Eficacia de la TAC helicoidal y la FDG-PET para la identificación de metástasis ganglionares mediastínicas en pacientes con carcinoma broncogénico no microcítico (CBNM) potencialmente resecable. Una perspectiva histórica.

Francisco Pozo-Rodríguez[1], José L. Martín de Nicolás[1], María A. Sánchez-Nistal[1], Antonio Maldonado[2], Santiago García de Barajas[1], R. Calero-García[1], Miguel A. Pozo[2], Pedro Martín-Escribano[1], Isabel Martín-García[1], y Ricardo García-Lujan[1].

(1) Hospital 12 de Octubre (2) Centro PET Complutense

1.- Introducción. En el año 1997 se publicó un trabajo (Goldwin 1977) que sistematizaba por primera vez la anatomía normal del mediastino en Tomografía Computarizada (TAC), y en 1994 el primero que abordaba de forma sistemática el papel de la Tomografía por Emisión de Positrones (FDG-PET) en la estadificación del CBNM (Wahl 1994). Desde principios de los 80 hasta la actualidad se han realizado cientos de estudios sobre la eficacia diagnóstica de la TAC y la PET en la estadificación de los pacientes con CBNM.

Los estudios sobre tests diagnósticos están diseñados para obtener estimadores de eficacia como la sensibilidad y especificidad, pero están amenazados por una amplia variedad de errores sistemáticos o sesgos que pueden afectar a su validez científica (Begg 1987; Begg and McNeil, 1988). Un espectro inadecuado de enfermedad, la ausencia de un buen estándar de referencia (*gold standard*), la incorporación del test problema en la definición del estándar de referencia, los sesgos de interpretación debidos a una lectura no ciega de los tests o al conocimiento de datos clínicos orientativos, el sesgo de verificación diagnóstica, la gestión de los resultados no interpretables, la variación del criterio de positividad y la variación interobservador son los principales sesgos que amenazan los estudios de diagnóstico. El tamaño de la muestra, y su distribución dentro de la tabla son elementos importantes para asegurar la estabilidad numérica de los índices diagnósticos. Los estándares de calidad de los estudios han sido tradicionalmente pobres (Sheps and Schechter 1984; Reid, Lachs, and Feinstein, 1995; Ramos Rincon and Hernandez 1998 ;Feinstein 2002; Bossuyt 2004) si bien en los últimos años se aprecia una tendencia positiva (Benson 2004). En un intento de superar esta situación, THE STANDARDS FOR REPORTING OF DIAGNOSTIC ACCURACY (STARD) STEERING GROUP ha sistematizado la metodología y desarrollado un cheklist para los autores y revisores de artículos (Bossuyt 2003)

Las revisiones sistemáticas y metanálisis no resuelven el problema sino que añaden otros, derivados de la heterogeneidad de los artículos originales. Esta heterogeneidad se debe a diferencias en el diseño, forma de reclutamiento y tamaño de la muestra, espectro de enfermedad, equipos y técnica usados, definición de positividad de los tests índice (CT y PET), grado de acuerdo interobservador, técnica y grado de exhaustividad de la exploración mediastínica, y lectura ciega e independiente de los tests, entre otros. En el metanálisis es preciso distinguir entre

los conceptos de variabilidad estadística, que puede ser cuantificada y tratada, y la variabilidad clínica, que es más difícil de identificar, y que debe ser examinada en profundidad porque puede quedar oculta por la contundencia de un estimador estadístico (Thompson 1994; Harper, Henson, and Reeves 2000; Lijmer, Bossuyt, and Heisterkamp 2002; Mulherin and Miller 2002).

2.- Revisiones Sistemáticas y metanálisis. Todas las revisiones sistemáticas publicadas sobre CT y PET se basan en un conjunto de 70 artículos distintos, una vez eliminados los duplicados.

Dales et al (1990) analizan 42 artículos sobre el valor diagnóstico de la CT, publicados entre 1980 y 1988 y que suman 3,194 pacientes. El metanálisis calcula una Sensibilidad y Especificidad globales de 0.83 y 0.81. En el subconjunto de 10 artículos que reúnen condiciones más favorables tales como CT de 4ª generación, tiempo de adquisición ≤4 s, grosor de corte ≤10 mm y espaciado ≤ 10 mm, el valor de los índices diagnósticos no se modificó. Los autores concluyen que mejorar estos resultados requeriría una aproximación diagnóstica que no se basara en el tamaño de los ganglios.

Dwamena et al (1999) estudian 29 artículos sobre CT publicados entre 1990 y 1998, que reúnen 2,226 pacientes. Catorce de estos trabajos estudian también la PET, con 514 casos acumulados. Sólo un 64% de los artículos cumple los criterios de calidad técnica de la PET y un 66% de la CT. Sólo un 45% de los artículos hacen lecturas ciegas e independientes de los tests índice y referencia. Un 31% de los estudios son retrospectivos y un 66% no alcanza el tamaño muestral crítico. El metanálisis obtiene para la CT una Sensibilidad y Especificidad media de 0.60 y 0.77y un odds ratio global de diagnóstico de 5.87. Para la PET estos resultados son 0.79, 0.91 y 43.4 respectivamente. Los autores concluyen que la PET es significativamente superior a la CT para la caracterización de los metástasis ganglionares mediastínicas en el CBNM.

Fisher et al (2001) analizan 17 estudios publicados entre 1993 y 2000 en relación con la PET. Los autores no proporcionan un listado de los artículos que utilizan para el análisis de la estadificación, por lo que no es posible cuantificar el número de casos ni conocer ulteriores detalles sobre cada original, ni especifican el método de análisis. Indican que la mayoría son de pobre calidad metodológica y calculan una Sensibilidad y Especificidad globales de 0,83 y 0.96 respectivamente.

Hellwig et al (2001) revisan 19 artículos de CT y 20 de PET, que califican metodológicamente como aceptables, publicados entre 1985-99. Para la estadificación disponen de 12 artículos y 505 pacientes para la CT y de 13 artículos y 529 pacientes para la PET. La sensibilidad y especificidad globales son 0.68 y 0.71 para la CT y 0.88 y 0.92 para la PET. Estos índices se mantienen si el análisis se limita a los 4 artículos que tienen un mínimo de 35 pacientes en los grupos con y sin metástasis ganglionares..

Bradbury I, Bonell E, Boynton J, et al (2002), Health Technology Board for Scotland Update 6/2003. www.htbs.co.uk identifican 32 artículos elegibles publicados entre 1995 y 2001. El metanálisis se realiza con 17 artículos, con 863 pacientes en el estudio. Calculan una especificidad global de 0,76 y una sensibilidad par de 0.65 . En los pacientes CT positivo (15 artículos y 357 casos) la especificidad global calculada de la PET es 0.76 y la sensibilidad global par 0.92 . En los pacientes CT negativo (15 artículos y 547 pacientes) la especificidad global calculada de la PET es 0.90 y la sensibilidad global par 0.86. Los autores concluyen que la PET parece discriminar entre ganglios metastásicos y no metastásicos tanto en ganglios grandes como pequeños.

Toloza et al (2003) revisan 20 artículos sobre CT con 3.438 pacientes y estiman una sensibilidad y especificidad globales de 0,57 y 0,82. El metanálisis de 18 estudios para PET (1.045 pacientes) produce una Sensibilidad y Especificidad globales de 0,84 y 0,89. Los autores no analizan en detalle la calidad de los estudios originales pero aluden expresamente a la marcada heterogeneidad de los estudios individuales.

Gould et al (2003) realizan una revisión sistemática y metanálisis muy detallados sobre el valor de diagnóstico de la CT y la PET con 32 artículos publicados entre 1994 y 2002. La descripción de los tests índice fue en general buena pero sólo el 28% indicó que descartaron pacientes con hiperglucemia y sólo un 32% especificó el uso de CT helicoidal o de un tiempo de adquisición ≤ 2 segundos. Más del 90% utilizó como test de referencia material de biopsia del mediastino pero sólo el 47% requirió toracotomía con muestreo sistemático. En un 44% de los casos las lecturas de la PET y la CT no eran ciegas al diagnóstico final y en más del 50% de los estudios no se interpretaron de forma independiente. La muestra fue reclutada prospectivamente en sólo el 50% de los estudios, y las características de los pacientes fueron incompletamente descritas en la mitad de los estudios. Sólo en el 45% de los artículos, el tamaño muestral era de al menos 35 casos con LN metástasis ó 35 sin ellas. El metanálisis de CT se realizó con 23 estudios (1.119 pacientes) y produjo una especificidad global estimada de 0.79 y una sensibilidad correspondiente de 0.59. En el metanálisis de PET se utilizaron 32 estudios (1.959 pacientes) y produjo una especificidad global estimada de 0.90 y una y una sensibilidad correspondiente de 0.90 es 0.81. En los pacientes CT positivo (12 artículos y 214 casos) la especificidad global estimada de la PET es 0.76 y la sensibilidad par de 0.91. En los pacientes CT negativo (14 artículos y 479 pacientes) la especificidad global estimada de la PET es 0.93 y sensibilidad par de 0.75.

Birim et al (2005) revisan la literatura hasta Enero de 2003 y seleccionan los artículos publicados en lengua inglesa que tienen altos estándares de calidad, de acuerdo al THE STANDARDS FOR REPORTING OD DIAGNOSTIC ACCURACY (STARD) STEERING GROUP (Bossuyt, 2003) y que estudien ambos la TAC y la PET. De los 49 artículos potencialmente relevantes seleccionan 17, que suman un total de 833 pacientes con un rango entre 18 y 102 por artículo. Los 17 artículos

están incluidos en revisiones previas, aunque no todos en todas, todos son prospectivos, y una puntuación de calidad de media 14.1 en una máxima posible de 16. No queda claro hasta que punto las lecturas PET y TAC se hicieron de forma mutuamente independiente y ciega, ni en cuantos pacientes la mediastinoscopia se hizo guiada por la imagen PET. La sensibilidad y especificidad global estimadas de la TAC fueron 0.59 (.50-.67) y 0.78 (.70-.84) y de la PET 0.83 (.77-.87) y 0.92 (.89-.95). Aunque no se comprueba heterogeneidad estadística en los estimadores globales, la Sensibilidad de la TAC se mueve en un rango de 0.20 a 0.81 y la Especificidad entre 0.44 y 1.00; los respectivos rangos para la PET son 0.66-1.00 y .81-1.00.

