

Concordancia interobservador en el diagnóstico tomográfico de dolor abdominal en estudio entre residentes de Imagenología y médicos imagenólogos

Interobserver agreement between radiology residents and radiology attendants in the tomographic diagnosis of abdominal pain

Jean P. Negreros-Osuna*, Lucia Zatarain-Bayliss, Edgar Dehesa-Lopez, Felipe J. Peraza-Garay.

Unidad de Imagenología (UNIMA), Hospital Civil de Culiacán, Centro de Investigación y Docencia en Ciencias de la Salud. Universidad Autónoma de Sinaloa.

***Autor de correspondencia:** Jean Paul Negreros Osuna
Matías Lazcano 1425-5 Colonia Chapultepec, Culiacán Sinaloa C.P. 80040
Tel: (669) 215-7155, Correo electrónico: negrerosjp@gmail.com

DOI <http://dx.doi.org/10.28960/revmeduas.2007-8013.v12.n1.005>

Recibido 02 de marzo 2021, aceptado 03 de septiembre 2021

RESUMEN

Objetivo: Determinar la concordancia de los diagnósticos tomográficos de pacientes con dolor abdominal entre los residentes de imagen y los médicos adscritos al servicio de UNIMA del Hospital Civil de Culiacán. **Material y Métodos:** Estudio prospectivo, observacional y transversal, tomando como datos de estudio 100 tomografías realizadas a pacientes con dolor abdominal que acudieron a UNIMA, que cumplían con los criterios de calidad para realizar un diagnóstico imagenológico adecuado y utilizando el coeficiente Kappa para calcular la concordancia entre 4 residentes y 2 médicos adscritos. **Resultados:** Se obtuvo una concordancia de 88% ($p < .0001$), entre el residente 1 y médicos adscritos, el residente 2 obtuvo una concordancia de 90% ($p < .0001$), los residentes 3 y 4, obtuvieron un 91% y 89% ($p < .0001$), respectivamente. **Conclusión:** La concordancia obtenida entre los residentes y médicos adscritos fue de un 89.5% ($p < .0001$) en promedio; Es interpretada según el coeficiente Kappa como casi perfecta, dando seguridad en que los estándares de enseñanza en el servicio de UNIMA son altos en el escenario de dolor abdominal en estudio. **Palabras Clave:** Concordancia interobservador, tomografía, dolor abdominal, Coeficiente Kappa.

ABSTRACT

Objective: To determine the concordance of the tomographic diagnoses of patients with abdominal pain between image residents and doctors assigned to the UNIMA service of the Hospital Civil de Culiacán. **Material and Methods:** Prospective, observational and cross-sectional study, taking as study data 100 tomographs performed on patients with abdominal pain who attended UNIMA, who met the quality criteria to make an adequate imaging diagnosis and using the Kappa coefficient to calculate the Concordance between 4 residents and 2 attendant doctors. **Results:** The agreement was calculated and obtained a concordance of 88% ($p < .0001$), between resident 1 and attendant doctors, resident 2 obtained a concordance of 90% ($p < .0001$), residents 3 and 4, obtained a 91% and 89% ($p < .0001$), respectively. **Conclusion:** The concordance obtained between residents and attendant doctors was 89.5% ($p < .0001$) on average; It is interpreted according to the Kappa coefficient as almost perfect, giving certainty that the teaching standards in the UNIMA service are high in the abdominal pain scenario. **Keywords:** Interobserver agreement, tomography, abdominal pain, Kappa coefficient.

Introducción

La validez de un estudio puede verse afectada si se utilizan mediciones poco fiables. Una fuente de error de medición es producto de la variabilidad interobservador, cuya magnitud es posible estimar a través de los estudios de concordancia.

Puede ser calculado siempre y cuando se contrasten dos observadores; para la evaluación de concordancia de tres o más observadores se usa el coeficiente Kappa de Fleiss.

La TC es la modalidad diagnóstica de elección en la mayor parte de las patologías abdominales, incluyendo los traumatismos.

