Preoperative staging of rectal cancer with MRI: correlation with pathologic staging

Soraia Filipa Macado Abreu a,b, Sandra Fátima Fernandes Martins a,b,c,*

a Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal
b ICVS/3B’s, PT Government Associate Laboratory, Braga/Guimarães, Portugal
c Surgery Department, Hospital Center of Trás-os-Montes e Alto Douro, Distrito de Vila Real, Portugal

A R T I C L E   I N F O
Article history:
Received 20 December 2014
Accepted 20 February 2015
Available online 20 April 2015

Keywords:
Rectal cancer
Pelvic magnetic resonance
Accuracy
Sensibility
Specificity

A B S T R A C T
Introduction: An accurate preoperative rectal cancer staging is crucial to the correct management of the disease. Despite great controversy around this issue, pelvic magnetic resonance (RM) is said to be the imagiologic standard modality. This work aimed to evaluate magnetic resonance accuracy in preoperative rectal cancer staging comparing with the anatomopathological results.

Methods: We calculated sensibility, specificity, positive (VP positive) and negative (VP negative) predictive values for each T and N. We evaluated the concordance between both methods of staging using the Cohen weighted $K(K_w)$, and through ROC curves, we evaluated magnetic resonance accuracy in rectal cancer staging.

Results: 41 patients met the inclusion criteria. We achieved an efficacy of 43.9% for T and 61% for N staging. The respective sensibility, specificity, positive and negative predictive values are 33.3%, 94.7%, 33.3% and 94.7% for T1; 62.5%, 32%, 37.0% and 57.1% for T2; 31.8%, 79%, 63.6% and 50% for T3 and 27.8%, 87%, 62.5% and 60.6% for N. We obtained a poor concordance for T and N staging and the anatomopathological results. The ROC curves indicated that magnetic resonance is ineffective in rectal cancer staging.

Conclusion: Magnetic resonance has a moderate efficacy in rectal cancer staging and the major difficulty is in differentiating T2 and T3.

© 2015 Sociedade Brasileira de Coloproctologia. Published by Elsevier Editora Ltda. All rights reserved.

Estadiamento pré-operatório do câncer do reto por ressonância magnética: correlação com estadiamento anatomopatológico

R E S U M O
Introdução: Um estadiamento pré-operatório do Câncer do Reto (CR) é essencial na gestão da doença. Apesar de grande controvérsia, a ressonância magnética pélvica (RM) é apontada...
Introduction

The incidence and mortality of oncological diseases have increased at an alarming rate worldwide, and according to World Health Organization, the incidence of cancer will increase to 22 million/year in the next two decades.¹

In Portugal, colorectal cancer (CRC) is the second most frequent cancer in both men and women, with an overall incidence of 14.5% and 15.7% mortality. In 2012, this disease was appointed as the second leading cause of cancer death in both genders.² Rectal cancer (RC), in particular, exhibited in northern region of this country in 2008 an incidence of 24.6/1,00,000 inhabitants; a progressive increase in this value since 1999 was found. In the district of Braga, its incidence in this year was 16.8/1,00,000 inhabitants.³

RC prognosis has improved greatly in recent decades and this was mainly due to advances in preoperative staging, which was reflected in the therapeutic approach,⁴ where a change was observed, from a purely surgical treatment to a multidisciplinary approach,⁵,⁶ lowering the lower local recurrence rate to 11% and improving the survival rate at 5 years to 58%,⁷ compared to previous values (27 and 48%, respectively).⁵

Preoperative staging of RC is divided into local and remote staging.⁸-¹⁰ The information (both clinical and pathological) gathered from the staging is grouped according to “Tumor-Node-Metastasis” (TNM) classification of the American Joint Committee on Cancer in different prognostic groups or anatomical stages.¹¹,¹² T and N stages are the best determinants of prognosis, being critical components of preoperative staging.¹³ Thus, in addition to the use of a reliable imaging modality and with high accuracy in preoperative staging of CR be a crucial prognosis factor,¹⁴ this procedure also helps in choosing the best therapeutic strategy, enabling a balance between oncological safety and quality of life of the patient.¹⁰,¹²