3.- Estudios observacionales publicados con posterioridad. Una serie de estudios publicados recientemente presentan resultados menos optimistas para la PET. La mayoría de ellos incurren en problemas metodológicos importantes, principalmente por errores sistemáticos de selección y verificación diagnóstica, que, paradójicamente, tienden a sobredimensionar los estimadores de eficacia diagnóstica.

Gonzalez-Stavinski et al (2003) concluyen que los resultados de la PET no confirman ni excluyen la afectación ganglionar mediastínica. **Cerfolio et al (2003)** obtiene valores particularmente bajos, con Sensibilidad y Especificidad estimadas de 43 y 75% para la TAC y 71 y 77% para la PET. **Verhagen et al. (2004)** realizaron un estudio prospectivo en 56 pacientes con CBNM comprobado, sin metástasis extratorácicas. Todos los pacientes sin afectación ganglionar mediastínica recibieron una toracotomía con muestreo sistemático ganglionar mediastínico. La prevalencia de afectación mediastínica fue del 54%, y la sensibilidad y especificidad de 56% (CI 35-72) y 81% (CI 58-91). **Kelly et al (2004)** hicieron un estudio retrospectivo con 69 pacientes que recibieron una TAC y una PET. La sensibilidad de la TAC y PET fue respectivamente de 46 y 62% y la especificidad de 86 y 98%. **Takamochi et al (2005)** estudiaron retrospectivamente 71 pacientes con CBNM que recibieron TAC, PET y estándar de referencia quirúrgico. La sensibilidad para la TAC y la PET fueron 29 y 83% respectivamente, y para la PET 39 y 79%.

Los trabajos de Pieterman (2000) y Pozo-Rodríguez (2005) son particularmente relevantes por sus características metodológicas.

El de **Pieterman et al (2000)**. Este es uno de los estudios más completos, metodológicamente correctos y exhaustivos, y está incorporado en las revisiones sistemáticas de Gould, Bradbury, Toloza y Birim. Se trata de un estudio prospectivo, con una muestra de 102 pacientes, reclutada de forma consecutiva a partir de casos de CBNM potencialmente operables, y un tamaño muestral suficiente. Se definieron con precisión los tests índice, y se utilizaron como estándar de referencia la mediastinoscopia, completada por linfadenectomía radical en los pacientes con cirugía resectiva. Los clínicos tomaron decisiones sin conocer los resultados de la PET. Desafortunadamente no se hizo explícito si las lecturas de PET y TAC fueron mutuamente ciegas

ni independientes respecto de los resultados clínicos, aunque en efecto los clínicos desconocían los resultados de la PET.

El de **Pozo-Rodríguez et al (2005-admitido para publicación en The Journal of Clinical Oncology)** es un estudio prospectivo, con evaluación ciega e independiente de los tests índice (TAC helicoidal y PET) y estándar (mediastinoscopia, y mediastinoscopia ampliada a toracotomía+seguimiento a 6 meses). La muestra de estudio se compone de 132 pacientes consecutivos con diagnóstico seguro de CBNM, potencialmente resecables de acuerdo con los estándares de práctica clínica en el H12X. La información clínica se completó con un seguimiento de 42 meses de mediana. Los elementos fuertes de este estudio son su robustez metodológica, que sigue las líneas de THE STANDARDS FOR REPORTING OD DIAGNOSTIC ACCURACY (STARD) STEERING GROUP (Bossuyt 2003) . La muestra de estudio es una buena representación de la población elegible, las lecturas de las pruebas (TACh y PET) son estrictamente ciegas e independientes de los resultados del test estándar, los decisiones clínicas se toman sin conocer los resultados de las pruebas problema, los radiólogos y los médicos nucleares desconocen los resultados de la PET y la TAC respectivamente, lo que permite estudiar el rendimiento conjunto de ambas pruebas, y el estándar de referencia se ha complementado con el conocimiento del estado clínico y vital de los pacientes durante un tiempo prolongado (mediana de 42 meses). Sin embargo el estudio presenta dos potenciales limitaciones. La primera es que es que el acuerdo interobservador para la lectura e interpretación de TACh y PET es bueno pero no excelente, lo que puede afectar a la generalización de los resultados; la segunda es que no se ha realizado linfadenectomía radical sistemática en las toracotomías resectivas, que puede sesgar los resultados en contra de la TAC y la PET. Sin embargo, comparando nuestros resultados con los de Pieterman (Pieterman 2000) y comprobando los resultados de supervivencia a más de 3 años, consideramos que este potencial sesgo es necesariamente pequeño y de poca significación clínica. En relación con el estándar de referencia MEDIASTINOSCOPIA , la Sensibilidad y Especificidad de la TACh son 0.86 (.70-.93) y 0.67 (.56-.75), de la PET 0.94 (.81-.98) y 0.59 (.49-.68), y de la TACh&PET 0.97 (.84-.99) y 0.44 (.34-.53)

4.- Ensayos Clínicos de diagnóstico.

Van Tinteren et al (2002) abordaron la cuestión de si la FDG-PET puede reducir el número de toracotomías fútiles (TF). La muestra de estudio está compuesta por 186 pacientes con radiología pulmonar sospechosa de malignidad o con carcinoma broncogénico no microcítico (CBNM) confirmado por histología, operables y resecables que se asignaron aleatoriamente a dos grupos: A (96, estadificación estándar) y B (92, estadificación estándar). La frecuencia de TF en el grupo A fue 41% y 21% en el grupo B (el NNT es 5 con IC95% de 3-14. Las limitaciones del estudio proceden principalmente de las características de la muestra, que contiene un espectro de enfermedad que va desde lesiones benignas a estadios muy avanzados de la enfermedad, y que

puede exagerar el valor clínico de la PET.

Viney, Boyer et al (2004) estudiaron el impacto de la PET en los desenlaces clínicos en 183 pacientes con diagnóstico histológico de CBNM, operables oncológica y funcionalmente de acuerdo con las pautas en activo en cada hospital (no incluyen PET), y remitidos al cirujano torácico para toracotomía resectiva IMPORTANTE, entre ellos había, en contra de las especificaciones del protocolo, 4 casos sin confirmación cito-histológica de CBNM. Se excluyen los menores de 18 años y las mujeres embarazadas. No se describe el estudio de operabilidad pero se especifica que todos los pacientes reciben una Rx. de tórax, una TAC de tórax, abdomen superior y cerebro, y una gamma grafía ósea en su caso, y que la mediastinoscopia (Ms-t) no era obligatoria antes de la randomización ("La Ms-t se usa raramente en Australia en pacientes con CBNM en estadios I ó II y que son operables funcionalmente"). Desde una oficina central y con una secuencia generada por ordenador se asignan los pacientes a los grupos de estudio: ECH (n=92) y ECH+PET (n=91). Los 4 pacientes sin confirmación histológica fueron randomizados al grupo ECH+PET. Los autores concluyen con que en pacientes apropiadamente estadificados como estadios I/II la PET puede proporcionar información adicional que mejore la asignación a un determinado tipo de tratamiento pero no conduce necesariamente a una disminución significativa en el número de toracotomías evitable. En nuestra opinión, los resultados del estudio son confusos entre otras cosas porque no se ha utilizado la mediastinoscopia en la estadificación pre-quirúrgica y porque se ha utilizado un protocolo quirúrgico 'agresivo' y distinto a otros estudios. Sin embargo, resulta aceptable 'la música' del estudio, que incide en que en los estadios precoces, con poca carga tumoral, la PET no identifica enfermedad metastásica en la mayoría de los pacientes.

5.- Estudios de evaluación económica. En términos de coste-efectividad y coste-utilidad, ¿cuánto han aportado la PET y la TACh a la estadificación clásica, basada fundamentalmente en protocolos clínicos y mediastinoscopia? En otra lección de este curso (F Pozo-Martín) se aborda el problema desde la perspectiva de la evaluación económica.

5.- Conclusiones. Por el momento, y de acuerdo con la opinión de reconocidos expertos (Kernstine 1999; Line BR White CS 2004), el debate sobre el valor de la PET en la estadificación permanece abierto, y todo apunta a que la PET complementa pero no sustituye a otros métodos de estadificación mediastínica. Habrá que esperar al desarrollo tecnológico de la fusión TAC-PET, la aparición de moléculas marcadoras más específicas, y la aplicación generalizada de técnicas endoscópicas con análisis molecular de las muestras biológicas, para conseguir el ideal de estadificar el CBNM con una gran confianza, en un tiempo muy breve, con la utilización de muy pocas pruebas diagnósticas y a un coste razonable.

Bibliografía

Begg, C. B. "Biases in the assessment of diagnostic tests." Stat.Med. 6.4 (1987): 411-23.

Begg, C. B. and B. J. McNeil. "Assessment of radiologic tests: control of bias and other design considerations." Radiology 167.2 (1988): 565-69.

Benson, T. "Scientific articles have hardly changed in 50 years: ...but Asher was asking why medical journals were so dull back then." BMJ 329.7461 (2004): 352.

Birim O, Kappetein AP Stijnen T Bogers AJ. Meta-analysis of positron emission tomographic and computed tomographic imaging in detecting mediastinal lymph node metastases in nonsmall cell lung cancer. Ann Thorac Surg 79, 375-81. 2005.

Bossuyt, P. M. "The quality of reporting in diagnostic test research: getting better, still not optimal." Clin.Chem. 50 (2004): 465-66.

Bossuyt, P. M. et al. "Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative." BMJ 326.7379 (2003): 41-44.

Bradbury I, Bonell E, Boynton J, Cummins E, Facey K, Iqbal K, Laking G, McDonald C, Parpia T, Sharp P, Single A, Walker A (2002). Positron Emission Tomography (PET) Imaging in Cancer Management. Health Technology Assessment Report 2. Glasgow: Health Technology Board for Scotlandscotland. "scotland." scotlnad (2004).