La historia clínica y la exploración física siguen siendo esenciales en la evaluación de las alteraciones abdominales para definir la etiología y también cuál de los estudios de imagen es más adecuado para establecer el diagnóstico correcto, en el caso de ser necesario.¹

La tomografía se utiliza en múltiples escenarios desde neoplasias hasta traumatismos abdominales, se suele recomendar al paciente que no ingiera alimentos sólidos durante 4 a 6 horas antes del estudio. El paciente se coloca en la camilla del escáner con los brazos por encima de la cabeza.²

Contraste oral: se administra un contraste positivo, yodado o baritado diluido en agua al 3-5%. El paciente ingiere entre 250 a 350 ml cada 20 a 30 minutos durante aproximadamente 1.5 horas antes de comenzar la prueba. En total se administran 1 a 1,5 L de contraste oral.

Se suele inyectar el contraste en una vena del antebrazo a un flujo de 2 a 3 ml/s y la adquisición de las imágenes se hace entre 60 y 70 segundos después de comenzada la inyección.³

La dosis de contraste suele ser de unos 100 a 150 ml de contraste de unos 300 mg de yodo por mililitro. Puede ajustarse la dosis al peso del paciente. La inyección de un bolo de suero tras el contraste aumenta la intensidad del realce al arrastrar el contraste de las venas hacia el corazón.

Lo habitual es adquirir las imágenes del abdomen y la pelvis en una sola serie, pero también

puede esperarse de 2 a 4 minutos antes de adquirir las imágenes de la pelvis para que la vejiga y las venas pelvianas se rellenen de contraste.⁴

Las imágenes de TC del abdomen deben reconstruirse con algoritmos <<estándar>>, que reducen de forma notable el ruido de la imagen en relación con los filtros que se utilizan para pulmón o hueso.

Es importante elegir las <<ventanas>> de visualización apropiadas. Una ventana intermedia es útil para una visión general del abdomen, pero debe cambiarse a una ventana más estrecha para evaluar lesiones en órganos sólidos y a una ventana más amplia para detectar gas extraluminal o para evaluar el hueso o las áreas de los pulmones que incluimos en el estudio abdominal.⁵

El dolor abdominal es uno de los síntomas más frecuentes de consulta en el servicio de urgencias, generalmente asociado a un proceso intra-abdominal, si bien puede ser debido a un trastorno extra-abdominal.

El término abdomen agudo (AA) no debería asociarse a intervención quirúrgica sino a la necesidad de realizar un diagnóstico y tratamiento precoz. Por lo que éste se puede definir como un dolor de instauración más o menos rápida que puede o no requerir tratamiento quirúrgico.⁶

El dolor visceral está originado por un estímulo de los receptores del peritoneo visceral (estiramiento o contracción de víscera hueca, disten-

sión de la cápsula de un órgano sólido, inflamación, isquemia) y es transmitido por vía simpática a las astas dorsales y a centros nerviosos superiores; es sordo y difuso.

El dolor parietal, es transmitido por fibras somáticas a la medula espinal y se localiza en el lugar de la lesión.

El dolor puede ser continuo, suele indicar afectación peritoneal o intermitente que suele indicar afección a víscera hueca.⁷

Desde un punto de vista práctico y de acuerdo con la presentación clínica y a los mecanismos etiopatogénicos, casi todas las patologías incluidas dentro del AA podrían encuadrarse en alguno de los siguientes grupos.

Así pues, tenemos: dolor abdominal y shock, peritonitis generalizada, peritonitis localizada, obstrucción intestinal y causas médicas.

Dolor abdominal y shock: las causas más frecuentes son la rotura de un aneurisma aórtico y de un embarazo ectópico, en los que la única opción es la cirugía inmediata. Otras causas debido a la pérdida de fluidos serían la oclusión intestinal, la isquemia mesentérica o la pancreatitis aguda grave.

Peritonitis generalizada: las causas más frecuentes son la perforación gastroduodenal y colónica (incluyendo apendicular). En estos casos, el tratamiento debe ser quirúrgico.