Nowadays, there are several imaging modalities for preoperative staging of RC, including computed tomography (CT), pelvic magnetic resonance imaging (MRI), endorectal ultrasound (EUS), and positron emission tomography (PET).¹⁰ EUS and MRI are the main instruments of preoperative staging.¹⁵ However, there is no consensus about the best method.⁵,¹⁶ Several studies suggest MRI as being superior to EUS,⁵,¹⁷-²⁰ and therefore MRI is the routine imaging modality for preoperative staging of RC.⁵,²¹ As to T staging, MRI shows an acuity between 55 and 86%, as to N staging, the accuracy varies from 39 to 95%.⁵

In RC staging, the main role of MRI lies in the evaluation of tumors in advanced and occlusive stages.¹⁵,²² With respect to stages T3 and T4, sensitivity and specificity of 74 and 76% (for T3) and 82 and 96% (for T4) were observed, respectively.²² For T1 and T2 stages, MRI is less sensitive versus EUS, with no difference relative to T3 and T4 stages.¹⁶

Detection of N is the biggest challenge of any imaging examination, with 66% sensitivity of MRI in its evaluation and 76% specificity.⁵

This study aims to evaluate the accuracy of MR in the preoperative staging of RC, compared to the anatomopathological result of the surgical specimen.

Materials and methods

Population

The target population for this study consists of 216 patients with histological diagnosis of RC treated in the General Surgery Service, Hospital de Braga (HB) between January 1st, 2007 and December 31, 2013.

Inclusion criteria for this study were: patients with histological diagnosis of rectal adenocarcinoma; patients with a conclusive preoperative staging by MRI, and patients with pathology staging results based on the surgical specimen.

Exclusion criteria were the following: patients with a histological diagnosis differing from the above; patients with a diagnosis of RC who did not undergo MRI or for whom such analysis was inconclusive; patients without results from pathology staging, and patients undergoing primary treatment.
Sample

A convenience sample of 41 patients who meet inclusion/exclusion criteria previously defined was studied.

Data collection

Clinical and staging data collected include: age, gender, oncological history, tumor location, and T/N staging by means of MRI.

Pathological data comprise the histological type and TNM staging.

Follow-up data are related to the occurrence of relapse and death.

This project was approved by HB’s Ethics Committee and also by Ethics Subcommittee for Life and Health Sciences.

Statistical analysis

The collected data were organized in an Excel (Microsoft® Office 2010) database, and the Statistical Package for Social Sciences (SPSS) version 22.0 (SPSS Inc., Chicago, IL, USA) was also used.

A descriptive analysis of the variables under study, to provide frequencies, means and standard deviations, was performed.

Sensitivity, specificity, positive and negative predictive values (PV) of RM staging, related to T and N, were compared with pathology results. For this purpose, the online tool MedCalc® available in http://www.medcalc.org/calc/diagnostic_test.php, was used. Efficacy was calculated by the formula \((TP+TN/n)\) and the confidence interval (CI) by the formula: \(\frac{P-Z \times \sqrt{P(1-P)/n}; P+Z \times \sqrt{P(1-P)/n}}{N}\).

The agreement between the staging results obtained by MRI and anatomopathological study was assessed by calculating the value of weighted Cohen’s Kappa \((K_w)\); to this end, we used the online tool VassarStats, available in http://vassarstatis.net/kappa.html. A value of \(K_w\) between 0.00 and 0.20 indicates poor agreement; between 0.21 and 0.40 reveals a considerable agreement; between 0.41 and 0.60, points to a moderate agreement; between 0.61 and 0.80 indicates a good agreement; and between 0.81 and 1.00 shows excellent agreement.23

Finally, using SPSS program, we used ROC curves for a comparative study of MRI and anatomopathological staging, by means of area under the curve (AUC) calculation. ROC curve with AUC assessment is a good predictor of the accuracy of a test. The closer the AUC is to 1, the better the examination is. Values <0.50 represent a poor or ineffective test; between 0.50 and <0.70 indicate an test with average or reasonable accuracy; and values ≥0.80 predict a good or excellent test.24

For all tests, it was assumed a significance of 0.05 and a confidence interval (CI) of 95%.

Results

Sample characterization

The study population consists of 41 subjects who meet inclusion and exclusion criteria. Of these, 58.5% (n = 24) are male, aged between 31 and 88 years.