Cerfolio, R. J. et al. "The role of FDG-PET scan in staging patients with nonsmall cell carcinoma." Ann.Thorac Surg 76.3 (2003): 861-66.

Dales RE, Stark RM Raman S. "Computed tomography to stage lung cancer. Approaching a controversy using meta-analysis." Am Rev Respir Dis 141 (1990): 1096-101.

Dwamena, B. A. et al. "Metastases from non-small cell lung cancer: mediastinal staging in the 1990s--meta-analytic comparison of PET and CT." Radiology 213 (1999): 530-36.

Feinstein, A. R. "Misguided efforts and future challenges for research on "diagnostic tests"." J Epidemiol.Community Health 56 (2002): 330-32.

Goldwin RL, Heitzman ER Proto AV. "Computed tomography of the mediastinum. Normal anatomy and indications for the use of CT." Radiology 124.1 (1977): 235-41.

Gonzalez-Stawinski GV, Lemaire A Merchant F O'Halloran E Coleman RE HarpoleDH D'Amico TA. "A comparative analysis of positron emission tomography and mediastinoscopy in staging non-small cell lung cancer." J Thorac Cardiovasc Surg 126 (2003): 1900-05.

Gould, M. K. et al. "Test performance of positron emission tomography and computed tomography for mediastinal staging in patients with non-small-cell lung cancer: a meta-analysis." Ann.Intern.Med. 139.11 (2003): 879-92.

Harper, R., D. Henson, and B. C. Reeves. "Appraising evaluations of screening/diagnostic tests: the importance of the study populations." Br.J Ophthalmol. 84 (2000): 1198-202.

Hellwig, D. et al. "[Meta-analysis of the efficacy of positron emission tomography with F-18-fluorodeoxyglucose in lung tumors. Basis for discussion of the German Consensus Conference on PET in Oncology 2000]." Pneumologie 55.8 (2001): 367-77.

Kelly, R. F. et al. "Accuracy and cost-effectiveness of [18F]-2-fluoro-deoxy-D-glucose-positron emission tomography scan in potentially resectable non-small cell lung cancer." Chest 125.4 (2004): 1413-23.

Kernstine, K. H. et al. "PET, CT, and MRI with Combidex for mediastinal staging in non-small cell lung carcinoma." Ann.Thorac.Surg. 68 (1999): 1022-28.

Lijmer, J. G., P. M. Bossuyt, and S. H. Heisterkamp. "Exploring sources of heterogeneity in systematic reviews of diagnostic tests." Stat.Med 21 (2002): 1525-37.

Line BR White CS. "Positron Emission Tomography Scanning for the Diagnosis and Management of Lung Cancer." Current Treatment Options in Oncology (BIOMED) 5 (2004): 63-73.

MacManus, M. R. et al. "FDG-PET-detected extracranial metastasis in patients with non-small cell lung cancer undergoing staging for surgery or radical radiotherapy--survival correlates with metastatic disease burden." Acta Oncol. 42.1 (2003): 48-54.

Mulherin, S. A. and W. C. Miller. "Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation." Ann.Intern.Med 137 (2002): 598-602.

Pieterman, R. M. et al. "Preoperative staging of non-small-cell lung cancer with positron-emission tomography." N.Engl.J.Med. 343.4 (2000): 254-61.

Ramos Rincon, J. M. and Aguado Hernandez, I. "[Research on diagnostic tests in Medicina Clinica. A methodological assessment]." Med Clin.(Barc.) 111 (1998): 129-34.

Reid, M. C., M. S. Lachs, and A. R. Feinstein. "Use of methodological standards in diagnostic test research. Getting better but still not good." JAMA 274 (1995): 645-51.

Sheps, S. B. and M. T. Schechter. "The assessment of diagnostic tests. A survey of current medical research." JAMA 252 (1984): 2418-22.

Takamochi K, Yoshida J Murakami K Niho S Ishii G Nishimura M Nishiwaki Y Suzuki K Nagai K. Pitfalls in lymph node staging with positron emission tomography in non-small cell lung cancer patients. Lung Cancer. 47, 235-42. 2005.

Thompson, S. G. "Why sources of heterogeneity in meta-analysis should be investigated." BMJ 309 (1994): 1351-55.

Toloza, E. M., L. Harpole, and D. C. McCrory. "Noninvasive staging of non-small cell lung cancer: a *systematic* review of the current evidence." Chest 123 (2003): 137S-46S.

Van Tinteren, H. et al. "Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial." Lancet 359.9315 (2002): 1388-93.

Verhagen, A. F. et al. "FDG-PET in staging lung cancer; How does it change the algorithm?" Lung Cancer 44.2 (2004): 175-81.

Viney, R. C. et al. "Randomized controlled trial of the role of positron emission tomography in the management of stage I and II non-small-cell lung cancer." J Clin Oncol. 22.12 (2004): 2357-62.

Wahl, R. L. et al. "Staging of mediastinal non-small cell lung cancer with FDG PET, CT, and fusion images: preliminary prospective evaluation." Radiology 191 (1994): 371-77.

Auhor and publication year	Type of study	Years of publication of the OA	Test performed and number of OA	Total patients	Se % (95% IC)	Sp % (95% IC)
Dales 1990	MA	1980-88	CT 42	3194	83 (78-87)	82 (78-85)
Dwamena 1999	MA	1990-98	CT 29	2226	60 (r 25-89)	76 (r 44-95)
			PET 14	514	79 (r 62-97)	91 (r 79-99)
Hellwig 2001	MA	1985-99	CT 12	505	68 (59-70)	76 (73-80)
			PET 13	529	88 (85-95)	92 (91-94)
Toloza 2003	MA	1991-01	CT 20	3438	57 (49-66)	82 (77-86)
			PET 18	1045	84 (78-89)	89 (83-93)
Bradbury 2002	MA	1995-01	CT 12	863	65 (61-69)	76 (70-81)
			PET/CT+ 15	357	92 (87-95)	78 (69-82)
			PET/CT- 15	547	86 (79-91)	90 (87-93)
Gould 2003	MA	1994-02	CT 23	1119	59 (52-66)	79 (iqr 66-89)
			PET 32	1959	81 (74-86)	90 (iqr 82-96)
			PET/CT+ 12	214	91 (79-96)	78 (iqr 68-100)
			PET/CT- 14	479	75 (59-87)	93 (iqr 92-100)
Birim 2005	MA	1994-03	CT	833	59 (50-67)	78 (70-84)
			PET	Same	83 (77-87)	92 (89-95)

Tabla 1. Resultados metanálisis

MA, meta-analysis ; OA, original article ; r, range; iqr, interquartile range; NA, not available.
(1) Weighted mean of sensitivity and specificity (original data were arcasine and squared root transformed; then weighted by the inverse of variance). (2) Mean sensitivities and specificities. Summary ROC curve methods. (3) Estimated sensitivity at the pooled specificity estimate (95% CI); pooled specificity estimate (95% CI). Summary ROC curve methods. (4) Sensitivity at point on Summary ROC curve corresponding to median specificity (and 95% CI); Median specificity (Inter Quartile Range, IQR). Summary ROC curve methods.
PET/CT+: PET results among patients with CT positive
PET/CT-: PET results among patients with CT negative

Tabla 2. resultados artículos recientes no incluidos en las revisiones sistemáticas

Tabla 3. Características de los pacientes estudiados en el HU12X

	Pacients included N=132	Pacientes not included N=62	p-value
Gender (male)	125 (95.5%)	56 (90.3%)	
Age years, median (IQR)	68 (60-71)	67 (61-72)	
Tumor side (right)	75 (57.6%)	38 (61.3%)	
Centrality (central)	59 (44.7%)	19 (30.7)	0.08
Pre-mediastinoscopy stage			
IA	28 (21.2%)	16 (25.8%)	0.41
IB	90 (68.2%)	38 (61.3%)	
IIB	14 (10.6%)	7 (11.3%)	
IIIB	0	1 (1·6)	
Histologic type			
Squamous	67 (50.8%)	28 (45.2%)	0.57
Large cells	37 (28.0%)	20 (32.3%)	
Adenocarcinoma	22 (16.7%)	13 (20.9%)	
Carcinoma	6 (4.5%)	1 (1.6%)	

Auhor and publication year	Type of study	Years of publication of the OA	Test performed and number of OA	Total patients	Se % (95% IC)	Sp % (95% IC)
González-Stav	OA	2003	PET	202	64 (49-76)	77 (70-83)
Cerfolio	OA	2003	CT	400	43 (NA)	75 (NA)
			PET	Same	71 (NA)	77 (NA)
Verhagen	OA	2004	PET	56	56 (35-72)	81 (58-91)
Kelly	OA	2004	CT	69	46 (NA)	86 (NA)
			PET	Same	62 (NA)	98 (NA)
Takamochi	OA	2005	CT	71	29	83
			PET	Same	39	79

Mediastinoscopies	132	62	
Number of stations sampled	964	NA	
Number of stations involved	77	NA	
Cases N2,3 by mediastinoscopy	36 (28%)	14 (22.6%)	0.59
Cases T4central,N0,1	1		
Thoracotomies	92 (69.9%)	35 (56.3)	0.08
Number of stations sampled	253	NA	
Number of stations involved	26	NA	
Cases N2,3	5 (3.7%)	0	
Cases T4central,N0,1	10 (7.5%)	2 (3.2%)	
Definitive stage *			
IA	16 (12.1%)	10 (16.3%)	0.11
IB	42 (31.8%)	26 (41.9%)	
IIA	4 (3.1%)	2 (3.2%)	
IIB	16 (12.1%)	4 (6.5%)	
IIIA	33 (25%)	6 (9.7%)	
IIIB	21 (13.9%)	14 (22.6%)	
Composite Gold Standard	132		
N2-3	41 (31%)		
T4central or N2,3	52 (39.3%)		
Mediastinal disease whithin 6 months	0		

Tabla 4. Estimates of efficacy of hCT and PET for the diagnosis of metastatic mediastinal lymph nodes (95% CI)