Peritonitis localizada: Las causas más frecuentes son la apendicitis aguda y las ginecológicas en mujeres jóvenes en los cuadrantes inferiores, la colecistitis aguda en el superior derecho

y la diverticulitis aguda en el inferior izquierdo. La valoración quirúrgica resulta fundamental.⁸

Hay que recordar que la mitad de los dolores abdominales que acuden a la urgencia quedan sin diagnóstico etiológico, denominándose dolor abdominal inespecífico (DAI). Se define como un dolor que dura un máximo de 7 días, sin causa inmediata aparente y que no requiere intervención quirúrgica.

Se ha descrito recientemente la asociación entre el DAI, el síndrome de colon irritable y la enfermedad celíaca. Las mujeres representan el 75% de ingresos con DAI.⁹

Las emergencias comprenden el 50% de la práctica quirúrgica general con el abdomen agudo siendo el que contribuye mayoritariamente a estas referencias. La evaluación clínica continúa siendo muy importante como primer paso, sin embargo, se ha encontrado que es precisa solo en un 47-76% de los pacientes.¹⁰

En Norte-América la TC para un abdomen agudo es frecuentemente obtenida a través del departamento de urgencias evitando admisiones innecesarias, y aumentando el nivel de certeza diagnóstica. Un estudio en Reino Unido mostro una reducción en la estancia hospitalaria y la mortalidad con el uso de TC temprana (dentro de las primeras 24 h de presentación), aunque carecía de poder y no mostro una reducción significativa.¹¹

En un estudio prospectivo de pacientes sometidos a TC en presencia de abdomen agudo los

diagnósticos confirmados más frecuentes fueron por orden decreciente: dolor abdominal no específico, diverticulitis aguda/absceso, apendicitis aguda, colitis/EII, perforación visceral, patología tubo-ovárica y otras.¹²

La mayoría de los estudios que evalúan la concordancia interobservador lo hacen solo entre dos médicos radiólogos especialistas y generalmente evalúan su acuerdo en una patología específica limitando la validación en cuanto a la utilidad de un método diagnóstico.^{13 14}

Son pocos los estudios que evalúan la concordancia entre los hallazgos de estudios de una o más regiones corporales de los cuales no solo obtenemos datos sobre la fiabilidad de la interpretación sino los diagnósticos más comunes e indicaciones de dichos estudios.¹⁵

Lo mismo sucede al momento de evaluar a residentes de radiología con los médicos radiólogos especialistas en los que la mayoría evalúan la concordancia en el diagnóstico de Tromboembolia pulmonar, además ser pocos los que hacen este tipo contraste entre residentes y especialistas.^{16 17 18}

Otros estudios carecen de una población de estudio significativa ya sea por la complejidad del protocolo de TC que están evaluando o rareza de la patología, lo que impide extrapolar sus resultados a otras demografías, limitando también su validez interna.^{19 20}

Material y métodos

Es una prueba diagnóstica, tipo observacional, prospectivo y transversal.

El universo de estudio fueron pacientes mayores de edad que acudieron al servicio de UNIMA con dolor abdominal en estudio para la realización de una tomografía.

El lugar donde se realizó fue el servicio de UNIMA del Hospital Civil de Culiacán. Teniendo como periodo de realización del 01 de Mayo del 2017 al 01 de Marzo del 2019.

Los criterios de inclusión fueron estudios tomográficos de pacientes mayores de 18 años con diagnóstico de dolor abdominal en estudio que acudieron al servicio de UNIMA de Hospital Civil de Culiacán.

Los criterios de exclusión fueron estudios tomográficos con criterios de calidad deficientes para establecer un diagnóstico confiable y pacientes con estudios de imagen previos de su padecimiento actual.

Se tomaron como criterios de eliminación, estudios tomográficos que no contaron con interpretación de forma individual por el médico adscrito y posteriormente por médico residente del servicio de UNIMA.

Para el análisis estadístico se utilizó el coeficiente Kappa para determinar la concordancia entre residentes y médicos especialistas en Imagenología, traduciendo dichos resultados a la escala propuesta por Landis y Koch para establecer el grado de acuerdo entre ambos.

Se estimó un tamaño de muestra de 81 pacientes suponiendo un porcentaje de acuerdo del 70%, con una precisión del 10% y una confianza del 95%, utilizando el coeficiente Kappa.

