivo diferente para lesão do ligamento deltoide.

Material e Métodos: Efetuamos um estudo coorte prospetivo. A amostra dividiu-se em grupo I (teste *stress* manual) e grupo II (radiografia gravitacional) com fratura maleolar associada ou não a lesão do ligamento deltoide. Os testes foram exclusivamente realizados pelo autor, embora a sua interpretação tenha sido concretizada por avaliadores independentes, cegos para o resultado entre observadores.

Resultados: Na amostra de 65 doentes, obtivemos no grupo I uma sensibilidade de 97%, especificidade e valor preditivo positivo de 100%. Enquanto no grupo II a sensibilidade foi de 100%, o valor preditivo positivo 97% e negativo 100%.

Verificámos que o espaço túbio talar medial  $\geq 5$  vs  $\geq 4$ mm tem uma sensibilidade 100% vs 91%, especificidade 100% vs 100%, valor preditivo positivo 100% vs 100% e negativo 100% vs 60%.

Em ambos os testes observou-se uma excelente concordância intraobservador e interobservador. Conclusões. Um espaço túbio talar medial  $\geq 5$  mm tem valor preditivo superior na avaliação da integridade e competência do ligamento deltoide. O teste de *stress* manual e gravitacional são uma ferramenta eficaz, tornando-se fundamental na decisão terapêutica na fratura maleolar isolada com suspeita de lesão medial.

**Palavras chave:** *Fratura maleolar; lesão do ligamento deltoide; radiografia gravitacional; teste stress manual; espaço túbio talar medial*

## ABSTRACT

**Background:** Can the manual and gravity stress test quality to separate populations with a malleolar fracture with or without a deltoid ligament injury?

The author intends to assess validity and accuracy of the manual external rotation stress test comparing it with gravity stress radiography. Secondly evaluate whether the medial clear space  $\geq 4$  or  $\geq 5$  mm has different predictive value for lesion of the deltoid ligament.

**Methods:** A consecutive series of patients with isolated lateral malleolar fractures were prospectively enrolled. The sample was divided into group I (manual stress) and group II (gravity stress) with malleolar fracture with or without deltoid ligament injury. The medial clear space on the ankle's mortise radiography was measured by blinded observers. Inter and intraobserver reliability was assessed.

**Results:** In the sample of 65 patients, we obtained in group I a sensitivity of 97%, specificity and positive predictive value of 100%. While in the group II the sensitivity was 100%, the positive predictive value 97% and negative 100%.

We found that the medial clear space  $\geq 5$  vs  $\geq 4$  mm had a sensitivity of 100% vs 91%, specificity 100% vs 100%, positive predictive value 100% vs 100% and negative 100% vs 60%.

In both tests an excellent intraobserver and interobserver reliability was observed.

**Conclusion:** A medial clear space  $\geq 5$  mm has a higher predictive value in assessing the integrity of the deltoid ligament. The manual and the gravity stress test are an effective tool in assessing the deltoid competence in isolated lateral malleolus fractures

**Key words:** *Ankle fracture; deltoid ligament injury; gravity stress radiography; manual stress radiography; medial clear space*

## INTRODUÇÃO

As fraturas maleolares têm aumentado nas últimas décadas, os estudos epidemiológicos de Court-Brown et al, calcularam uma incidência de 125 / 100.000 / ano<sup>1</sup>. Elas são mais comuns nos doentes do género feminino a partir dos 50 anos, vítimas de acidentes de baixa energia, ou em jovens do género masculino no seguimento de acidentes desportivos ou viação.

A congruência das superfícies articulares durante a carga, os complexos ligamentares estáticos, e as unidades músculo tendinosas dinâmicas, são os principais contribuintes para a estabilidade da articulação do tornozelo. O ligamento deltoide é responsável pela estabilização medial, limitando a translação anterior, posterior e lateral do astrágalo na articulação talocrural. O seu feixe profundo é o principal bloqueador da rotação externa do astrágalo, sendo o contribuinte mais importante para a estabilidade do tornozelo<sup>2,3,4</sup>.

Identificar uma lesão do ligamento deltoide numa fratura maleolar lateral isolada e alinhada, continua a ser fundamental para diferenciar o seu grau de estabilidade. A fratura maleolar associada a rotura ligamentar, deve ser considerada instável e tem indicação cirúrgica, de forma a manter o astrágalo estável na pinça maleolar<sup>5,6,7,8</sup>.

Quando se utiliza um teste para fazer a distinção entre indivíduos com resultados normais (sem lesão) ou anormais (com lesão), é importante compreender como é que as características se encontram distribuídas entre as populações. Há também que determinar o ponto de corte acima do qual o teste é considerado positivo e abaixo do qual é considerado negativo<sup>9</sup>.

A medição do espaço túbio talar medial é comumente usada para avaliar a competência do ligamento deltoide e a estabilidade nas fraturas do tornozelo. Michelson em 2001 no seu estudo em 8 cadáveres, demonstrou que a presença na radiografia de *stress* gravitacional de um desvio talar > 2 mm ou um ângulo talar > 10°, era indicativo de instabilidade por rotura completa do deltoide<sup>5</sup>. Estudos de McConnel, Gill, Schock e DeAngelis determinam que um espaço túbio talar medial  $\geq 4$  mm, ou  $\geq 1$  mm que o espaço túbio talar superior, em **radiografia** a 10° de rotação interna, constitui fator preditivo de lesão

dos feixes profundos deste ligamento<sup>6,7,8,10</sup>.