In 12.2% (n = 5) of patients, there is a personal history of another cancer.

As for tumor localization, 14.6% (n = 6) of the cases are located in the lower third, 68.3% (n = 28) in the middle third and 17.1% (n = 7) in the upper third of the rectum. Disease recurrence was observed in 22% (n = 9) patients. Until August 2014 there were 10 deaths in the study population.

T staging

Regarding MR tumor staging, 7.32% (n = 3) of the tumors are classified as T1, 65.9% (n = 27) are staged as T2 and 26.8% (n = 11) as T3. With regard to anatomopathological staging of surgical specimens, 7.32% (n = 3) neoplasms are classified as T1, 39% (n = 16) are staged as T2 and 53.7% (n = 22) as T3 (Table 1).

Comparing MR staging versus anatomopathological results based on surgical specimen (Table 2), we noted substaging in 4.88% of cases (n = 2) staged by MR as T1; overstaging in 4.88% (n = 2) and substaging in 36.6% (n = 15) of tumors staged by MRI as T2; and overstaging in 9.76% (n = 4) of patients staged as T3.

The sensitivity of MRI in the preoperative staging of RC relative to T was calculated, and sensitivities of 33.3% (95% CI, 5.47–88.5) for T1, 62.5% (95% CI, 35.5–84.7) for T2 and 31.8% (95% CI 13.9–54.9) for T3 were observed. As for specificity, this parameter is 94.7% (95% CI 82.2–99.2) for T1, 32% (95% CI, 15–53.5) for T2 and 79% (95% CI 54.4–93.8) for T3. The positive PV calculated is 33.3% (95% CI 5.47–88.5) for T1, 37.0% (95% CI 19.4–57.6) for T2 and 63.6% (95% CI 30.9–88.9) for T3. In relation to negative PV, values of 94.7% (95% CI 82.2–99.2) for T1, 57.1% (95% CI 28.9–82.2) for T2 and 50% (95% CI 31.3–68.7) for T3 were found. RM shows efficacy of 43.9% for T staging, 90.2% in particular for T1 staging, 43.9% for T2 staging and 53.7% for T3 staging.

| Table 1 – MR and anatomopathological staging related to T and N. |
|------------------|------------------|------------------|
|                  | RM staging       | Anatomopathological staging |
|                  | Frequency (%)    | Frequency (%)     |
| T0               | 0 (0)            | 0 (0)             |
| T1               | 3 (7.2)          | 3 (7.2)           |
| T2               | 27 (65.9)        | 16 (39.9)         |
| T3               | 11 (26.8)        | 22 (53.7)         |
| T4               | 0 (0)            | 0 (0)             |
| N0               | 36 (87.8)        | 23 (56.1)         |
| N+               | 5 (12.2)         | 18 (43.9)         |

| Table 2 – Comparison between RM and anatomopathological staging, related to T. |
|------------------|------------------|------------------|
|                  | RM staging       | Anatomopathological staging |
|                  | T1 (%)           | T2 (%)           | T3 (%) | Total (%) |
| Anatomopathological staging | | | |
| T1                | 1 (2.44)         | 2 (4.88)         | 0 (0)  | 3 (7.32)  |
| T2                | 2 (4.88)         | 10 (24.4)        | 4 (9.76)| 16 (39.0) |
| T3                | 0 (0)            | 15 (36.6)        | 7 (17.1)| 22 (53.7) |
| Total             | 3 (7.32)         | 27 (65.9)        | 11 (26.8)| 41 (100)  |
The value of $K_w$ for $T$ staging was calculated, in order to determine the correlation between MR versus anatomopathological staging. There is poor agreement between the two forms of staging: $K_w = 0.14$ (95% CI, 0–0.38) $p > 0.05$ (Table 3).

The evaluation of MRI staging accuracy, by formulating ROC curves and AUC determination, reveals an AUC value of 0.26 (95% CI, 0.00–0.54), $p = 0.18$ for $T1$, 0.46 (95% CI 0.28–0.64), $p = 0.63$ for $T2$ and 0.61 (95% CI, 0.43–0.78), $p = 0.24$ for $T3$ (Fig. 1).