Las variables estudiadas fueron los posibles diagnósticos más comunes entre ellos la litiasis renal, pancreatitis y obstrucción intestinal, siendo en su mayoría variables del tipo categórica nominal y categórica ordinal en el caso de ser diagnósticos que cuentan con una escala de gravedad.

Los estudios de tomografía abdominal obtenidos que cumplieron con los estándares de calidad para su correcta interpretación primero fueron valorados aleatoriamente por los residentes de imagenología de segundo año.

Posteriormente dichos estudios fueron interpretados aleatoriamente por 2 médicos adscritos al servicio de UNIMA, ninguno de los participantes tuvo información clínica acerca del paciente.

Se tomó como “patrón de referencia” (Gold Standar) la interpretación final escrita por los médicos adscritos, en los casos en los que existió duda sobre el diagnóstico final se hizo un consenso entre ambos médicos el cual fue tomado en su caso como el “patrón de referencia”.

Una vez obtenidos la totalidad de los datos, se utilizó para su análisis estadístico el coeficiente Kappa, determinando la concordancia entre los hallazgos, se realizó la interpretación de los resultados y las conclusiones respecto al estudio.

Resultados

En total se obtuvieron 100 estudios de tomografía computarizada en pacientes mayores de edad que acudieron al servicio de UNIMA con clínica dolor abdominal, los cuales cumplieron con los estándares de calidad para su correcta interpretación.

La población estudiada tuvo una distribución del 58%(n=58) para el sexo femenino y del 42%(n=42) para el masculino, con una media de edad de 44 años en el sexo femenino con una desviación estándar de 17.5 y de 52 años en el masculino, con una desviación estándar 16.8

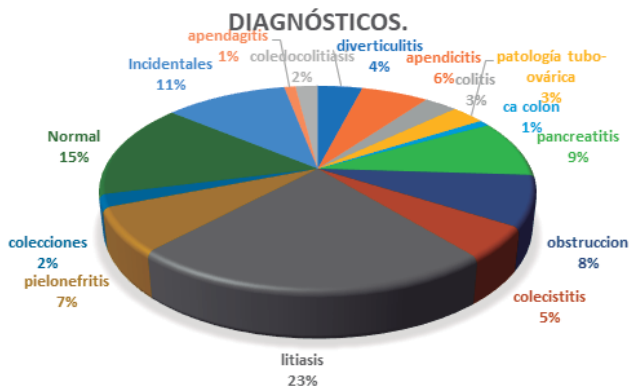
Se obtuvo un diagnóstico tomográfico concluyente en 74%(n=74) de los estudios, 15%(n=15) de los estudios tomográficos fueron concluidos como normales, en un 11%(n=11) estudios no se llegó a un diagnóstico concluyente y se identificaron hallazgos inespecíficos. El diagnóstico más común como causa del dolor abdominal fue la litiasis renal tanto en los pacientes masculinos como en los femeninos, con un 23 % (n = 23) del total de casos

El segundo y tercer diagnóstico más común fueron pancreatitis aguda y obstrucción intestinal, con un 9% (n=9) para pancreatitis y 8%(n=8) para obstrucción intestinal los cuales representaban en conjunto un 17%(n=17) del total de casos.

Otros diagnósticos encontrados fueron 4%(n=4) de tomografías con hallazgos de diverticulitis, 3 presentados en mujeres y 1 hombre, 3%(n=3)

de tomografías con hallazgos sugestivos de colitis en pacientes masculinos, 3%(n=3) de tomografías con datos de patología tubo-ovárica y 1%(n=1) con hallazgos de CA colorrectal en un paciente masculino (grafica 1).

Grafica 1. Porcentaje de diagnósticos tomográficos en dolor abdominal.



La concordancia interobservador obtenida en el estudio fue de 89.5% en promedio con un Kappa promedio de 0.8780 (p<.0001) entre los 4 residentes y 2 médicos adscritos al servicio de UNIMA en el diagnostico tomográfico de dolor abdominal en estudio.

