ORIGINAL REPORT

Abdominal ultrasonography versus abdominal CT in the follow-up of testicular germ cell tumors

E. Murias Quintana a,*, E. Esteban b, A. López-Muñiz c, A. Saiz a, S. Costilla a, J.R. Jiménez a

a Servicio de Radiodiagnóstico, Hospital Universitario Central de Asturias, Oviedo, Spain  
b Servicio de Oncología Médica, Hospital Universitario Central de Asturias, Oviedo, Spain  
c Departamento de Morfología y Biología Celular, Universidad de Oviedo, Oviedo, Spain

Received 17 November 2009; accepted 18 August 2010

Abstract

Introduction: Testicular germ cell tumors are common in young men. There is a wide consensus regarding the key points in their diagnosis and treatment, although some elements, including the best approach to follow-up, are being reviewed and revised.

We present a statistical study that uses tools for the evaluation of diagnostic tests to compare the usefulness of abdominal ultrasonography (US) with that of abdominal CT, taken as the gold standard, in the detection of liver metastases and retroperitoneal adenopathies in patients with testicular germ cell tumors.

Material and methods: We analyzed a total of 308 diagnostic tests (154 CT studies and 154 US studies) from 59 patients with at least one year’s follow-up at our institution. Patients underwent abdominal US before abdominal CT following a standard protocol.

Results: Compared to the gold standard, abdominal US had 95% sensitivity, 92% specificity, 82% positive predictive value (PPV), and 98% negative predictive value (NPV).

Conclusion: US is very sensitive and can be used in protocols for the follow-up of primary testicular tumors to rule out the disease.

© 2009 SERAM. Published by Elsevier España, S.L. All rights reserved.

PALABRAS CLAVE
Seminoma; Neoplasias testiculares; Tumor de células germinales testiculares; Estadificación del cáncer del testículo;

Comparación de la ecografía abdominal con la tomografía computarizada abdominal en el seguimiento del tumor de células germinales testicular

Resumen

Introducción: El tumor de células germinales testicular es una neoplasia frecuente que afecta a varones jóvenes. En la literatura médica existe un consenso en los puntos clave que se refieren a su diagnóstico y tratamiento, aunque hay elementos que se encuentran actualmente en revisión, entre ellos los métodos de seguimiento de esta enfermedad.


* Corresponding author.
E-mail address: emurias@hotmail.com (E. Murias Quintana).

2173-5107/$ - see front matter © 2009 SERAM. Published by Elsevier España, S.L. All rights reserved.
Introducción

Los tumores de células germinales del testículo son frecuentes en el momento adulto, siendo el más común tumor maligno en hombres entre los 20 y 30 años de edad. Se caracteriza por la presencia de un dolor sincomún en el testículo, que puede ser asociado con dolor lumbar. En casos de seminoma, la ecografía abdominal, la tomografía computarizada (TC) y la resonancia magnética (RM) son las técnicas habituales para su diagnóstico. En los casos de no-séminoma, se considera la realización de tomografía computarizada abdominal y testicular. Los tumores de células germinales sitiales son considerados menos biológicamente agresivos que los sarcomas de células germinales.

Material y métodos: Analizamos 308 pruebas diagnósticas, 154 estudios de TC abdominal y 154 ecografías abdominales correspondientes a 59 pacientes seguidos durante un periodo de al menos un año en nuestra institución. Los pacientes se estudiaron mediante una ecografía abdominal que precedía a la TC abdominal, protocolo de seguimiento estándar.

Resultados: La ecografía abdominal presenta, respecto a la prueba de referencia, una sensibilidad del 95%, una especificidad del 92%, un VPP del 82% y un VPN del 98%.

Conclusión: La ecografía es una prueba con gran sensibilidad que se puede utilizar para descartar la recidiva abdominal en los protocolos de seguimiento de la neoplasia primitiva de testículo.

© 2009 SERAM. Publicado por Elsevier España, S.L. Todos los derechos reservados.