Já Park em 2006, no seu estudo em 6 cadáveres alerta que espaço túbio talar medial  $\geq 5$  mm tem maior valor preditivo na determinação do grau de instabilidade<sup>11</sup>.

Uma vez que as decisões clínicas são frequentemente baseadas em diferenças de 1 a 2 milímetros, é crucial perceber qual a concordância na medição. Trabalhos em cadáver de Metitiri e Ghorbanhoseini alertam para o elevado grau de erro e falta de precisão nas medições<sup>12</sup>.

É assim fundamental determinar a fiabilidade ou reprodutibilidade do teste escolhido. Um facto é que se os resultados obtidos não se puderem reproduzir, o valor e a utilidade desse teste será mínimo, independentemente da sua sensibilidade ou especificidade<sup>9</sup>.

O teste *stress* manual tem sido usado classicamente, mas como é difícil padronizar e nem sempre é tolerado pelo utente, a sua praticidade tem sido questionada. Também alguns autores têm alertado que a **radiografia** de *stress* gravitacional sobrestima os resultados. Por conseguinte continua a não haver consenso quanto ao melhor instrumento a utilizar como método de avaliação da integridade e competência do ligamento deltoide<sup>13,14,15,16</sup>.

Poderão estes testes, ter qualidade para separar populações de indivíduos com fratura maleolar com ou sem lesão do ligamento deltoide, para que possam ser tomadas decisões racionais acerca da sua utilização e interpretação?

Com o nosso estudo pretendemos avaliar se o teste de *stress* manual ou a **radiografia** de *stress* gravitacional, poderá revelar-se uma ferramenta eficaz na avaliação da competência do ligamento deltoide, tornando-se assim fator fundamental na decisão terapêutica na fratura do maléolo lateral isolada.

Assim o objetivo primário consiste em determinar a validade e a fiabilidade do teste *stress* manual, quando comparada com outro método de imagem a radiografia *stress* gravitacional. E secundariamente determinar se o espaço túbio talar medial  $\geq 4$  mm ou  $\geq 5$  mm tem um valor preditivo diferente para lesão do ligamento deltoide.

## MATERIAL E MÉTODOS

No período compreendido entre 10 Fevereiro de 2016 e 30 Novembro de 2017, obtivemos uma população de 105 doentes com fratura maleolar isolada, admitidos no serviço de urgência.

Foram incluídos no estudo, os doentes com maturidade esquelética, as fraturas do maleolo lateral com suspeita de lesão do ligamento deltoide ou com avulsão do maleolo medial até 5mm, fratura maleolo lateral com encurtamento > 2mm ou desvio rotacional. Foram excluídas, trauma prévio a nível maleolar, fratura isolada do maleolo medial, posterior ou sem lesão do ligamento deltoide, as operadas por outros colegas ou com mais de 15 dias após o episódio traumático agudo e ainda os transferidos para outra instituição hospitalar. Obtivemos assim, uma amostra de 65 utentes.

Efetuamos um estudo observacional longitudinal, coorte prospetivo comparativo.

Utilizámos a **radiografia stress** gravitacional como “padrão de ouro” para determinar o estado do ligamento deltoide de cada um dos indivíduos da nossa população.

A amostra dividiu-se em dois grupos: o grupo I (teste *stress* manual) com fratura maleolar associada a lesão do ligamento deltoide (doentes expostos) ou sem lesão do ligamento deltoide (doentes não expostos) e o grupo II (**radiografia stress** gravitacional) com fratura maleolar associada a lesão do ligamento deltoide (doentes expostos) ou sem lesão do ligamento deltoide (doentes não expostos).

Neste estudo a dimensão do espaço tíbio talar medial foi centrado nos <4 mm, ou ≥4 mm e ≥5 mm, permitindo assim a separação de indivíduos sem ou com lesão do ligamento deltoide.

**A radiografia stress gravitacional foi efetuada com o paciente em decúbito lateral sobre o lado lesionado, ficando o tornozelo e o pé pendente pelo efeito da gravidade, fora da marquês de Raio X.**

O teste *stress* manual foi realizado no bloco operatório sob anestesia, com o pé em dorsiflexão e 15° de rotação interna, aplicando força de rotação externa e visualizado em intensificador de imagem, marca Philips.

As imagens dos dois testes, foram gravadas

digitalmente através do sistema IMPAX®. A fim de se evitar o enviesamento da análise, os testes foram exclusivamente realizados pelo autor da investigação, embora a sua interpretação tenha sido concretizada por avaliadores independentes, cegos para o resultado entre observadores. Assim, a população em estudo foi dividida entre quatro observadores, que receberam instruções para medir o espaço tíbio talar medial. Foram utilizados critérios de leitura semelhantes para os dois testes, com medição em mm do espaço tíbio talar medial entre o bordo medial do astrágalo e o bordo lateral do maleolo medial, desenhando uma linha perpendicular a 0,5 cm do rebordo superior da superfície articular medial (**Figura 1**).