Los valores individuales obtenidos para el primer residente evaluado fueron una concordancia interobservador de 88% con un coeficiente Kappa de 0.8651 (p<.0001).

Los diagnósticos en los que el residente 1 tuvo una discordancia respecto a los médicos adscritos fueron principalmente colecistitis, patología tubo ovárica, apendicitis y litiasis renal.

En el segundo residente se obtuvieron una concordancia interobservador de un 90% y un coeficiente Kappa de 0.8859 (p<.0001).

Los diagnósticos en los que hubo discordancia entre el residente 2 y los médicos adscritos fueron patología tubo-ovárica, colitis, coledocolitiasis y colecistitis, en total hubo discordancia en 10%(n=10) tomografías.

El tercer residente obtuvo un 91% de concordancia interobservador con un coeficiente Kappa de 0.8975 (p < .0001).

Los diagnósticos en los que no hubo acuerdo entre el residente 3 y los médicos adscritos fueron colitis, colecistitis y coledocolitiasis, hubo discordancia en 7%(n=7) tomografías en total. Finalmente, el cuarto residente evaluado obtuvo una concordancia interobservador de 89% con un coeficiente Kappa de 0.8738 (p < .0001). (Tabla 1).

Tabla 1. Concordancia interobservador en diagnostico tomográfico por dolor abdominal.

Residente	Acuerdo	Acuerdo esperado	Kappa	Desviación estándar	P
1	88%	11.05%	0.8651	0.0336	<.0001
2	90%	12.39%	0.8859	0.0357	<.0001
3	91%	12.23%	0.8975	0.0355	<.0001
4	89%	12.84%	0.8738	0.0363	<.0001

En este cuarto residente los diagnósticos en los que hubo discordancia fueron patología tubo-ovárica, colecistitis, colitis y coledocolitiasis, en 9% (n=9) de tomografías en total.

Todos los residentes tuvieron discordancia con los médicos adscritos en cuanto al reporte de hallazgos incidentales que no explicaban el dolor abdominal, así como la decisión de concluir un estudio como normal.

La mayoría de las diferencias entre diagnósticos se pueden explicar por la falta de experiencia entre los residentes y médicos adscritos, sobre todo en diagnósticos en los que la tomografía no es el método ideal, así como en hallazgos que pueden ser muy sutiles.

Discusión

La concordancia interobservador obtenida en el estudio fue de 89.5% en promedio entre los 4 residentes y 2 médicos adscritos al servicio de UNIMA en el diagnóstico tomográfico de dolor abdominal en estudio.

En el presente estudio se tenía como objetivo identificar el nivel de acuerdo entre residentes de imagenología y médicos adscritos al servicio de UNIMA, con el propósito de evaluar la docencia y aprendizaje.

Ya que está bien establecido que la tomografía es un método diagnóstico eficaz en la detección de patologías abdominales agudas, se realizó en cambio una evaluación de el diagnóstico final concreto dado por cada participante en el escenario de paciente con dolor abdominal en estudio.^{3 4}

En el estudio por Keedy et al, en el que se tenía como objetivo valorar el acuerdo interobservador en la caracterización de cambios cirróticos por tomografía multicorte.¹³

Se utilizó el análisis estadístico Kappa en el que obtuvieron un valor de 0.70, encontraron que la detección general de cirrosis fue la observación más objetiva y precisa.

Concluyeron que la tomografía tiene una alta precisión y concordancia interobservador en el diagnóstico de cirrosis, pudiendo ser una prueba diagnóstica suplementaria en pacientes con contraindicaciones para biopsia o hallazgos equívocos.

En otro estudio realizado por Constantino et al se tenía como objetivo obtener el nivel de concordancia interobservador para el diagnóstico de tromboembolismo pulmonar por medio de la tomografía computarizada multidetector, la cual es una de las mejores herramientas diagnósticas para su detección.¹⁶

Ellos tuvieron la participación de 4 radiólogos con diferentes grados de experiencia ya que este es uno de los factores que interfiere el correcto diagnóstico, así como la experiencia del personal operador del equipo y se refleja en las diferencias de interpretación de estudios.