To summarize, in usual clinical practice, the imaging modality used in the follow-up of testicular germ cell carcinoma is abdominal CT. It is known that CT is associated with high radiation burden, especially if total accumulated exposure is taken into account since it is young patients who are being treated; moreover, the iodinated contrast agents that are required for CT scanning may have, although in very few cases, side effects.

This study raises the hypothesis that sector imaging obtained with abdominal ultrasound (US) is a diagnostic method that can be useful in the follow-up of testicular tumors. The use of US would reduce the dose of radiation that these patients receive.

The aim of the present study is to assess the diagnostic accuracy of US versus CT in the diagnosis of recurrence of testicular germ cell tumor in the abdominal cavity (presence of retroperitoneal adenopathic masses and focal liver lesions).

Material and methods

A prospective, epidemiologic, observational and descriptive study was designed to monitor the one-year progression of 59 consecutive patients, randomly selected between January 1, 2003 and January 1, 2008. These patients were diagnosed with and treated for testicular neoplasm and joined a follow-up protocol. The body mass index was not an exclusion criteria and none of the patients were excluded because of lack of availability of diagnostic means.

In compliance with the protocols accepted by our institution, the standard imaging follow-up was performed using abdominal CT during the first year: 4 abdominal CT studies separated by a 3-month interval (CT multidector Aquilion 16 Toshiba medical systems, protocol: oral and intravenous (iv) contrast 120 ml at 3 ml/s, starting 40 s after injection, covering the region from the lung bases to the pelvis).

Before the first CT exam, the patient was informed and asked for consent to have an abdominal US performed (ultrasound Apio XG Toshiba medical systems) after the CT and
before the subsequent follow-up studies. No patient refused to participate in the study.

The US examinations were performed by a team of radiologists (5 experimented specialists) using single reading and before the CT examination. The double-blind could not be guaranteed.

A section was included in the report where the radiologist indicated, according to their opinion and experience, the degree of reliability of the US: unreliable (if the radiologist had serious doubts regarding the US findings); reliable (if the radiologist was relatively convinced of the US diagnosis), and diagnostic (if the radiologist was convinced beyond doubt of the US diagnosis). CT and US criteria for tumor recurrence are shown in Tables 1 and 2.

In 19 of the 59 patients, the imaging follow-up protocol was modified by reducing the intervals between CT studies, because of recurrence of the disease and its treatment. Nonetheless, the US examinations continued to be performed in all these patients before the follow-up and diagnostic CT examinations. No CT study was considered insufficient to determine a diagnosis and no patient had contraindications for intravenous contrast administration.

US and CT studies corresponding to the diagnosis of extension prior to treatment were excluded since this study was designed to assess follow-up examinations. A total of 154 abdominal CT examinations were obtained and as many US (Tables 1 and 2).

No data regarding the size and number of either hepatic or retroperitoneal lesions were collected; only data about the presence or absence of signs of tumor recurrence or tumor extension and location were compared.

Subsequently these results were analyzed using statistical tools that assessed the internal validity (sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio, percentage agreement and kappa index, see Table 3) of US versus abdominal CT taken as the reference standard.

### Results

The average interval between the US and the CT study was of 8.3 days (range 2–12 days). Out of the 154 imaging tests, the US examinations were classified as positive (50) or negative (104); and the abdominal CT examinations discriminated diseased (43) from healthy (111). Out of the positive US (50), 3 were unreliable, 18 reliable and 29 diagnostic. Of the negative US (104) 2 were unreliable, 50 reliable and 52 diagnostic (Table 4):

- Sensitivity: TP/TP + FN = 95%.
- Specificity = TN/TN + FP = 92%.
- Positive likelihood ratio = S/1 – SPC = 11.8.
- Negative likelihood ratio = 1 – S/SPC = 0.05.
- PPV = positive predictive value = TP/TP + FP = 82%.
- NPV = negative predictive value = TN/TN + FN = 98%.
- Percentage agreement = TP + TN/N = 0.93.
- Kappa index = Po − Pe/1 − Pe; Po = (f1c1 + f2c2)/N² = 0.57; kappa index = 0.83.