Figura 1 – Medição em mm do espaço tíbio talar medial a) na radiografia gravitacional b) na radiografia stress manual.

Para calcular a concordância percentual observada e a concordância percentual esperada apenas por acaso da medição do espaço tíbio talar medial, adicionamos os números em que as leituras de ambos os observadores estiveram de acordo, dividimos essa soma pelo número total de medições lidas e multiplicámos o resultado por 100. Calculou-se ainda a estatística Kappa<sup>9,17</sup>. Obteve-se o valor Kappa pela equação: (Concordância percentual observada) - (Concordância percentual esperada apenas por acaso) / 100% - (Concordância percentual esperada apenas por acaso).

Efetuamos ainda a análise estatística descritiva dos dados, com recurso ao *Software Statistical Package for Social Sciences (SPSS)* versão 23, apresentando-se os dados categóricos como percentagens, enquanto as variáveis contínuas foram caracterizadas com recurso a medidas de tendência central e de dispersão. Foi utilizado o teste Chi-Square para avaliar a associação entre duas variáveis, o teste Kolmogorov- Smirnov para verificar a normalidade da população em estudo e o teste de Mann-

Whitney para avaliar a significância. Como forma de avaliação da relação linear entre os dois testes, utilizou-se o coeficiente de correlação Pearson. O nível de significância considerado foi de  $p \leq 0,05$ , de acordo com o habitualmente preconizado a nível internacional.

Todos os participantes assinaram o Consentimento Informado, respeitou-se a Lei 67/98 de 26 de Outubro da Proteção de Dados, revogada pela Deliberação 1704/2015, assim como autorização pela comissão de Ética do Hospital.

## RESULTADOS

Na população de 105 doentes com fratura maléolo lateral isolada, foram excluídos 40 utentes: 26 (24,8%) com fratura do maléolo lateral sem lesão do ligamento deltoide, 2 operados por outros colegas, 2 com mais de 15 dias após o episódio traumático agudo e 10 transferidos para outra instituição hospitalar.

Dos 65 doentes elegíveis, 25 (38,5%) eram do género masculino e 40 (61,5%) do feminino. A média de idade foi de 50,7 (Max 85 e Min 14) com desvio padrão de 17,5. Em relação à lateralidade, 31 (47,7%) eram à direita e 34 (52,3%) à esquerda. 47 (72,3%) foram vítimas de queda, 15 (23,1%) de acidente desportivo e 3 (4,6%) de acidente de viação.

No grupo I, a instabilidade foi confirmada em 60 (92,3%). Registaram-se os seguintes valores: espaço túbio talar medial  $<4\text{mm}$  em 5 (7,7%),  $\geq 4\text{mm}$  em 24 (36,9%) e  $\geq 5\text{mm}$  em 36 (55,4%). Os indivíduos com espaço  $\geq 5\text{mm}$ , 17 (56,7%) classificaram-se predominantemente como 44B2.1 e 9 (56,3%) 44B2.3 (p value= 0,54).

Enquanto no grupo II, a instabilidade foi confirmada em 62 (95,3%), através da medição de um espaço túbio talar medial  $\geq 4\text{mm}$  em 35 (53,8%) e  $\geq 5\text{mm}$  em 27 (41,5%). Os indivíduos com espaço  $\geq 4\text{mm}$ , 20 (66,7%) classificaram-se predominantemente como 44B2.1 e os com espaço  $\geq 5\text{mm}$ , 9 (30,0%) como 44B2.1 e 9 (56,3%) como 44B2.3 (p value = 0,023). A média do espaço túbio talar medial medida em mm no grupo I com lesão de ligamento deltoide (doentes expostos) foi de 6,3 (max 17,7 e min 4) com desvio padrão de 3,2 e intervalo de confiança 95% de 5,5 a 7,2. Nos casos sem lesão do ligamento

deltoide (doentes não expostos) foi de 3,3 (max 3,7 e min 2,6) com desvio padrão de 0,6 e intervalo de confiança 95% de 1,8 a 4,9.

A média do espaço túbio talar medial medida em mm, no grupo II com lesão de ligamento deltoide (doentes expostos) foi de 5,89 (max 23 e min 4) com desvio padrão de 3,4 e intervalo de confiança 95% de 5,0 a 6,7. Nos casos sem lesão do ligamento deltoide (doentes não expostos) foi de 3,1 (max 3,8 e min 2,6) com desvio padrão de 0,6 e intervalo de confiança 95% de 1,6 a 4,7.

No grupo I (**teste stress manual**) e no II (**radiografia gravitacional**) obtivemos uma sensibilidade de 97% e 100%, enquanto a especificidade foi de 100% e 60%. Obtivemos ainda um valor preditivo positivo (VPP) de 100% e um valor preditivo negativo (VPN) de 60%, contra um VPP de 97% e VPN de 100% respetivamente.

Assim registou-se no grupo I, 60 verdadeiros positivos, 3 verdadeiros negativos e 2 falsos negativos enquanto no grupo II a destacar 60 verdadeiros positivos, 2 falsos positivos, 3 verdadeiros negativos (**Quadro 1**).