Los 4 radiólogos evaluaron 46 diferentes tomografías computarizadas multidetector realizadas por sospecha de tromboembolismo pulmonar agudo, tuvieron que verificar su presencia o ausencia, y de ser positivo localizarla y cuantificarla.

Se utilizó el análisis estadístico Cohen K (Kappa) para obtener la concordancia la cual fue de 0.82 en promedio entre los 4 participantes, sin embargo, esto fue en trombosis masivas, en el caso de segmentarias y subsegmentarias hubo un acuerdo de solo 0.47, por lo que recomendaron una segunda evaluación por un experto en estos casos.

Una de las preocupaciones para el presente estudio era demostrar que, aunque los residentes tienen una menor experiencia están capacitados para dar un diagnóstico correcto en el escenario de dolor abdominal.

En los diagnósticos en los que se requiere mayor experiencia ya que no es el método ideal como en el caso de colecistitis o coledocolitiasis hubo una menor concordancia entre residentes y médicos adscritos, pero en los diagnósticos en los que la tomografía tiene alta precisión la concordancia promedio fue de 89.5% con coeficiente Kappa de 0.87.

Se obtuvieron resultados satisfactorios de los 4 residentes con acuerdos que van desde 0.86 a 0.89 ($p < .0001$) (ver resultados), dando certeza de que los diagnósticos emitidos de forma preliminar en "horario fuera de oficina" por residentes de guardia son confiables.

Algunas de las fortalezas de nuestro estudio fue que logramos acumular un número alto de estudios tomográficos que presentaban unos estándares de calidad adecuados para su interpretación, siendo un total de 100 estudios, la muestra calculada para tener unos resultados estadísticamente significativos fue de 81 estudios.

También obtuvimos la participación de todos los residentes que cursaban el segundo año de la residencia en imagenología diagnóstica y terapéutica los cuales son los responsables

de guardia, y además contamos con la participación de los médicos adscritos encargados del área de tomografía para que dieran su evaluación de los estudios.

La diversidad de diagnósticos tomográficos que se obtuvo fue amplia, lo que provee un panorama general de la patología más frecuente en la población de estudio, y al mismo tiempo da certeza en que los resultados obtenidos en cuanto a la concordancia son estadísticamente significativos.

Las limitaciones del estudio fueron que se evaluó solamente la concordancia en un escenario clínico específico, por lo que sería importante evaluar el acuerdo interobservador en escenarios menos frecuentes o en los que no se tiene un entrenamiento tan extenso como es el caso de patología musculoesquelética o cardiovascular.

Conclusión

Se obtuvo un resultado positivo con una muy buena concordancia interobservador en el diagnóstico tomográfico en pacientes con dolor abdominal entre los residentes de radiología y los médicos radiólogos en el servicio de UNIMA del Hospital Civil de Culiácan.

La tomografía es el método ideal para evaluar la mayoría de las patologías abdominales que requieren una atención inmediata por lo que el contar con la confianza de que se dispone del

mejor diagnóstico radiológico posible en todo momento es fundamental.

En los diagnósticos más comunes e importantes no hubo diferencia significativa en las interpretaciones, solamente se sugeriría una segunda evaluación "no urgente" en casos donde no se encuentre una causa evidente del dolor.

Es recomendable realizar estudios similares para evaluar el grado de acuerdo en otras áreas de la radiología como sistema nervioso central, sistema musculoesquelético y cardiovascular, ya que son áreas en las que hay menor cantidad de casuística en comparación a la patología abdominal.