The 2 patients with negative abdominal US that the CT classified as diseased (false negatives) had liver lesions, both in the dome. In both cases, the radiologist rated the US as “reliable”.

In 9 patients, the US diagnosed signs of tumor recurrence but the CT did not confirm these findings (false positives). In 8 of these patients, the US diagnosed retroperitoneal involvement and the radiologist rated them as “unreliable”. In the other patient, the US diagnosed retroperitoneal involvement and the radiologist rated it as “reliable”.

If true positives are analyzed, of the 41 US studies that diagnosed tumor recurrence confirmed with abdominal CT, 31/41 were retroperitoneal lesions and 10/41 hepatic recurrence. Nine of these 10 patients with hepatic recurrence also had retroperitoneal recurrence (Figs. 1 and 2).

### Discussion

It is justified to carry out studies that compare diagnostic tests that do not use ionizing radiation with CT in order to reduce the radiation burden on young oncology patients undergoing follow-up. There are numerous

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Diagnostic criteria of the imaging tests.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A patient is considered healthy if:</td>
<td>The abdominal CT was reported as normal CT pathological findings are not indicative of recurrence or tumor extension</td>
</tr>
<tr>
<td>A patient is considered diseased if:</td>
<td>Abdominal CT shows signs of recurrence or tumor extension</td>
</tr>
<tr>
<td>An US is considered positive if:</td>
<td>It shows radiological signs of recurrence or tumor extension</td>
</tr>
<tr>
<td>An US is considered negative if:</td>
<td>It is reported as normal if pathological findings are not indicative of recurrence or tumor extension</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Radiological signs of recurrence or tumor extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signs of tumor recurrence</td>
<td></td>
</tr>
<tr>
<td>1. Retroperitoneum: retroperitoneal masses or adenopathy with short axis &gt; 1 cm in the renal hilum or in the perivascular region</td>
<td></td>
</tr>
<tr>
<td>2. Pelvis: inguinal or iliac masses or adenopathy with short axis &gt; 1 cm</td>
<td></td>
</tr>
<tr>
<td>3. Liver: hepatic masses described as ”suspicious”, ”compatible with” or ”suggestive of” hepatic metastases. Masses that do not match the criteria for hepatic cysts or typical hemangiomas, not known through the diagnostic tests</td>
<td></td>
</tr>
</tbody>
</table>
examples of this type of studies in medical literature concerning retroperitoneal and hepatic disease and, in the particular case of testicular germ neoplasms, there are studies that confirm the efficacy of US in the evaluation of patients who have undergone lymphadenectomy. Therefore, in our opinion it is justified to consider this diagnostic test as part of the follow-up protocols. 25-31

Abdominal US allows for the selection of healthy individuals. In our series, due to its high sensibility (95%) with high NPV (98%), a patient diagnosed as healthy by negative abdominal US is highly likely of being so. Similarly, the low negative likelihood ratio of US allows us to state that a negative US most likely does not come from a diseased patient, thus both factors allow the US to rule out healthy patients. The kappa index value exceeds 0.8 over 1, suggestive that there is a "very good" degree of agreement between the diagnostic tests and it rules out that the results might be due to chance.

Abdominal US in our study shows high specificity (92%), but not a very high PPV (82%), which seems to prove that US is not useful to confirm that a patient is actually diseased, since it produces too many false positives. It can be inferred from this that, if positive, the abdominal US should be confirmed with other tests with higher specificity such as abdominal CT.