Procuramos ainda determinar o valor preditivo para lesão do ligamento deltoide tendo em conta a diferença do espaço túbio talar medial de  $\geq 4\text{mm}$  ou  $\geq 5\text{mm}$ . Verificámos que um espaço maior que cinco tem uma sensibilidade, especificidade, VPP e VPN de 100%, respetivamente. Enquanto um espaço maior que quatro tem uma especificidade e VPP de 100% e uma sensibilidade 91% e VPN de 60% (**Quadro 2**).

No diagrama apresentado no quadro 3 e 4, as leituras efetuadas pelo 1º observador foram esquematizadas em tabulação cruzada com a segunda leitura realizada em tempo diferente (intraobservador) e com as do 2º observador (interobservador). Avaliando a variação intraobservador no teste *stress manual*, verificou-se uma concordância percentual de 95%, enquanto na interobservador foi de 88%, havendo uma muito forte associação entre as variáveis em estudo (Qui Square) **com p value=0,000**.

Na **radiografia gravitacional**, na variação intraobservador verificou-se uma concordância percentual de 97%, enquanto na interobservador foi de 91%, havendo também uma muito forte associação entre as variáveis em estudo (Qui Square) com p value=0,000.

Teste	Estado do ligamento deltoide teste gravitacional		
	Com lesão	Sem lesão	Total
stress	VP= 60	FP= 0	60
manual	FN= 2	VN= 3	5
	62	3	65

Teste	Estado do ligamento deltoide teste stress manual		
	Com lesão	Sem lesão	Total
gravitacional	VP= 60	FP= 2	62
	FN= 0	VN= 3	3
	60	5	65

Quadro 1: Comparação dos resultados de um teste dicotómico (teste stress manual e gravitacional) com o estado do ligamento deltoide. VP= verdadeiro positivo FP= falso positivo VN=verdadeiro negativo FN=falso negativo.

Teste	Estado do ligamento deltoide teste gravitacional		
	Com lesão	Sem lesão	Total
stress	$\geq 4$ mm	$< 4$ mm	
manual	20	0	20
	2	3	5
	22	3	25

Teste	Estado do ligamento deltoide teste manual		
	Com lesão	Sem lesão	Total
stress manual	$\geq 5$ mm	$< 4$ mm	
	23	0	23
	0	3	3
	23	3	26

Quadro 2: Comparação dos resultados de um teste dicotómico (teste stress manual e gravitacional) com o estado do ligamento deltoide especificando valor medio do espaço tibio talar medial  $\geq 4$  mm ou  $\geq 5$  mm com lesão e  $< 4$ mm sem lesão.

Teste	Medição espaço tibio talar medial							Total
	1º Observador	Intraobservador			Interobservador			
		$< 4$ mm	$\geq 4$ mm	$\geq 5$ mm	$< 4$ mm	$\geq 4$ mm	$\geq 5$ mm	
stress manual	$< 4$ mm	4	1	0	5	0	0	5
	$\geq 4$ mm	1	22	1	0	19	5	24
	$\geq 5$ mm	0	0	36	0	3	33	36
	Total	5	23	37	5	22	38	65

Quadro 3: Variação do observador no teste stress manual: concordância percentual. O tracejado assinala a concordância.

Teste	Medição espaço tibio talar medial							Total
	1º Observador	Intraobservador			Interobservador			
		$< 4$ mm	$\geq 4$ mm	$\geq 5$ mm	$< 4$ mm	$\geq 4$ mm	$\geq 5$ mm	
gravitacional	$< 4$ mm	3	0	0	3	0	0	3
	$\geq 4$ mm	0	34	1	0	29	6	35
	$\geq 5$ mm	0	1	26	0	0	27	27
	Total	3	35	27	3	29	33	65

Quadro 4: Variação do observador no teste gravitacional: concordância percentual. O tracejado assinala a concordância.

No teste stress manual, consideremos então as leituras no quadro 5, em que por exemplo 7,7 do total de 5 indivíduos com espaço  $< 4$ mm ou 33,8% de 24 indivíduos com espaço  $\geq 4$  mm ou ainda 58,5% dos 36 com espaço  $\geq 5$  mm. Assim a concordância percentual esperada apenas por acaso no intraobservador foi de 45,2% e a estatística Kappa de 0,91. Enquanto no interobservador foi 45,5% e o valor da estatística Kappa foi 0,78.

Ainda se considerarmos no quadro 6, por exemplo para a radiografia gravitacional que 4,6% do total de 3 indivíduos com espaço  $< 4$ mm, ou 53,8% dos 35 indivíduos com espaço  $\geq 4$  mm ou ainda

41,55% dos 27 com espaço  $\geq 5$  mm, a concordância esperada apenas por acaso no intraobservador foi de 46,3% e a estatística Kappa de 0,94. Enquanto no interobservador foi de 45,2% e a estatística Kappa teve um valor de 0,84.