INDEX TESTS	SINGLE GOLD STANDARD Ms-t N2,3 N=132			COMPOSITE GOLD STANDARD *Ms-t N2,3 +Toracot N2,3 + 6 m f-up* *N=129*		
	+ve	-ve		+ve	-ve	
hCT Threshold N2,3			Se 0.86 (0.70-0.93) Sp 0.67 (0.56-0.75) PPP 0.49 (0.36-0.62) NPP 0.93 (0.84-0.98) LR+ 2.58 (1.89-3.56) LR- 0.21 (0.09-0.44)			Se 0.83 (0.67-0.91) Sp 0.70 (0.60-0.78) PPP 0.57 (0.43-0.69) NPP 0.90 (0.80-0.96) LR+ 2.81 (1.99-4.03) LR- 0.24 (0.12-0.45)
+ve	31	32		34	26	
-ve	5	64		7	62	
PET Threshold N2,3			Se 0.81 (0.64-0.89) Sp 0.76 (0.66-0.83) PPP 0.56 (0.41-0.70) NPP 0.91 (0.83-0.96) LR+ 3.36 (2.86-5.00) LR- 0.21 (0.09-0.44)			Se 0.80 (0.65-0.89) Sp 0.80 (0.69-0.86) PPP 0.65 (0.50-0.78) NPP 0.90 (0.81-0.95) LR+ 3.94 (2.58-6.16) LR- 0.25 (0.13-0.43)
+ve	29	23		33	18	
-ve	7	73		8	70	

PET Threshold N1c,2,3			Se	0.94 (0.81-0.98)		Se 0.95 (0.82-0.98)
			Sp	0.59 (0.49-0.68)		Sp 0.63 (0.52-0.71)
			PPP	0.47 (0.35-0.59)		PPP 0.54 (0.42-0.66)
+ve	34	39	NPP	0.97 (0.88-0.99)	39	33 NPP 0.96 (0.88-0.99)
			LR+	2.32 (1.82-3.04)		LR+ 2.54 (1.95-3.40)
-ve	2	57	LR-	0.09 (0.03-0.31)	2	55 LR- 0.08 (0.02-0.26)

Tabla 5. Estimates of the efficacy of hCT&PET combined in parallel for the diagnosis of metastatic mediastinal lymph nodes (95% CI).

INDEX TEST	**SINGLE GOLD STANDARD Ms-t N2,3 N=132**				**COMPOSITE GOLD STANDARD Ms-t N2,3 +Thoracot N2,3 + 6 m f-up** *N=129*			
	+ve	**-ve**			**+ve**	**-ve**		
hCT&PET Threshold N2,3			Se	0.92 (0.76-0.96)			Se	0.90 (0.76-.0.95)
			Sp	0.55 (0.45-0.64)			Sp	0.59 (0.48-0.68)
			PPP	0.43 (0.32-0.55)			PPP	0.49 (0.36-0.62)
+ve	33	43	NPP	0.95 (0.85-0.98)	37	36	NPP	0.93 (0.83-0.98)
			LR+	2.05 (1.60-2.64)			LR+	2.21 (1.69-2.93)
-ve	3	53	LR-	0.15 (0.05-0.40)	4	52	LR-	0.17 (0.009-0.27)
hCT&PET Threshold N1c,2,3			Se	0.97 (0.84-0.99)			Se	0.98 (0.86-0.99)
			Sp	0.44 (0.34-0.53)			Sp	0.48 (0.37-0.58)
			PPP	0.39 (0.29-0.50)			PPP	0.51 (0.39-0.63)
+ve	35	54	NPP	0.98 (0.88-1.00)	40	46	NPP	0.98 (0.88-1.00)
			LR+	1.73 (1.44-2.11)			LR+	1.87 (1.54-2.33)
-ve	1	42	LR-	0.06 (0.01-0.33)	1	42	LR-	.05 (0.009-0.27)

Mediastinal staging with Positron-Emission Tomography and Contrast-Enhanced Spiral Computer Tomography in the clinical management of potentially operable lung cancer patients: an economic evaluation.

Francisco Pozo-Martin[1], Jack Dowie[2], Francisco Pozo-Rodriguez[3] JL Luis Martin de Nicolas[3], Maria Antonia Sanchez-Nistal[3], y Antonio Maldonado[4]

Within the TNM system, mediastinal staging is performed using invasive and non-invasive staging procedures: mediastinoscopy (Ms-t) and chest Computed Tomography (CT). Both techniques are somewhat inaccurate. Inaccurate N staging triggers the undesired effect of indicating surgery in patients that have undetected locally advanced disease (N2 or N3 disease). Conversely, the staging of resectable patients (N0 or N1 disease) as patients with mediastinal lymph node metastasis (N2 or N3) will deny them the possibility of surgical cure. Thus, the main goal of mediastinal staging procedure is to minimize the number of false positive and false negative patients.

Including PET in any staging strategy for staging NSCLC requires the availability of sophisticated equipment and facilities that substantially increase healthcare costs. However, these costs may be offset by a better selection of patients for beneficial resective surgery, increased life expectancy and better quality of life. Thus, assessing the cost-effectiveness of PET in the staging and treatment of NSCLC patients is of primary importance. Several attempts have been made to evaluate the cost-effectiveness of PET in lung cancer staging (see section 2.- of this report), but to the knowledge of this author there have been no attempts at evaluating **both** the **actual costs** (rather than charges) to the healthcare system and the consequences, **quality of life**, of adding PET to the standard staging strategy of NSCLC patients.

This report brings together high quality clinical data from a sample of NSCLC patients in the Hospital 12 de Octubre (Madrid, Spain), physician estimates of quality-adjusted life expectancy for those patients, and detailed hospital cost estimation to perform a cost-utility analysis, from the viewpoint of hospital, of the following 3 strategies for the clinical management of NSCLC: 1) The standard clinical management, which uses Ms-t as the *sole* mediastinal staging procedure; 2) A clinical management in which performing Ms-t is conditioned to the results of spiral CT. CT+ve patients would be reassessed with Ms-t, CT-ve patients would go straight to resective surgery, and 3) A clinical management in which performing Ms-t is conditioned to the results of both spiral CT and PET. Patients for which any or both CT and PET are +ve would be reassessed with Ms-t. Patients for which both CT and PET are -ve would go straight to resective surgery.

In order to to this, the authors of this report will model the alternative clinical management strategies, their costs and outcomes (quality-adjusted life expectancy) into a decision tree and will

compare them in terms of incremental cost-effectiveness. Sensitivity analysis will be performed to test the robustness of the model to changes in uncertain parameters.

Evidence on the cost-effectiveness of PET from the medical literature: a literature review. Systematic review of the medical literature with the following research question: What is the cost-effectiveness of clinical management strategies involving Positron-Emission Tomography (PET) versus other non-PET strategies in the staging of patients diagnosed with lung cancer?

1.- Aurum Health Research, Johannesburg, Republic of South Africa. 2.- London School of Public Health, UK. 3.- Hospital 12 de Octubre, Madrid. 4.- Centro PET Complutense, Madrid

Clinical management strategies for staging NSCLC. In the studies identified, standard clinical management of NSCLC patients does not include PET as a staging procedure as seen in **Table 1**. However it is defined, standard clinical management is the baseline comparator in all the cost-effectiveness analyses that were reviewed. The alternative strategies, intervention strategies, all include PET as an add-on staging procedure, never as free-standing. These intervention strategies are reported in **Table 2.**

Are PET-including intervention strategies more EFFECTIVE than standard practice in the clinical management of NSCLC patients? See Table 3**. A sounded discussion on this issue will be provided in this meeting by F Pozo-Rodríguez et al.**

Are PET-including intervention strategies COST-EFFECTIVE compared to standard practice in the clinical management of NSCLC patients? Evidence from the literature review.

Table 4 shows the ICERs (when reported) for the alternative PET-including clinical management strategies of the four studies in this review that use Life Expectancy (LE) as health outcome in their cost-effectiveness analyses (cite). **In conclusion,** even though results have to be handled with much care, there seems to be some evidence from the literature that including PET in the staging and treatment of NSCLC patients is cost-effective relative to alternative strategies at acceptable Willingness to Pay (WTP) per Quqlity Adjusted Life Year (QALY). Particularly, in relation to some alternative interventions. For example, when PET is used after CT only in CT-ve patients[13,14,15]. Also, when Ms-t is used in those patients where PET and CT results contradict each other (i.e. in

patients that have both CT+ve and PET-ve results, and in patients that have CT-ve and PET+ve results)[2].

The "Hospital 12 de Octubre Study".

1.- Aim: **to perform a cost-utility analysis of 3 clinical management strategies for NSCLC patients: 1) Standard clinical management, which uses Ms-t as the sole mediastinal staging procedure. 2) clinical management in which performing Ms-t is conditioned to the results of spiral CT. 3) clinical management in which performing Ms-t is conditioned to the results of both spiral CT and PET.**