<table>
<thead>
<tr>
<th>Name of the test</th>
<th>Definition</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (S)</td>
<td>Probability of the US being positive in a diseased individual</td>
<td>TP/TP + FN</td>
</tr>
<tr>
<td>Specificity (SPC)</td>
<td>Probability of the US being negative in a healthy individual</td>
<td>TN/TN + FP</td>
</tr>
<tr>
<td>Likelihood ratio (positive [LR+] and negative [LR−])</td>
<td>The likelihood or probability ratio compares the possibility that a result comes from a diseased versus the possibility that this result comes from a healthy individual</td>
<td>LR+ = S/1 − SPC; LR− = 1 − S/SPC</td>
</tr>
<tr>
<td>Predictive value (positive [PPV] and negative [PVN])</td>
<td>Probability that an individual with a positive result has the disease (PPV) and that an individual with a negative result does not have the disease (NPV)</td>
<td>PPV = TP/TP + FP; NPV = TN/TN + FN</td>
</tr>
<tr>
<td>Percentage agreement (PA)</td>
<td>The proportion of agreement between both diagnostic tests for all subjects indicates the percentage of times that both tests agree (without taking into account the agreement occurring by chance)</td>
<td>PA = TP + TN/N</td>
</tr>
<tr>
<td>Kappa index (KI)</td>
<td>Statistical tool that allows for the estimation of the degree of agreement between both tests expected to occur through chance</td>
<td>KI = Po − Pe/1 − Pe; Where Po is the observed agreement (percentage agreement) and Pe is the agreement expected to occur through chance, estimated taking into account the probability of independent events, both for a positive and negative result Po = (f1c1 + f2c2)/N²</td>
</tr>
</tbody>
</table>

FN: false negatives; FP: false positives; TN: true negatives; TP: true positives.

Table 3: Statistical tools for internal validity in diagnostic test comparison.

<table>
<thead>
<tr>
<th>Table 4: Contingency table of imaging tests.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal CT</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Positive US</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>US</td>
</tr>
</tbody>
</table>
Abdominal ultrasonography versus abdominal CT

The false positives, that is, the cases diagnosed by US that have not been confirmed by CT occurred in the retroperitoneum. In 8 out of the 9 cases, the radiologist rated the US as unreliable, therefore, it can be concluded that the experience and opinion of the radiologist is fundamental to reduce false positives and to increase the specificity and PPV. This, however, does not represent a problem in follow-up protocols since in case of positive US, an abdominal CT would be performed in order to confirm or rule out the presence of recurrence.

Figure 1  (A and B) Ultrasound shows examples of tumor recurrence due to hepatic metastases.

The limitation of US in retroperitoneal assessment is well known, a finding that has not been confirmed in this research, probably because young patients in good clinical health predominated in our series.

It can therefore be inferred that abdominal US allows us, with a very low number of false negatives, to rule out the disease both in retroperitoneum and in liver. The clinical application of US would be justified as initial assessment in the follow-up of testicular neoplasms, used in a similar way as in other imaging protocols such as initial assessment of a trauma patient, of acute abdominal disease and of acute renal failure, as well as in the follow-up of atherosclerotic and heart diseases, and screening in obstetrics and gynecology conditions.

In our case, US could be used as an initial diagnostic test for recurrence, in conjunction with abdominal CT as a more specific test to confirm the findings, regardless of whether there are pathological data or if the US is a priori of low quality. The abdominal CT examination could be prescribed with longer time intervals, for example reducing the number of examinations to 3 during the first year (at the beginning, 6 and 12 months) with intervening US examinations (at 2, 4, 8 and 10 months). This would allow a lesser number of ionizing radiation tests and a reduction by half in the follow-up intervals thus benefiting the patient.

These results and this protocol could be complemented with tumor marker tests that in case of being elevated could mean a false negative and an abdominal CT would be then indicated. This reflection is already included in other neoplasm follow-up protocols, such as breast, cervical, ovarian and liver neoplasms, among others.

Conclusion

US has a good ability to rule out the presence of both liver and retroperitoneal disease in patients with testicular neoplasms undergoing follow-up. If it were to be included in the follow-up protocols, the cumulative radiation exposure could be reduced in these patients.

Authorships

1. Responsible for the integrity of the study: EMQ.
2. Conception of the study.
3. Design: EMQ, EE, ALM.
4. Acquisition of data: EMQ, EE, SC, JRJ.
5. Analysis and interpretation of data: AS.
6. Statistical analysis: EQM, ALM, AS, SC.
8. Drafting of the manuscript: EMQ.
9. Critical review with intellectually relevant contributions: JRJ.
10. Approval of the final version: EMQ, EE, ALM, AS, SC and JRJ.

Conflict of interests

The authors declare not having any conflict of interests.
References