Na análise estatística, quando se verificou a normalidade da população em estudo com o teste Kolmogorov- Smirnov, obteve-se para cada um dos grupos um valor de 0,000, assim rejeitou-se a hipótese de normalidade, pelo que se optou pelo teste de Mann- Whitney, obtendo-se um valor de 0,000, no grupo I e de 0,005 no grupo II, com uma evidência muito forte de uma diferença entre

		Medição espaço tíbio talar medial						Total
		Intraobservador			Interobservador			
		<4mm	≥4 mm	≥5 mm	<4mm	≥4 mm	≥5 mm	
Teste <i>stress</i> manual	<4mm	0,4	0	4,6	0,4	1	3,6	5 (7,7%)
	≥4 mm	3,6	8,5	11,9	2,6	8,1	13,3	24 (36,9%)
	≥5 mm	1	14,5	20,5	2	12,9	21,1	36 (55,4%)
	Total	5	23	37	5	22	38	65
		(7,7%)	(35,4%)	(56,9%)	(7,7%)	(33,8%)	(58,5%)	

Quadro 5: Variação do observador no teste *stress* manual: concordância percentual esperada, apenas por acaso. O tracejado assinala a concordância.

		Medição espaço tíbio talar medial						Total
		Intraobservador			Interobservador			
		<4mm	≥4 mm	≥5 mm	<4mm	≥4 mm	≥5 mm	
Teste gravitacional	<4mm	0,1	0	0	0,1	0,1	2,8	3 (4,6%)
	≥4 mm	0	18,8	1	2,9	15,6	16,5	35 (53,8%)
	≥5 mm	0	1	11,2	0	13,3	13,7	27 (41,5%)
	Total	3	35	27	3	29	33	65
		(4,6%)	(53,8%)	(41,5%)	(4,6%)	(44,6%)	(50,8%)	

Quadro 6: Variação do observador no teste gravitacional: concordância percentual esperada, apenas por acaso. O tracejado assinala a concordância.

expostos e não expostos. Utilizando o teste Chi-Square, houve também uma muito forte associação entre os dois testes em estudo com **p value= 0,000**. Procurando avaliar se existe uma relação linear entre o teste *stress* manual e gravitacional, utilizou-se o coeficiente de correlação Pearson, obtendo-se um valor de 0,712, valor próximo de 1, revelando haver uma forte evidência de correlação direta entre as

duas variáveis (Quadro 7).

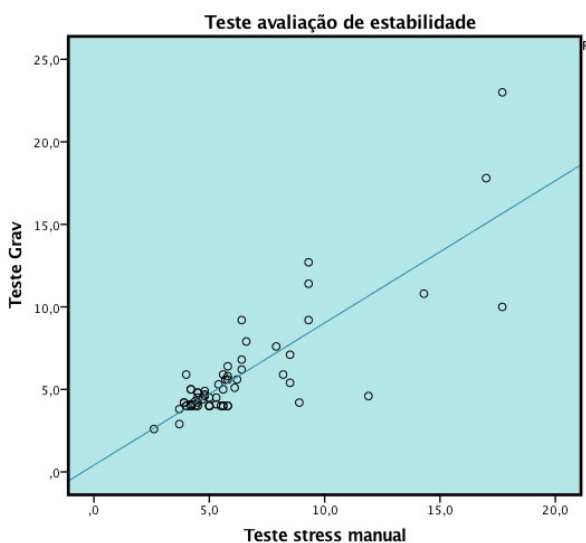
Comparando a validade e a fiabilidade através de representação gráfica, obtivemos uma distribuição dos resultados dos testes desenhada por uma curva dispersa centrada no verdadeiro valor (Quadro 8).

## DISCUSSÃO

A fratura do maleolo lateral é o padrão mais comumente observado nas fraturas do tornozelo. As isoladas são rotineiramente tratadas de forma não cirúrgica. No entanto, quando associadas a lesão do complexo ligamentar medial, devem ser consideradas instáveis e consensualmente tratadas cirurgicamente por forma a restaurar a anatomia e a estabilidade do tornozelo.

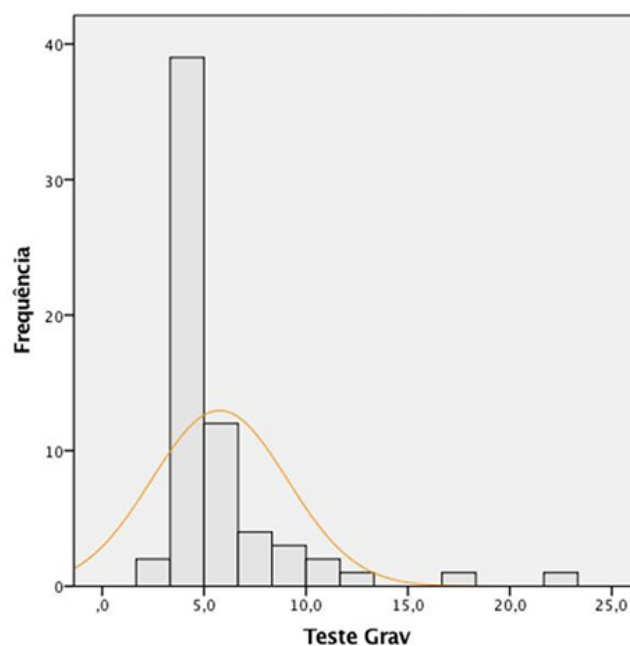
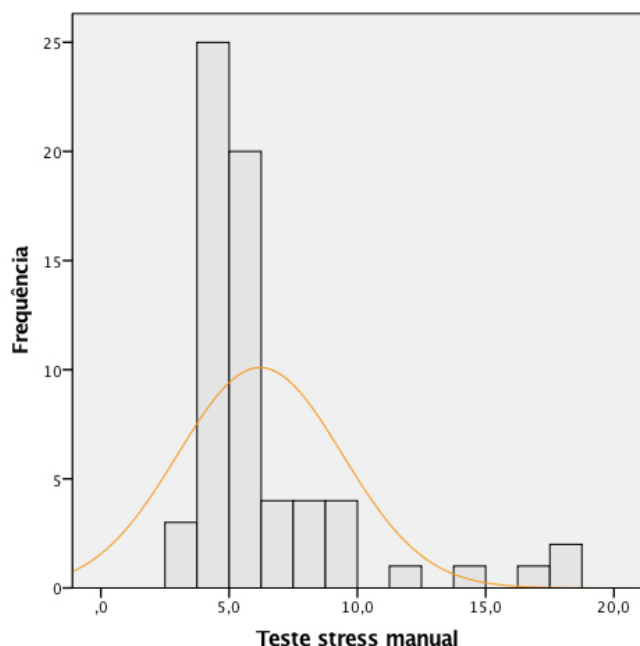
Michelson et al<sup>5</sup> e Gill et al<sup>7</sup> recomendam a radiografia gravitacional, considerando-a tão confiável como o teste de *stress* manual, mas tendo como vantagem a menor exposição à radiação ionizante.