**Estadiamento $N$**

With respect to $N$ staging done by MR and shown in Table 1, 87.8% ($n = 36$) of all tumors are classified as $N0$, and 12.2% ($n = 5$) as $N+$. As to anatomopathological staging of surgical specimens, 56.1% ($n = 23$) of the tumors do not have lymph node involvement; in 26.8% ($n = 11$) there is a $N1$ staging, and 9.76% ($n = 4$), received a $N2$ staging. In 7.32% ($n = 3$) of tumors is not possible to classify $N$, since the minimum of 12 lymph nodes were not obtained for the analysis. However, since the analyzed ganglia were metastasized, for the sake of the statistical analysis these ganglia are considered as being $N+$. Thus, the anatomopathological analysis, shown in Table 3, reveals nodal involvement in 43.9% ($n = 18$) of cases, and no such involvement in the remaining 56.1% ($n = 23$).

Comparing the staging relative to $N$ obtained by RM and by anatomopathological analysis, a substaging is noted in 31.7% ($n = 13$) of cases, and an overstaging in 7.32% ($n = 3$). Table 4 compares both forms of staging.

**$M$ staging**

As for $M$ staging, 90.2% ($n = 37$) of patients are staged as $M0$ and 4.88% ($n = 2$) as $M1$; in 4.88% ($n = 2$) of cases it is not possible to determine the existence of distant metastases.

<table>
<thead>
<tr>
<th>$T$</th>
<th>$K_w$ (95%CI)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.14 (0–0.38)</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>N</td>
<td>0.16 (0–0.42)</td>
<td>&gt;0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N$ staging</th>
<th>RM staging</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N0$ (%)</td>
<td>$N+$ (%)</td>
</tr>
<tr>
<td>20 (48.7)</td>
<td>13 (31.7)</td>
</tr>
<tr>
<td>3 (7.31)</td>
<td>5 (12.2)</td>
</tr>
<tr>
<td>23 (56.1)</td>
<td>18 (43.9)</td>
</tr>
</tbody>
</table>

**Fig. 1 – ROC curves related to $T$ and $N$.**
Discussion

A precise preoperative staging of CR is critical for the proper management of this disease, since the therapeutic strategies should be individualized. Moreover, this staging is also a predictor of prognosis. Therefore, the imaging techniques used in tumor staging proved to be decisive, and it is extremely important to ascertain their effectiveness. EUS and MR are the main tools of preoperative staging, but there is no consensus on which is the best method; however, several studies indicate MR to be superior to EUS; therefore, MRI is the standard imaging modality for preoperative staging of RCs.

This study was designed with the aim to determine the accuracy of MR in the preoperative staging process for RC.

With regard to T staging, when comparing the staging performed by MR with the anatomopathological staging, a substaging occurred in 4.88% (n = 2) of cases staged by RM as being T1; there was overstaging in 4.88% (n = 2) of patients and substaging in 36.6% (n = 15) of those tumors staged by MRI as T2 and an overstaging in 9.76% (n = 4) patients staged as T3. Given that, in general, superficial tumors (T1 and T2) without metastatic lymph nodes are treated solely with surgery, while locally advanced tumors or those with nodal metastases are subjected to primary treatment prior to surgical recession, an undertreatment occurred in 36.6% (n = 15) of patients, considering that these subjects were staged as T2, and their anatomopathological results showed that actually these were T3, and thus did not perform primary therapy.

The largest number of cases of incorrect staging by MRI was observed in the distinction between stages T2 and T3. This may in part be explained by the presence of a desmoplastic reaction in peritumoral tissues, making it difficult to distinguish between spiculation of perirectal fat, caused simply by fibrosis, and that that contains viable tumor cells.

As for the 4 patients who were overstaged by RM as T3, when actually these were T2 tumors, these could have been overtreated. However, that did not happen for reasons unrelated to this study.

With respect to the calculation of the MRI sensitivity for T staging, values of 33.3% for T1, 62.5% for T2, and 31.8% for T3 were observed. These values are lower than those observed in several other studies. However, Beaumont et al. obtained even smaller values in relation to T1 and T2 stages, namely 27% for T1 and 59% for T2.

The values for specificity were of 94.7% for T1, 32% for T2 and 79% for T3. With respect to T1 and T3, the results were consistent with the bibliography. In the case of T2, these values were lower than those found in most of the referred studies.