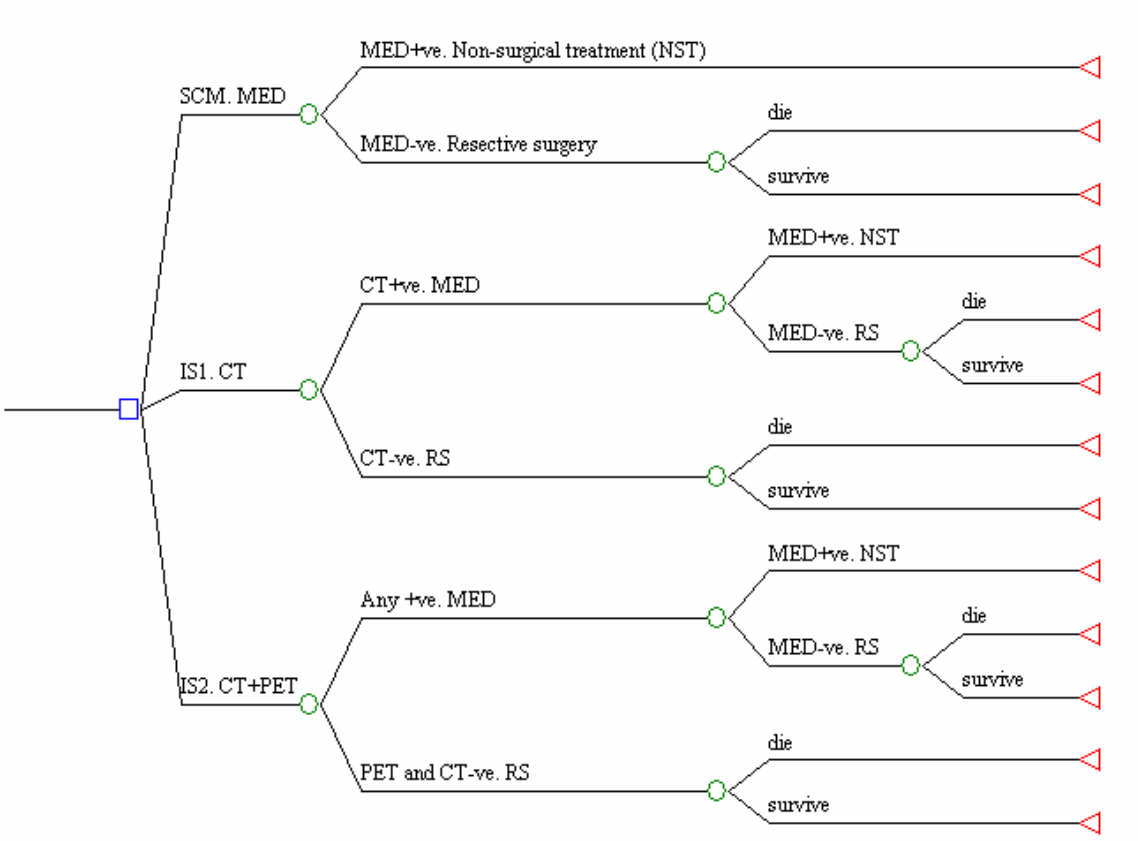
2.- Specific objectives. **2.1.- To measure the quality-adjusted life expectancy (QALE), and to measure the costs of NSCLC patients under 3 alternative clinical management strategies. 2.2.- To calculate the incremental cost-effectiveness ratios (ICERs) for the different strategies. 2.3.- To undertake a sensitivity analysis of key uncertain parameters. 2.4.- To discuss the implications for resource allocation**

3.- Methods.

3.1.- Study design. Described elsewhere (F Pozo-Rodríguez et al)

3.2.- Alternative Clinical Strategies. *Standard clinical management (SCM):* Mediastinal staging using Ms-t alone followed by non-surgical treatment (chemotherapy and/or radiotherapy) in Ms-t+ve patients and resective surgery (with or without concurrent chemotherapy/radiotherapy) in MS-T-ve patients. *Intervention strategy 1 (IS1) :* Mediastinal staging using spiral CT followed by straightforward resective surgery (with or without concurrent chemotherapy/ radiotherapy) in CT-ve patients and Ms-t in CT+ve patients (MS-T+ve patients go to non-surgical treatment, Ms-t-ve patients go to resective surgery with or without concurrent chemotherapy/radiotherapy). *Intervention strategy 2(IS2):* Mediastinal staging using spiral CT followed by PET in all patients. Ms-t is performed if either PET or CT (or both) are +ve. When results of both CT and PET are –ve, the patient is referred for straightforward resective surgery with or without concurrent chemotherapy/radiotherapy.

3.3.- Decision analysis was used to assess which of the 3 strategies (*SCM, IS1, IS2*) is optimal, given 1) the number of patients diagnosed with N2/N3 disease and the number of patients diagnosed with N0/N1 disease under each strategy, 2) the consequences in terms of QALE accruing to all patients under each strategy, 3) the actual costs borne by the hospital under each strategy, and 4) a given willingness to pay (WTP) of 50,000 EUR per QALY gained. Diagram 1 shows the conceptual outline of this study design. The decision tree model is shown below



3.4.- Data was collected for the following variables before transforming them into the appropriate variables for decision tree analysis:

3.4.1.- Number of N2/N3+ve and N2/N3-ve patients with each staging procedure, and 2) number of patients that die in surgery for each clinical management strategy. This data came from the follow-up study (February 1999-February 2001) of 132 patients perforMs-t in hospital H. The study database was analysed to extract clinical information on 1) and 2).

3.4.2.- Quality-adjusted life expectancy (QALE) accruing to all patients. Expected QALYs were calculated for all 132 patients on the basis of assigning each patient 1) an initial health state, 2) an individual survival / life expectancy according to its pathological or post Ms-t stage, and 3) individualized changes in quality of life derived from the treatment and complications that each patient underwent. QALY calculations were divided into four steps:

3.4.2.1.- Step1. Patient sample QALY calculations. Four initial health state values, associated with 4 types of NSCLC patients, were elicited by a lung specialist, using the EuroQol EQ-5D descriptive instrument (see Table 5). Each of all 132 patients was assigned to one of the four types in order to obtain their quality of life status at the beginning of their treatment. This five-dimensional quality of life description was translated into a number between 0 and 1 (where 0 is the worst possible health state and 1 is perfect health) using the Spanish EuroQol tariff.

3.4.2.2.- Step 2. Patient sample QALY calculations. Duration of life was evaluated for every patient, based on either 1) actual survival if the patient had already died, or 2) survival data adjusted for stage, estimated from the Spanish Lung Cancer Cooperative Group if the patient was not dead at the end of December 2002.

3.4.2.3.- Step 3. Patient sample QALY calculations. Beginning with the initial health state, the expected change in quality of life along life expectancy (health profile) was modelled for 5 types of patient. These 5 health profiles were developed taking into account 1) the treatment each patient underwent, and 2) the complications experienced by each patient. Profiles and changes in quality of life for each profile were based in previous results from the literature and in the judgement of an expert oncologist. The different profiles modelled are displayed in **Table 6.** A simple algebraic

formula was used to calculate the area under each health profile, given in **Table 7**. The resulting value is the number of expected QALYs accruing to each health profile.

3.4.2.4.- Step 4. Patient sample QALY calculations. QALYs accruing to each patient at the end of each branch of the decision tree were calculated on an EXCEL spreadsheet by assigning each patient the profile corresponding to his treatment/complications after surgery (profile 1, 2, 3, 4, or 5), and entering into the algebraic formula corresponding to his profile his survival time/LE and the initial health state belonging to his type (type 1, 2, 3, 4). See appendix 3 for an exhaustive list of the values accruing to each patient at each branch of the tree.

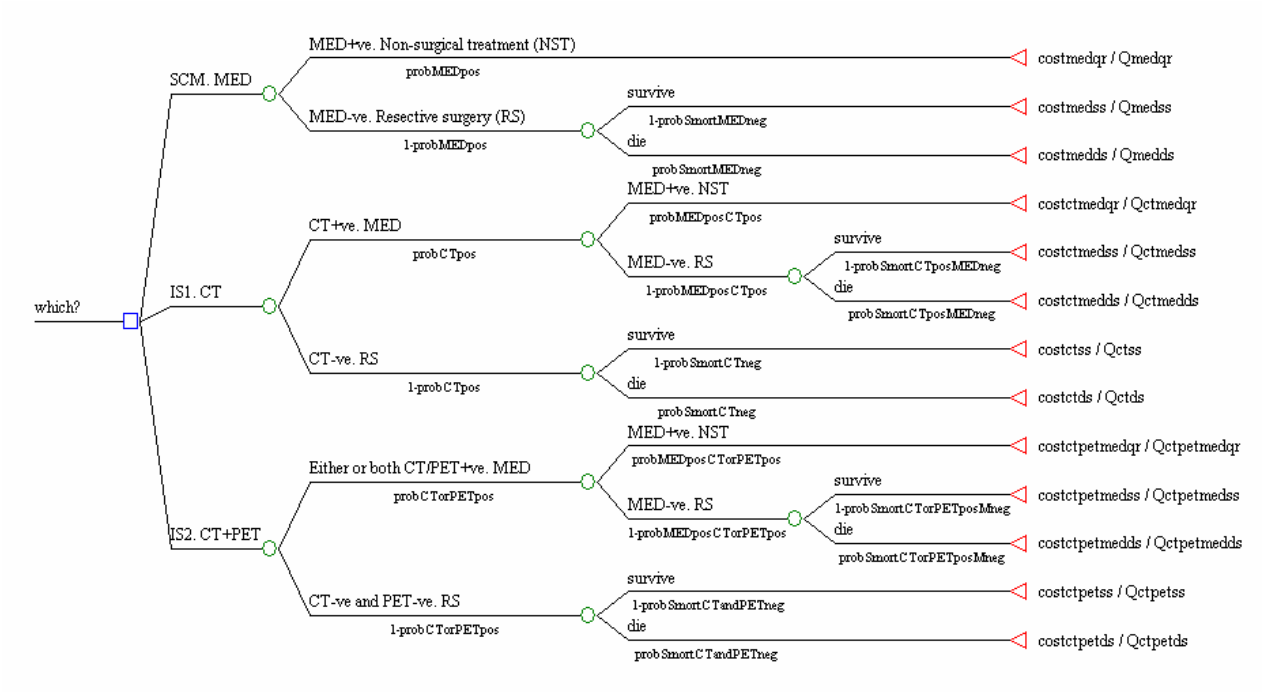
3.4.3.- Total costs for each staging procedure (PET, CT, Ms-t) and for each treatment (resective surgery, chemotherapy, radiotherapy). In order to calculate total costs for each procedure and each treatment, detailed resource use for an average (i.e. standard) procedure and treatment, as well as unit costs, were obtained directly from the Hospital 12 de Octubre. The cost of each PET scan used in the baseline cost-effectiveness analysis was not calculated in detail because Hospital 12 de Octubre purchases each PET procedure from the Centro PET Complutense, which is a private company. Since this study was conducted from the perspective of the hospital (provider organization), the baseline cost of each PET that was taken into account in this study is the price at which it was bought from PET Center. **Table 8** shows the information obtained on resource use and the sources from which that information was obtained. **Table 9** shows the sources of information for unit costs. Total costs of each procedure were calculated on an EXCEL spreadsheet by multiplying each unit of average resource consumption by its unitary cost. Overhead costs and capital costs were allocated across treatments and procedures (see **Table 10**). The price of PET and the detailed breakdown of average costs for each of the other 2 staging proceduresprocedure (MS-T/CT), as well as for each treatment (resective surgery, chemotherapy, radiotherapy) are shown in **Appendix 3.**

3.5.- Patient sample. The patient sample on which this study is based in a prospective series of 132 consecutive cases attended at Hospital 12 de Octubre between February 1999 and February 2001. Inclusion criteria: histologically diagnosed NSCLC, no distant metastases (M1 category of the TNM classification), informed consent. Non-inclusion criteria: any anatomic alteration that precludes Mediastinoscopy, pregnancy, and diabetes.