A medição do espaço tíbio talar medial provou ser um meio amplamente aceite de determinar radiograficamente a integridade do ligamento deltoide profundo e consequentemente a sua competência. Contudo, persistem dúvidas quanto ao melhor instrumento a utilizar na avaliação da



Quadro 7: Correlação linear entre os dois testes *stress* manual e gravitacional.





Quadro 8: Resultado de testes validos e fiáveis, com curva centrada no valor medio.

estabilidade nas fraturas maleolares isoladas. Os indicadores do nosso projeto assentam sobre uma população de doentes (N=105) com fratura maleolar isolada, admitidos no serviço de urgência. A elegibilidade dos participantes foi garantida através do cumprimento de critérios de inclusão e exclusão, minimizando a influência de possíveis variáveis de confundimento. Obtivemos assim, uma amostra de 65 indivíduos. A fim de se evitar o enviesamento da análise, os exames foram exclusivamente realizados pelo autor da investigação, embora a sua interpretação tenha sido concretizada por avaliadores independentes, cegos para os resultados, eliminando assim o elemento subjetivo da leitura. A amostra foi dividida por quatro observadores, que receberam instruções para medir o espaço tíbio talar medial nas imagens gravadas no sistema IMPAX®, dos dois testes. Foram utilizados critérios de leitura semelhantes, com medição em mm do espaço tíbio talar medial entre o bordo medial do astrágalo e o bordo lateral do maleolo medial, desenhando uma linha perpendicular a 0,5 cm do rebordo superior da superfície articular medial, à semelhança do preconizado por Travis Motley<sup>14</sup> e Rungprai et al<sup>19</sup>. Também Murphy et al<sup>18</sup>, alertaram para a variabilidade de métodos de medição do espaço tíbio talar medial, aconselhando a realização da **radiografia** a 15° de rotação interna e em carga comparando com o tornozelo contra

lateral, sugerindo que a medição perpendicular tinha maior fiabilidade interobservador e intraobservador quando comparado à medição oblíqua ou ao espaço tíbio talar superior.

No nosso estudo no Grupo I, foi realizado no bloco operatório, o teste *stress* manual. A instabilidade foi confirmada em 60 (92,3 %). A média do espaço tíbio talar medial medida em mm no grupo I com lesão de ligamento deltoide (doentes expostos) foi de 6,3 (max 17,7 e min 4) com desvio padrão de 3,2 e intervalo de confiança 95% de 5,5 a 7,2. Nos casos sem lesão do ligamento deltoide (doentes não expostos) foi de 3,3 (max 3,7 e min 2,6) com desvio padrão de 0,6 e intervalo de confiança 95% de 1,8 a 4,9 (**p = 0,000**).

Enquanto no grupo II, a instabilidade foi confirmada em 62 (95,3 %) e a média do espaço tíbio talar medial medida em mm, nos doentes expostos foi de 5,9 (max 23 e min 4) com desvio padrão de 3,4 e intervalo de confiança 95% de 5,0 a 6,7 e nos não expostos foi de 3,1 (max 3,8 e min 2,6) com desvio padrão de 0,6 e intervalo de confiança 95% de 1,6 a 4,7 (**p = 0,000**).

Estes valores são sobreponíveis aos resultados apresentados por Gill et al<sup>7</sup> em 25 indivíduos, comparando o teste *stress* manual com o gravitacional. Neste estudo, no grupo supinação rotação externa (SER) II o espaço tíbio talar medial médio foi de 4,15 mm e 4,26 (p = 0,50) e no grupo

SER IV, foi de 5,21 e 5,00 mm respetivamente ( $p = 0,69$ ). Houve diferenças significativas entre o grupo SER II e SER IV em relação ao espaço túbio talar medial nos dois grupos,  $p < 0,02$  e  $p < 0,05$  respetivamente.