The positive PV calculated was 33.3% for T1, 37.0% for T2 and 63.6% for T3. Regarding negative VP, these values were 94.7% for T1, 57.1% for T2 and 50% for T3. These values were lower than those observed by Uçar et al. and Akasu et al., except in the case of negative VP for T1, which was similar to that calculated by Iannicelli et al.

The diagnostic efficacy for tumor staging has been benefited from improvements due to the development of MR techniques; in early studies, the efficacy reached about 60%; currently, this indicator is between 92 and 94% for T stage and 63% for N stage. Usually the efficiency increases with T stage and varies, according to some authors, between 67 and 94% or 55 and 86%. Efficacy calculated for T stage was 43.9%, particularly 90.2% for T1, 43.9% for T2, and 53.7% for T3. The results for T in general and for T2 and T3 were lower than those found by other authors. The effectiveness for T1 stage was similar that in another study.

N detection is the most challenging detection of any imaging examination. The criterion “size” for detection of lymph node metastases is a poor predictor, since non-tumor enlarged nodes can exist, and the reverse is also true. The irregular contour and heterogeneous signal intensity are more specific criteria for metastasization.

In this study there was substaging in 31.7% (n = 13) of cases classified as N0, and overstaging in 7.32% (n = 3); in 31.7% (n = 13), undertreatment of substaged patients occurred.

The sensitivity, specificity, positive VP and VP negative relative to N were 27.8%, 87%, 62.5% and 60.6%, respectively. The value for sensitivity was lower than those found in several previous studies; however, the specificity showed values higher than those found in these same studies.

Positive and negative VPs were similar to those obtained in other studies. The effectiveness of MR for N staging was 61%, a figure similar to that found by several authors, varying between 39 and 95%.

Positive agreement for T staging results obtained by MR and anatomopathological results, evaluated by $K_v$ value for T and N, were given as: $K_v = 0.14$, $p < 0.05$ and $K_v = 0.16$, $p < 0.05$, respectively. These figures revealed a poor agreement between the two staging forms; additionally, they differ from most studies consulted, where $K_v$ values between 0.71 and 0.89 for T and between 0.40 to 0.56 for N were obtained.

The study by Tytherleigh MG et al. was that that obtained the closest values to ours, namely: $0.37 (p < 0.001)$ for T staging and 0.25 ($p < 0.002$) for N staging.

The evaluation of MR accuracy through ROC curves with AUC determination showed values for AUC of 0.26, $p = 0.18$ for T1, 0.46, $p = 0.63$ for T2, 0.61, $p = 12.24$ for T3, and 0.40, $p = 0.40$ for N staging. Thus, RM proved to be an ineffective or poor method for N staging. These values were discordant with those found by other authors, ranging from 0.81 to 0.94 for T staging and from 0.57 to 0.78 for N staging.

This study has some limitations, such as the relatively small number of patients, particularly those staged as T1. There are also biases associated with retrospective studies, such as selection bias, and the results should be validated by future prospective multicenter studies. MR and anatomopathological staging were not always made by the same radiologist and pathologist; this may be an error factor.

In addition, RM began operating in RC staging in HB in 2007. As in any modality, there is a learning curve associated with this imaging technique, and this may have contributed to some of the reported staging errors.

Finally, publication bias may be another reason for the disparate results found in most of the literature.

Conclusion

RC is a common disease, and its preoperative staging remains a topic of great concern and controversy. In this study, we
evaluated the accuracy of MR in the preoperative staging of RC, compared to anatomicopathological results of surgical specimens.

MR showed a moderate acuity in RC staging, with efficacy values between 43.9 and 90.2% for T staging and 61% for N staging. RM was more sensitive for T2 stage and more specific for T1 and N+ stages. This imaging modality revealed little sensitivity to both T1 and N+ (33.3 and 27.8%, respectively), being more useful for exclusion, than to confirm these stages. Most staging errors occurred on the distinction between T2 and T3 stages, resulting in cases of undertreatment.

The agreement between values obtained by MR and anatomicopathological results was poor for both T and N stages. Thus, in this study, it was an ineffective or poor method for RC staging.

Conflicts of interest
The authors declare no conflicts of interest.

REFERENCES