4.- Data analysis.

4.1.- Variable transformations for input in the decision tree. Data collected from the patient sample was transformed into 3 types of variables for input in the decision tree model: probabilities at each chance node, QALYs at each terminal node, and costs at each terminal node. See Tables 11-13.

4.2.- Cost-effectiveness analysis. DATA 4.0 software was used to build the decision tree model (see Diagram 3 on the next page for its full display), and to calculate incremental cost-effectiveness ratios (ICERs) in order to evaluate the most cost-effective strategy. One-way sensitivity analysis to evaluate the impact on ICERs of changes in uncertain parameters was conducted for each probability, each outcome and each cost. Also, an additional model was run using mean LE instead of mean QALE as health outcome. The decision tree model is shown below



5.- Results.

5.1.- Results of the model with baseline values

Table 11 shows the baseline values (range for sensitivity analysis) for all the variables in the decision tree model (units of measurement are QALYs for Q-variables and EUR for cost-variables). Table 12 displays the results of cost-effectiveness analysis.

Standard clinical management (strategy SCM) is the cheapest strategy (cost = 7843 EUR per patient) and it accrues a little over 2 QALYs per patient. The next cheapest alternative, Intervention Strategy 1 (IS1, i.e. staging using spiral CT+MS-T instead of MS-T alone) has an incremental cost of 796 EUR per patient. QALE gain using IS1 is 0.028 QALYs. ICER is 28,904 EUR per QALY gained. For a willingness-to-pay threshold of 50,000 EUR/QALY, this strategy is cost-effective. The most expensive clinical management strategy is IS2, which uses CT, PET and MS-T as staging procedures. Compared to IS1, it costs 544 extra EUR per patient, and it brings about a loss in QALE of 0.084 QALYs. It is a dominated strategy (i.e. both more expensive and less effective), both in relation to SCM and IS1.

5.2.- Results of the sensitivity analysis. One-way sensitivity analysis was perforMs-t on all probabilities, QALYs and costs using the ranges outlined in Section 4.4. Results are shown in Table 13. In a second analysis,a new model was run using LE instead of QALE as final outcome measure.

The results of the baseline model are robust to changes of up to 50% in the cost of each staging and treatment procedures. Strategy IS2 is always dominated by both IS1 and SCM. The minimum/maximum ICERs of IS1 versus SCM (18,336 EUR/QALY and 39,471 EUR/QALY) are in a range determined by changes in the cost of CT. In all cases, for a WTP of 50,000 EUR per QALY gain, IS1 is cost-effective.

Results of the baseline model are also robust to changes in the QALE of the patients in the sample. With the exception of 3 scenarios, IS2 is systematically dominated and IS1 is cost-effective (WTP threshold = 50,000 EUR/QALY) compared to SCM. ICERs of IS1 versus SCM move in a range between 1,442 EUR/QALY (for the maximum QALE accruing to patients found N2/N3+ve by CT and N2/N3-ve by MS-T, under strategy IS1) and 28,904 EUR/QALY (for the minimum QALYs accruing under any of the 3 scenarios under strategy IS2). However, IS2, the PET-including strategy, becomes cost-effective for the maximum QALE accruing to all patients under this strategy. Maximum QALE is 1.6 QALYs for patients referred straight to NST (but only 1 individual reaches 1 QALY

before dying), 4.33 QALYs for patients who are CT/PET+ve and MS-T-ve and survive surgery (40% of individuals reach more than 3.5 QALYs), and 6.2 QALYs for patients CT/PET-ve that survive surgery (around 10% of all patients reach more than 5 QALYs).

Results are least robust to changes in the tree probabilities. Again, IS2 is systematically dominated, except in one case. IS2 is cost-effective compared to SCM when the probability that PET and/or CT find N2/N3 disease is at its lowest. That is, when PET and CT refer the least patients to non-surgical treatment, by far the most expensive treatment. Across scenarios, IS1 changes from being cost-effective to being dominated by SCM. The small incremental costs and incremental effectiveness at baseline change substantially when the tree probabilities are changed.

When QALE is changed in favour of LE as the outcome of the decision tree model, effectiveness results change enough to suggest that, at baseline, IS1 is no longer cost-effective compared to SCM. Incremental cost of IS1 is still 796 EUR, but QALE gain is now negative (-0.012 LYs). In fact, both IS1 and IS2 become dominated by SCM under this scenario.

BIBLIOGRAFÍA

Gambhir SS, Hoh CK, Phelps ME, Madar I, Maddahi J. Decision Tree Sensitivity Analysis for Cost-Effectiveness of FDG-PET in the Staging and Management of Non-Small Cell Lung Carcinoma. J Nucl Ms-t 1996; 37: 1428-1436.

Stroobants S, Verschakelen J, Vansteenkiste J (2003). Value of FDG-PET in the Management of Non-Small Cell Lung Cancer. Eur J Radiol 2003; 45: 49-59.

Webb W, Gatsonis C, Zerhouni E et al. CT and MR Imaging in Staging Non-Small Cell Bronchogenic Carcinoma: Report of the Radiologic Diagnostic Oncology Group. Radiol 1991; 178: 705-713.

Pieterman, RM, et al. Pre-operative Staging of Non-small Cell Lung Cancer with Positron-Emission Tomography. N-Engl-J-Ms-t. 2000 343: 254-61.

Van Tinteren et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial Lancet. 2002; 359:1388-93

G J Herder, H van Tinteren, E F Comans, O S Hoekstra, G J Teule, P E Postmus, U Joshi, and E F Smit. Prospective use of Serial Questionnaires to Evaluate the Therapeutic Efficacy of 18F-fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) in Suspected Lung Cancer. Thorax 2003; 58: 47-51.

Dwamena.

Drummond MF, O'Brien B, Stoddart G, Torrance GW (1997). Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press. Chapter 3.

Verboom P, Herder G, Hoekstra OS, Smit EF, van den Bergh J, Grijseels EW. Staging of Non-Small Cell Lung Cancer and Application of FDG-PET: a Cost Modelling Approach. Int J Technol Assess Health Care 2002; 18: 576-585.

Kosuda S, Ichihara K, Watanabe M, Kobayashi H, Kusano S. Decision Tree Sensitivity Analysis for Cost-effectiveness of Chest 2-Fluoro-2-D(18F) Positron-Emission Tomography in Patients with Pulmonary Nodules (Non-small cell carcinoma) in Japan. Chest 2000; 117: 346-353

Valk PE, Pounds TR, Tesar RD, Hopkins DM, Haseman MK. Cost-Effectiveness of PET Imaging in Clinical Oncology. Nuc Ms-t Biol 1996; 23: 737-743.

Dietlein M, Weber K, Gandjour A, Theissen P, Lauterbach KW, Schicha H.
Cost-effectiveness of FDG-PET for the Management of Potentially Operable Non-Small Cell Lung Cancer: Priority for a PET-based Strategy After Nodal-negative CT Results. Eur J Nucl Ms-t 2000; 27: 1598-1609.

Scott WJ, Shepherd J, Gambhir SS. Cost-effectiveness of FDG-PET for Staging Non-Small Cell Lung Cancer: a Decision Analysis. Ann Thorac Surg 1998; 66: 1876-1883.

Bradbury I, Bonell E, Boynton J, Cummins E, Facey K, Iqbal K, Laking G, McDonald C, Parpia T, Sharp P, Single A, Walker A (2002). Positron Emission Tomography (PET) Imaging in Cancer Management. Health Technology Assessment Report 2. Glasgow: Health Technology Board for Scotland.

von Schultness GK, Steinert HC, Dummer R, Weder W. Cost-Effectiveness of Whole-body PET Imaging in Non-Small Cell Lung Cancer and Malignant Melanoma. Acad Radiol 1998; 5 (suppl): s300-s302.

Laupacis A, Feeny D, Detsky AS, Tugwell PX. How Attractive does a new Technology have to be to Warrant Adoption and Utilization? Tentative Guidelines for Using Clinical and Economic Evaluations.Can Ms-t Assoc J 1992; 146 (4): 473-81

Badia X. EuroQoL EQ-5D. Guia del Usuario. Version Espanola. 1999.

Table 1. Standard clinical management strategies. Results from the systematic review

Reference(s)	Standard clinical management strategy (baseline comparator)
Kosuda et al Dietlein et al Bradbury et al Verboom et al Valk et al	**Standard strategy 1 (S1).-** TNM staging (mediastinal staging in Dietlein et al) using CT, followed by mediastinoscopy (Ms-t) in all patients independently of the result of CT. Ms-t-ve patients (usually patients with no abnormal mediastinal lymph nodes > 10mm, although this is not specified in many articles!) go for tumour resection (thorachotomy) and Ms-t+ve patients are referred to non-surgical treatment (chemotherapy and/or radiotherapy)
Scott et al Gambhir et al	**Standard strategy 2 (S2).-** Mediastinal staging using CT, followed by Ms-t only in CT+ve patients. CT-ve patients are referred straight to surgery. Of all CT+ve patients that undergo mediastinoscopy, MS-T-ve patients are referred to surgery and MS-T+ve patients are referred to non-surgical treatment
Von Schultness et al	**Standard strategy 3 (S3).-** TNM staging using CT and bone scanning

Table 2. Intervention clinical management strategies. Results from the systematic review.