No nosso estudo, no grupo I do teste *stress* manual obtivemos uma sensibilidade de 97%, identificando assim como positivo 97% dos 62 indivíduos com a lesão do ligamento deltoide ( $n=60$ ). Enquanto o teste gravitacional tinha uma sensibilidade de 100%, identificando, portanto, como positivo 100% dos 60 indivíduos identificados pelo teste do grupo I. São os verdadeiros positivos do grupo I e II. A especificidade no grupo I foi de 100%, enquanto a do grupo II foi de 60%. Obtivemos ainda no grupo I um VPP de 100% e um VPN de 60%, com uma probabilidade de haver lesão em 40%, enquanto na radiografia gravitacional obtivemos um valor de 97% e 100% respetivamente. O ideal seria que todos os indivíduos com a lesão fossem corretamente designados pelo teste como positivos (verdadeiros positivos) e os sem lesão corretamente designados como negativos (verdadeiros negativos), contudo na prática tal é muito raro<sup>9</sup>. O que constitui um potencial problema uma vez que, se o individuo tem a lesão e o teste for negativo (falso negativo), iremos optar por um tratamento conservador ao invés de cirúrgico e se for positivo num sem lesão (falso positivo) estamos a condenar um individuo a uma cirurgia sem necessidade.

Quando se avaliou a lesão do ligamento deltoide tendo em conta a diferença do espaço túbio talar medial de  $\geq 4$  mm ou  $\geq 5$  mm, verificámos que um espaço maior que cinco tinha uma sensibilidade, especificidade, valor preditivo positivo e negativo de 100%, respetivamente. Enquanto um espaço maior que quatro tinha uma especificidade e valor preditivo positivo de 100% e uma sensibilidade de 91% e valor preditivo negativo de 60%, revelando ter menor valor na avaliação da competência do ligamento deltoide. Dados semelhantes foram obtidos com os estudos de Park et al<sup>11</sup>, estes autores mostraram que o espaço túbio talar medial de 5 mm nas **radiografias** de *stress* manual, tinha uma sensibilidade, uma especificidade e valores preditivos positivo e negativo de 100%. Diferentes quando se considerou a medição de 4 mm, com uma sensibilidade de 100% e igual valor preditivo

negativo, contudo, menor especificidade e valor preditivo positivo de 87% e 60% respetivamente.

Estudos recentes em cadáver de Metitiri e Ghorbanhoseini<sup>12</sup> alertaram para o grau de erro de medição e falta de precisão nas medições do espaço túbio talar medial, fundamentais na decisão de tratamento. Também Seidel et al<sup>20</sup>, referem que as **radiografias** *stress* gravitacional, podem sobrestimar as fraturas maleolares laterais isoladas do tipo SER com indicação cirúrgica.

Por conseguinte é fundamental determinar se a leitura entre o mesmo observador e observadores independentes estão de acordo, para além do que seria de esperar também que acontecesse apenas por acaso. De facto, podem ocorrer diferenças entre duas ou mais leituras dos resultados de um mesmo teste feitas pelo mesmo observador (variação intraobservador) assim como entre observadores (variação interobservador), pelo facto de que dois indivíduos raramente obtêm o mesmo resultado, devendo ser expressa a extensão de concordância em termos quantitativos<sup>9</sup>.

Utilizamos o valor de estatística Kappa, proposta por Cohen em 1960<sup>17</sup>. Assim o Kappa exprime até que ponto é que a concordância observada excede a esperada apenas por acaso relativamente ao máximo que os observadores poderiam esperar melhorar na sua concordância.

Landis e Kosch<sup>21</sup> sugeriram que um Kappa maior que 0,75 representa uma excelente concordância, Kappa abaixo de 0,40 representa uma concordância fraca e um valor entre 0,40 e 0,75 entre intermédia e boa.

No nosso estudo quando consideramos a **radiografia** *stress* manual, na avaliação intraobservador verificou-se uma concordância percentual de 95%, uma concordância percentual esperada apenas por acaso de 45,2% e um valor estatística Kappa de 0,91, enquanto no interobservador foi de 88%, 45,5% e 0,78 respetivamente, o que representa uma excelente concordância entre intraobservador e interobservador.

E quando consideramos a **radiografia** gravitacional na avaliação intraobservador, obtivemos uma concordância percentual de 97%, uma concordância percentual esperada apenas por acaso de 46,3% e um valor de estatística Kappa de 0,94, enquanto no interobservador foi de 91%, 45,2% e 0,84

respetivamente, o que representa também uma excelente concordância.

Como ponto fraco do estudo, as imagens obtidas dependeram muito da experiência do técnico de imagiologia, da qualidade da imagem e da própria colaboração do utente, também é relevante a dificuldade em padronizar a força necessária a aplicar no teste de *stress* manual.

Este último teve a vantagem de ter sido efetuado sob anestesia eliminando o viés subjetivo da dor ou desconforto do doente. Contudo, é condicional a um exame de diagnóstico prévio como a **radiografia** gravitacional.