Referencias

- Herring W. Radiología básica + StudentConsult: Aspectos fundamentales; Elsevier Health Sciences Spain;2012. p. 105-106.
- Prokop M. General principles of MDCT. Eur J Radiol. 2003 Mar;45 Suppl 1:S4-10.
- Fenchel S, Fleiter TR, Merkle EM. Multislice helical CT of the abdomen. Eur rad. 2002 Jul;12 Suppl 2:S5-10
- Cahir JG, Freeman AH, Courtney HM. Multi-slice CT of the abdomen. BJR. 2004;77(suppl_1):S64-S73.
- Reiner BI, Siegel EL, Hooper FJ. Accuracy of interpretation of CT scans: comparing PACS monitor displays and hard-copy images. AJR. 2002 Dec;179(6):1407-10.
- Grundmann RT, Petersen M, Lippert H, Meyer F. The acute (surgical) abdomen - epidemiology, diagnosis and general principles of management. Z Gastroenterol. 2010 Jun;48(6):696-706.
- Robinson DA, Lin E. Cope's Early Diagnosis of the Acute Abdomen: 20th Edition. Ann Surg. 2001;233(4):594.
- Moshe S, Lane R, Rogers P, Assalia A. Schein's Common Sense Emergency Abdominal Surgery: An Unconventional Book for Trainees and Thinking Surgeons; Berlin Heidelberg; Springer 2010.
- Knopp RK, Dries D. Analgesia in acute abdominal pain: what's next? Ann. Emerg. Med. 2006 Aug;48(2):161-3.
- Rosen MP, Sands DZ, Longmaid HE, 3rd, Reynolds KF, Wagner M, Raptopoulos V. Impact of abdominal CT on the management of patients presenting to the emergency department with acute abdominal pain. AJR. 2000 May;174(5):1391-6.
- Ng CS, Watson CJ, Palmer CR, See TC, Beharry NA, Housden BA, et al. Evaluation of early abdominopelvic computed tomography in patients with acute abdominal pain of unknown cause: prospective randomised study. BMJ (Clinical research ed). 2002 Dec 14;325(7377):1387
- Salem TA, Molloy RG, O'Dwyer PJ. Prospective study on the role of the CT scan in patients with an acute abdomen. Colorectal disease: the official journal of the Association of Coloproctology of Great Britain and Ireland. 2005 Sep;7(5):460-6.

13. Keedy A, Westphalen AC, Qayyum A, Aslam R, Rybkin AV, Chen MH, et al. Diagnosis of cirrhosis by spiral computed tomography: a case-control study with feature analysis and assessment of interobserver agreement. *JCAT*. 2008 Mar-Apr;32(2):198-203.
14. Stafira JS, Sonnad JR, Yuh WT, Huard DR, Acker RE, Nguyen DL, et al. Qualitative assessment of cervical spinal stenosis: observer variability on CT and MR images. *AJNR*. 2003 Apr;24(4):766-9.
15. Colledge N, Sellar R, Wardlaw J, Lewis S, Mead G, Wilson J. Interobserver agreement in magnetic resonance brain and neck imaging. *J Neuroimaging*. 2006;16(1):47-51.
16. Costantino G, Norsa AH, Amadori R, Ippolito S, Resta F, Bianco R, et al. Interobserver agreement in the interpretation of computed tomography in acute pulmonary embolism. *Am. J. Emerg. Med*. 2009 Nov;27(9):1109-11
17. Verweij JW, Hofstee HM, Golding RP, van Waesberghe JH, Smulders YM. Interobserver agreement between on-call radiology residents and radiology specialists in the diagnosis of pulmonary embolism using computed tomography pulmonary angiography. *JCAT*. 2009 Nov-Dec;33(6):952-5.
18. Ginsberg MS, King V, Panicek DM. Comparison of interpretations of CT angiograms in the evaluation of suspected pulmonary embolism by on-call radiology fellows and subsequently by radiology faculty. *AJR*. 2004 Jan;182(1):61-6.
19. Ippolito D, Casiraghi AS, Talei FC, Bonaffini PA, Fior D, Sironi S. Intraobserver and Interobserver Agreement in the Evaluation of Tumor Vascularization With Computed Tomography Perfusion in Cirrhotic Patients With Hepatocellular Carcinoma. *JCAT*. 2016 Jan-Feb;40(1):152-9.
20. Petralia G, Preda L, Raimondi S, D'Andrea G, Summers P, Giugliano G, et al. Intra- and Interobserver Agreement and Impact of Arterial Input Selection in Perfusion CT Measurements Performed in Squamous Cell Carcinoma of the Upper Aerodigestive Tract. *AJNR*. 2009 June 1, 2009;30(6):1107-15.