Reference(s)	PET-including clinical management strategy (intervention).
Dietlein et al Bradbury et al	**Intervention strategy 1 (I1).-** TNM staging (Mediastinal staging in Dietlein et al) with CT followed by PET in all patients. Independently of the results of CT, PET-ve patients go straight to surgery. PET+ve patients go all to non-surgical treatment.
Scott et al Dietlein et al Bradbury et al	**Intervention strategy 2 (I2).-** TNM staging (mediastinal staging in Dietlein et al) with CT followed by PET in all patients. Independently of the results of CT, PET-ve patients go straight to surgery and PET+ve patients undergo Ms-t, followed by surgery in Ms-t-ve patients and non-surgical treatment in Ms-t+ve patients (Scott, Dietlein). One variation of this strategy sends PET N0/N1 M1 patients straight to non-surgical treatment.
Dietlein et al	**Intervention strategy 3 (I3).**- TNM staging (Mediastinal staging in Dietlein et al) with CT followed by PET in CT-ve patients only. CT+ve patients go straight to non-surgical treatment. CT-ve patients are referred to non-surgical treatment if PET is +ve, and to surgery if PET is –ve.
Scott el al	**Intervention strategy 4 (I4).-** Mediastinal staging with CT followed by PET in CT-ve patients only. CT+ve patients go straight to non-surgical treatment. CT-ve patients are referred to Ms-t if PET is +ve, followed by non-surgical treatment in Ms-t+ve patients and surgery in Ms-t-ve patients. If PET is –ve, patients are referred straight to surgery.
Dietlein et al	**Intervention strategy 5 (I5).-** Mediastinal staging with CT followed by PET in all patients. In patients that are both CT-ve and PET+ve, Ms-t is performed followed by surgery in Ms-t-ve patients and non-surgical treatment in Ms-t+ve patients. For all other categories, independently of CT results, PET+ve patients undergo non-surgical treatment and PET-ve patients go to surgery.
Bradbury et al	**Intervention strategy 6 (I6).-** TNM staging with CT followed by PET in all patients. Patients that are both CT+ve and PET-ve undergo MS-T (aswell as patients that are both CT-ve and PET+ve), followed by surgery in Ms-t-ve patients and non-surgical treatment in Ms-t+ve patients.
Scott et al	**Intervention strategy 7 (I7).-** Mediastinal staging with CT followed by PET in all patients. MS-T is perforMs-t on all patients (followed by non-surgical treatment or surgery) except those that are both CT-ve and PET-ve, who are referred straight to surgery.
Gambhir et al	**Intervention strategy 8 (I8).-** Mediastinal staging using PET followed by CT in all patients. PET+ve patients all undergo Ms-t independently of CT results, followed by non-surgical treatment for Ms-t+ve and surgical treatment for Ms-t-ve patients.
Gambhir et al	**Intervention strategy 9 (I9).-** Mediastinal staging using PET followed by CT in all patients. Patients both PET-ve and CT-ve go straight to surgery. Patients that are both PET-ve and CT+ve (or PET+ve and CT-ve) go to MS-T (followed by non-surgical treatment if Ms-t+ve or surgery if Ms-t-ve). Patients both CT+ve and PET+ve go straight to non-surgical treatment.
Bradbury et al	**Intervention strategy 10 (I10).-** TNM staging with CT followed by MS-T in all patients. Ms-t-ve patients go to PET, followed by surgery if PET-ve and non-surgical treatment if PET+ve. Ms-t+ve patients are referred straight to non-surgical treatment.
Von Schultness et al	**Intervention strategy 11 (I11).-** PET and CT only if PET indicates operable disease (no further details are given).
Verboom et al	**Intervention strategy 12 (I12).-** Conventional work-up and PET (no further details are given).

Table 3. Evidence for increased effectiveness using PET in the clinical management of NSCLC patients. Results from the literature review.

Study	**Source of effectiveness data**	**Outcomes with each strategy**	**Comments regarding evidence for increased effectiveness.**
Verboom et al	Clinical trial (n=188)	Futile surgery (I12) = 21% Futile surgery (S1) = 41% (p=0.003)	Strong evidence.
Valk et al	Clinical series (n=99)	15 thorachotomies avoided fue to PET finding 6 non-surgical treatments avoided due to PET findings	No p-values or C.I.s given. No sample size necessary for detection of significant effect reported.
Von Schultness et al	2 studies in the literature	- PET staging was accurate, in detecting N3 in 4 of 65 patients, which were understaged as N2 disease with CT. - Unsuspected metastases were identified with PET in 14 out of 95 patients (…) 9 of those 14 patients were initially diagnosed as N0/N1 or N2 disease (…) Only with the PET was the disease recognised as inoperable.	Weak evidence due to: - insufficient information on statistical significance of results provided. - Small sample sizes with no power calculations.
Dietlein et al	Literature review	LE under each strategy: CT+M (S1): 3.308 yrs CT+PET (I3): 3.322 yrs CT+PET+M (I2): 3.328 yrs CT+PET+M (I5): 3.282 yrs CT+PET (I1): 3.255 yrs Max LE gain= LE(I2)-LE(S1)= 0.02 yrs	Weak evidence (Max LE gain = 1 week.) Patients included all had normal size Ms-tiastinal lymph nodes (which excludes 75% of patients with NSCLC). Thus, limited external validity.
Scott et al	Literature review	LE under each strategy: CT+ M (S2): 4.921 yrs CT+M+PET (I4): 4.928 yrs CT+M+PET (I7): 4.930 yrs CT+M+PET (I2): 4.928 yrs Max LE gain= LE(I7)-LE(S2)= 0.009 yrs	Weak evidence (Max LE gain < 1 week.)
Gambhir et al	Literature review	LE gain (I8 – S2) = 2.96 days	Weak evidence
Kosuda et al	Literature review and clinical series (n=56)	LE under each strategy: CT alone (S3): 10.33 yrs/patient CT+ FDG PET (I11): 10.94 yrs/patient LE gain (I11-S3): 0.61 yrs	No external validity: - Standard practice in this study not used anywhere else. - 30% patients included not NSCLC patients
Bradbury et al	Literature review	CT+ve patients: CT+M (S1): 71.86 QALYs CT+M+PET (I10): 66.03 QALYs CT+PET+M (I6): 66.16 QALYs CT+PET+M (I2): 72.11 QALYs CT+PET (I1): 66.17 QALYs QALE (I2)-QALE (S1)= 0.15 QALYs CT-ve patients: CT+M (S1): 189.63 QALYs CT+PET+M (I2): 190.96 QALYs CT+PET (I1): 181.97 QALYs CT+PET +M (I6): 181.39 QALYs CT+M+PET (I10): 181.28 QALYs QALE (I2)-QALE(S1)= 1.33 QALYs	CT+ve patients: QALYs gained with most effective intervention strategy in a cohort of 100 patients = 0.15 QALYs CT-ve patients: QALYs gained with most effective intervention strategy in a cohort of 100 patients = 1.33 QALYs Weak evidence of increased effectiveness.

Table 4. Incremental cost-effectiveness ratios from the literature review. Outcome measure: Life Expectancy (LE).

Study	**ICER**
Dietlein et al	ICER (I3-S1) = 142 EUR per life year saved ICER (I2-S1) = 11,100 EUR per lys ICER (I2-I3) = 36,667 EUR per lys ICER (I5-I3) = 15,325 EUR per lys ICER (I5-I2) = 18,109 EUR per lys ICER (I1-I3) = 15,716 EUR per lys ICER (I1-I2) = 17,438 EUR per lys
Scott et al	ICER (I4-S2)= $25,286 per lys ICER (I7-I4) = $70,889 per lys ICER (I2-I7) = $137,857 per lys.
Gambhir et al	Not reported
Kosuda et al	Not reported

Table 5. Patient types for which initial health state valuations were elicited using EQ-5D (Spanish tariff).

Type 1	Type 2	Type 3	Type 4
Patients with a weight loss of at least 10 kg prior to Ms-tiastinal staging and one or more of the following symptoms: hemoptysis persistent cough difficulty breathing chest pain fever	Patients with a weight loss of at least 10 kg prior to Ms-tiastinal staging and none of the following symptoms: hemoptysis persistent cough difficulty breathing chest pain fever	Patients with no weight loss prior to Ms-tiastinal staging and one or more of the following symptoms: Hemoptysis persistent cough difficulty breathing chest pain fever	Patients with no weight loss prior to Ms-tiastinal staging and none of the following symptoms: hemoptysis persistent cough difficulty breathing chest pain fever

Table 6. Quality of life profiles modelled for each treatment.

Patient referred to surgery (N0/N1 disease)	Patient referred to non-surgical treatment (N2/N3 disease)
Profile 1: no surgical complications, no chemotherapy/radiotherapy	Profile 5. Patient treated with chemotherapy and radiotherapy
Profile 2: no surgical complications, patient given chemotherapy and radiotherapy	
Profile 3: surgical complications, no chemotherapy/radiotherapy	
Profile 4: surgical complications, patient given chemotherapy and radiotherapy	

Table 7. Quality of life profiles developed for each treatment. Source: REFERENCE and expert opinion.

QOL profiles	Estimated change in QOL	Algebraic formula
Profile 1	Surgery induces a fall in QOL of 0.05 from the initial health state (Uo) during the first month. QOL remains stable for the rest of the time except the last 6 months of life, during which QOL drops to 0.2.	QALYs (Profile 1): 1/12 x (Uo+(Uo-0.05)x(LE-7)+(0.2x6))
Profile 2	Same as Profile 1 except that chemotherapy /radiotherapy induce a drop in QOL of 0.4 during the 4 months of non-surgical treatment. After non-surgical treatment, QOL rises by .4 until the last 6 months of life, in which QOL drops to 0.2	QALYs (Profile 2): QALYs (Profile 1) – 4/12x0.4
Profile 3	Same as Profile 1 except that complications induce a drop in QOL to 0 during the first month after surgery	QALYs (Profile 3) 1/12 x (Uo+(Uo-0.05)x(LE-6)+(0.2x6))
Profile 4	Same as Profile 3 with the addition of changes in QOL produced by Profile 2	QALYs (Profile 4): QALYs (Profile 3)- 4/12x0.4
Profile 5	Chemotherapy induces a fall in QOL of 0.4 after 1 month in the initial health state, and remains stable for the 4 months of non-surgical treatment. It then rises by 0.2 until the last 6 months of life, during which QOL drops to 0.2	QALYs (Profile 5): 1/12x(Uo+4x(Uo-0.4)+(Uo-0.2)x(LE-11)+ +0.2x6)

Table 8. Departments from which average resource consumption was obtained for the alternative staging procedures and treatments. All services except PET Center are in hospital 12 de Octubre.