Apesar de considerarmos serem necessários mais estudos comparativos, comprovou-se o valor destes dois instrumentos na avaliação da competência do ligamento deltoide, obtendo-se através de representação gráfica uma distribuição dos resultados desenhada por uma curva dispersa centrada no verdadeiro valor, indicador que os resultados obtidos foram ao mesmo tempo validos e fiáveis. Também o coeficiente de correlação Pearson de 0,712, valor próximo de 1, mostrou haver uma forte evidência de correlação direta entre os dois testes.

## CONCLUSÃO

No nosso estudo verificamos que o espaço túbio talar medial  $\geq 5$  mm tem uma sensibilidade, especificidade, valor **preditivo positivo e negativo**, superior na avaliação da integridade do ligamento deltoide.

O teste **stress manual** assim como o **gravitacional** revelaram ser uma ferramenta eficaz na avaliação da competência do ligamento deltoide, podendo considerar-se fator fundamental na decisão terapêutica em fraturas do maléolo lateral isoladas.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: a review. *Injury*. 2006; 37: 691-697
2. Pankovich AM, Shivaram MS. Anatomical basis of variability in injuries of the medial malleolus and the deltoid ligament: I. Anatomical studies. *Acta Orthop Scand*. 1979; 50 (2): 217-223
3. Stufkens SA, Bekerom M, Knupp M, Hintermann B, van Dijk CN. The diagnosis and treatment of deltoid ligament lesions in supination-external rotation ankle fractures: a review. *Strat Traum Limb Recon*. 2012; 7: 73-85
4. Campbell K, Michalski M, Wilson K, Goldsmith M, Wijdicks C, LaPrade R, et al. The Ligament Anatomy of the Deltoid Complex of the Ankle: A Qualitative and Quantitative Anatomical Study. *J Bone Joint Surg Am*. 2014; 96 (62): 1-10
5. Michelson JD, Varner KE, Checcone M. Diagnosing deltoid injury in ankle fractures: the gravity stress view. *Clin Orthop Relat Res*. 2001; 387: 178-182
6. McConnell T, Creevy W, Tornetta P 3rd. Stress examination of supination external rotation-type fibular fractures. *J Bone Joint Surg Am*. 2004; 86: 2171-2178
7. Gill JB, Risko T, Raducan V, Grimes JS, Schutt RC Jr. Comparison of manual and gravity stress radiographs for the evaluation of supination-external rotation fibular fractures. *J Bone Joint Surg*. 2007; 89 (5): 994-999
8. Schock HJ, Pinzur M, Manion L, Stover M. The use of gravity or manual-stress radiographs in the assessment of supination external rotation fractures of the ankle. *J Bone Joint Surg Br*. 2007; 89 (8): 1055-1059
9. Gordis L. Avaliar a Validade e a Fiabilidade dos Testes de Diagnostico e de Rastreio. *Epidemiologia*, 4ed. New York: Lusodidacta- Soc. Port. De Material Didactico, Ida; 2011. p. 85-108.
10. DeAngelis NA, Eskander MS, French BG. Does medial tenderness predict deep deltoid ligament incompetence in supination-external rotation type ankle fractures?. *J Orthop Trauma*. 2007; 21 (4): 244-247
11. Park SS, Kubiak EN, Egol KA, Kummer F, Koval KJ. Stress radiographs after ankle fracture: the effect of ankle position and deltoid ligament status on medial clear space measurements. *J Orthop Trauma*. 2006; 20 (1): 11-18
12. Metitiri O, Ghorbanhoseini M, Zurakowski D, Hochman M, Nazarian A, Kwon J. Accuracy and Measurement Error of the Medial Clear Space of the Ankle. *Foot & Ankle International*. 2017; 38 (4): 443-451
13. Egol KA, Amirtharajah M, Tejwani NC, Capla EL, Koval KJ. Ankle stress test for predicting the need for surgical fixation of isolated fibular fractures. *J Bone Joint Surg*. 2004; 86 (11): 2393-2398
14. Motley T, Clements JR, Moxley K, Carpenter B, Garrett A. Evaluation of the Deltoid Complex in Supination External Rotation Ankle Fractures. *The Foot and Ankle Online Journal*. 2010; 3 (4): 1
15. DeAngelis JP, Anderson R, DeAngelis NA. Understanding the superior clear space in the adult ankle. *Foot Ankle Int*. 2007; 28 (4): 490-493
16. Gougoulas N, Sakellariou A. When is a simple fracture of the lateral malleolus not so simple? How to assess stability, which ones to fix and the role of the deltoid ligament. *Bone Joint J*. 2017 Jul; 99 (7): 851-855
17. Cohen J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*. 1960; 20 (1): 37
18. Murphy JM, Kadakia AR, Irwin TA. Variability in radiographic medial clear space measurement of the normal weight-bearing ankle. *Foot Ankle Int*. 2012 Nov; 33 (11): 956-963
19. Rungprai C, Goetz JE, Arunakul M, Gao Y, Femino JE, Amendola A, et al. Validation and reproducibility of a biplanar imaging system versus conventional radiography of foot and ankle radiographic parameters. *Foot Ankle Int*. 2014 Nov; 35 (11): 1166-1175
20. Seidel A, Krause F, Weber M. Weightbearing vs gravity stress radiographs for stability evaluation of SER fractures of the ankle. *Foot and Ankle Surgery*. 2017; 23 (1): 1-28
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33: 159-174