Resource Consumption	Procedures/treatments					
	MS-T	CT	PET	Surgery	Chemotherapy	Radiotherapy
Materials	Thoracic surgery department	Radiodiagnosis Department	NA	Thoracic surgery Department	Oncology Department	Oncological radiotherapy Department
Pharmaceuticalproducts						
Personnel time						
Hospital stay						
Overheads	Accounting department	Accounting department		Accounting department	Accounting department	Accounting department
Capital depreciation	Not applicable			Not applicable	Not Applicable	

Table 9. Departments from which unit costs were obtained for the alternative staging procedures and treatments. All services except PET Center are in hospital H.

Resource Consumption	Procedures/treatments					
	MS-T	CT	PET	Surgery	Chemotherapy	Radiotherapy
Materials	Supplies department		NA	Supplies department		
Pharmaceuticalproducts	Hospital pharmacy			Hospital pharmacy		
Personnel time	Personel department			Personel department		
Hospital stay	Accounting department	Accounting department		Accounting department	Accounting department	Accounting department
Overheads						
Capital depreciation	Not applicable			Not applicable	Not applicable	

Table 10. Allocation of overhead costs and capital costs across treatments/procedures

Procedure/treatment	Overhead cost allocation	Capital depreciation allocation
MS-T	Fixed costs were allocated per patient-day spent in the thoracic surgery ward	Not applicable
CT	Fixed costs were allocated per CT	Equivalent annual cost of CT equipment

	procedure perforMs-t in the radiodiagnosis department	(no resale price, interest rate 5%) was divided into number of CTs perforMs-t annually
PET	Fixed costs were allocated per PET procedure perforMs-t in the PET Center	Equivalent annual cost of PET cyclotron equipment (no resale price, interest rate 5%) was divided into number of PETs perforMs-t annually
Surgery	Fixed costs were allocated per patient-day spent in the thoracic surgery ward	Not applicable
Chemotherapy	Fixed costs were allocated by patient treated in the oncology ward	Not applicable
Radiotherapy	Fixed costs were allocated by patient treated in the oncological radiotherapy department	Equivalent annual cost of accelerator equipment (no resale price, interest rate 5%) was divided into number of patients treated annually

Table 11. Baseline values (range) for all variables in the decision tree model.

Variable	Baseline value (range for sensitivity analysis)	Variable	Baseline value (range for sensitivity analysis)
ProbMS-Tpos	0.27 (0.2-0.36)	CostMS-T	626 (313-918)
ProbCTpos	0.48 (0.39-0.57)	CostCT	582 (291-873)
ProbCTorPETpos	0.58 (0.47-0.66)	CostPET	450 (450-900)
ProbMS-TposCTpos	0.49 (0.36-0.62)	CostSurgery	3,938 (1,969-5,907)
ProbMS-TposCTorPETpos	0.43 (0.32-0.55)	Cost Chemotherapy	6383 (3,191-9,574)
ProbSmortMS-Tneg	0.08 (0.04-0.16)	Cost	5836 (2,918-8,754)

		Radiotherapy	
ProbSmortCTposMS-Tneg	0.16 (0.05-0.33)	Costs below are derived from costs above using simple arithmetic formulae (range of values not available from DATA software)	
ProbSmortCTneg	0.04 (0.01-0.12)		
ProbSmortCTorPETposMS-Tneg	0.16 (0.07-0.31)		
ProbSmortCTandPETneg	0.02 (0.001-0.01)	CostMs-tqr	12,845
QMs-tqr	0.52 (0.03-2.20)	CostMs-tss	6,118
QMs-tss	2.78 (0.05-6.20)	CostMs-tds	4,564
QMs-tds	0 (0)	CostctMs-tqr	13,427
QctMs-tqr	0.58 (0.03-2.20)	CostctMs-tss	9,712
QctMs-tss	1.78 (0.07-4.33)	CostctMs-tds	5,146
QctMs-tds	0 (0)	Costctss	6,388
Qctss	3.07 (0.05-3.20)	Costctds	4,520
Qctds	0 (0)	CostctpetMs-tqr	13,877
QctpetMs-tqr	0.47 (0.03-1.60)	CostctpetMs-tss	9,375
QctpetMs-tss	1.83 (0.07-4.33)	CostctpetMs-tds	5,596
QctpetMs-tds	0 (0)	Costctpetss	6,757
Qctpetss	3.22 (0.05-6.20)	Costctpetds	4, 970
Qctpetds	0 (0)		

Table 12. Cost-effectiveness analysis.

Strategy	Cost (EUR)	Incremental cost (EUR)	Effectiveness (QALYs)	Incremental Effectiveness (QALYs)	ICER (EUR/QALY)
SCM	7843.7		2.007		
IS1	8639.6	795.9	2.035	0.028	28,904
IS2	9183.5	543.9	1.951	-0.084	Dominated

Table 13. One-way sensitivity analysis. ICERs at minimum and maximum values for probabilities, QALYs and costs

Variable	**Minimum value:** **Result**	**Maximum value:** **Result**
ProbMS-Tpos	0.2: IS1 and IS2 dominated by SCM	0.36: ICER (IS1-SCM) = 850 EUR/QALY IS2 dominated by IS1
ProbCTpos	0.39: ICER (IS1-SCM) = 1,812 EUR/QALY IS2 dominated by IS1	0.57: IS1 and IS2 dominated by SCM
ProbCTorPETpos	0.47: Extended dominance of strategy IS1 by SCM and IS2. ICER (IS2-SCM) = 4,000 EUR/QALY	0.66: ICER (IS1-SCM) =28,904 EUR/QALY IS2 dominated by SCM and IS1

ProbSmortMS-Tneg	0.04: Strategies IS1 and IS2 dominated by SCM	0.16: ICER (IS1-SCM) = 4,669 EUR/QALY IS2 dominated by SCM and IS1
Prob SmortCTposMS-Tneg	0.05: ICER (IS1-SCM) = 12,176 EUR/QALY IS2 dominated by SCM and IS1	0.33: IS1 and IS2 dominated by SCM
ProbSmortMS-Tneg	0.04: IS1 and IS2 dominated by SCM	0.16: ICER (IS1-SCM) = 4,670 EUR/QALY IS2 dominated by IS1
Prob SmortCTorPETpos MS-Tneg	0.07: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1	0.31: ICER (IS1-SCM) =28,904 EUR/QALY IS2 dominated by IS1
Prob SmortCTandPETneg	0.001: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1	0.01: ICER (IS1-SCM) =28,904 EUR/QALY IS2 dominated by IS1
CostMS-T	313 EUR: ICER (IS1-SCM) = 27,180 EUR/QALY IS2 dominated by IS1 and SCM	938 EUR: ICER (IS1-SCM) =30,621 EUR/QALY IS2 dominated by IS1 and SCM
CostCT	291 EUR: ICER (IS1-SCM) = 18,336 EUR/QALY IS2 dominated by IS1 and SCM	873 EUR: ICER (IS1-SCM) =39,471 EUR/QALY IS2 dominated by IS1 and SCM
CostPET	450 EUR: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1 and SCM	900 EUR: ICER (IS1-SCM) =28,904 EUR/QALY IS2 dominated by IS1 and SCM
CostSurgery	1,969 EUR: ICER (IS1-SCM) = 26,415 EUR/QALY IS2 dominated by IS1 and SCM	5907 EUR: ICER (IS1-SCM) =31,392 EUR/QALY IS2 dominated by IS1 and SCM

Table 13 (cont).

Variable	**Minimum value: Result**	**Maximum value: Result**
Cost Chemotherapy	3,191 EUR: ICER (IS1-SCM) = 29,516 EUR/QALY IS2 dominated by IS1 and SCM	9,574 EUR: ICER (IS1-SCM) = 28,291 EUR/QALY IS2 dominated by IS1 and SCM
Cost Radiotherapy	2,918 EUR: ICER (IS1-SCM) = 27,619 EUR/QALY IS2 dominated by IS1 and SCM	8,754 EUR: ICER (IS1-SCM) = 29,188 EUR/QALY IS2 dominated by IS1 and SCM
QMs-tqr	0.03 QALYs: ICER (IS1-SCM) = 4,979 EUR/QALY IS2 dominated by IS1 and SCM	2.20 QALYs: IS1 and IS2 dominated by SCM
QctMs-tss	0.07 QALYs: IS1 and IS2 dominated by SCM	4.33 QALYs: ICER (IS1-SCM) = 1,442 EUR/QALY

		IS2 dominated by IS1 and SCM
QctMs-tqr	0.03 QALYs: IS1 and IS2 dominated by SCM	2.20 QALYs: ICER (IS1-SCM) = 1,948 EUR/QALY IS2 dominated by IS1 and SCM
Qctss	0.05 QALYs: IS1 and IS2 dominated by SCM	3.20 QALYs: ICER (IS1-SCM) = 4,708 EUR/QALY IS2 dominated by IS1 and SCM
QctpetMs-tqr	0.03 QALYs: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1 and SCM	1.60 QALYs: Extended dominance of IS1 by SCM and IS2 ICER (IS2-SCM) = 5,951 EUR/QALY
QctpetMs-tss	0.07 QALYs: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1 and SCM	4.33 QALYs: Extended dominance of IS1 by SCM and IS2 ICER (IS2-SCM) = 2,101 EUR/QALY
Qctpetss	0.05 QALYs: ICER (IS1-SCM) = 28,904 EUR/QALY IS2 dominated by IS1 and SCM	6.2 QALYs: Extended dominance of IS1 by SCM and IS2 ICER (IS2-SCM) = 1,145 EUR/QALY