Learning curve for endorectal ultrasound in young and elderly: lights and shades

AMS 10.1515/med-2016-0074
received August 12, 2016; accepted September 20, 2016

Abstract: Aim of the study is to highlight difficulties faced by an inexperienced surgeon in approaching endorectal-ultrasound, trying to define when learning curve can be considered complete.

A prospective analysis was conducted on endorectal-ultrasound performed for subperitoneal rectal adenocarcinoma staging in the period from January 2008 to July 2013, reported by a single surgeon of Department of Oncology, Section of General Surgery, “San Luigi Gonzaga” Teaching Hospital, Orbassano (Turin, Italy); the surgeon had no previous experience in endorectal-ultrasound. Forty-six endorectal-ultrasounds were divided into two groups: early group (composed by 23 endorectal-ultrasounds, made from January 2008 to May 2009) and late group (composed by 23 endorectal-ultrasound, carried out from June 2009 to July 2013).

In our experience, the importance of a learning curve is evident for T staging, but no statistical significance is reached for results deal with N stage.

We can conclude that ultrasound evaluation of anorectal and perirectal tissues is technically challenging and requires a long learning curve. Our learning curve can not be closed down, at least for N parameter.

Keywords: Endorectal; Ultrasound; Learning curve; Pitfalls

1 Introduction

Endo-Rectal Ultrasound (ERU) in rectal cancer staging is the object of interest in more than 400 studies published in the last 30 years. Most of them are retrospective studies with a low sample size. These studies have very different results and it is not possible to extrapolate the actual ERU diagnostic accuracy: this reflects the different experience among different specialists [1-3]. Since 2007, many studies deal with the learning curve (LC) for ERU. According to international literature, if ERU is performed by an experienced radiologist it is not necessarily a LC [4]; instead, if ERU is performed by a surgeon, a LC is necessary. In this case, the literature emphasizes the importance of the LC because inexperience was cited as a risk factor for ultrasound staging errors [5].

The aim of this study is outlining the LC for ERU and pointing out main pit-falls during medical training in ERU.

2 Material and methods

We analyzed ERU performed for consecutive subperitoneal rectal adenocarcinoma staging in the period from January 2008 to July 2013, reported by a single surgeon (trained by an expert surgeon in ERU), who had no previous experience in ERU. The sonographic reports were compared with the pathology reports according to the pTNM classification of the sixth American Joint Committee on Cancer (AJCC) [6]. The only exclusion criteria was neoadjuvant therapy. Forty-six ERUs were divided into two groups: early group (composed by 23 ERUs, made from January 2008 to May 2009) and late group (composed by 23 ERU, carried out from June 2009 to July 2013).
We analyzed accuracy, overstaging and substaging risk in both early and late group. The statistical significance was evaluated using chi-square test.

**Ethical approval**: The research related to human use has been complied with all the relevant national regulations, institutional policies and in accordance the tenets of the Helsinki Declaration, and has been approved by the authors’ institutional review board or equivalent committee.

### 3 Results

The sample is divided into two groups:
- early: 4 pT0, 5 pT1, 7 pT2, 6 pT3, 1 pT4
- late: 8 pT0, 4 pT1, 5 pT2, 6 pT3, 0 pT4

Statistical analysis is focused on accuracy, overstaging and substaging risk for T stages (Tables 1 and 2).

The same analysis was conducted for N parameter. Total number for uN is lower because in early group 4 patients are uT0/uTis and 3 T2 are treated with TEM (Transanal Endoscopic Microsurgery), so no lymph nodes are removed for histological examination; in late group, 6 patients are uT0/uTis and 4 patients (3 uT1 and 1 uT2) were treated with TEM.

Twenty-nine patients were analyzed for uN parameter, 16 in early group (10 uN0, 6 uN+), 13 in late group (6 uN0, 7 uN+) (Tables 3 and 4).

Statistical analysis for N staging is completed by sensitivity, specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) (Table 5).

#### Table 1: Accuracy, overstaging and substaging risk for T stages – Early group

<table>
<thead>
<tr>
<th>EARLY</th>
<th>uT0/uTis</th>
<th>uT1</th>
<th>uT2</th>
<th>uT3</th>
<th>uT4</th>
<th>Total</th>
<th>Accuracy</th>
<th>Overstaging</th>
<th>Understaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT0/uTis</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>75%</td>
<td>25%</td>
<td>-</td>
</tr>
<tr>
<td>pT1</td>
<td>1</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>0%</td>
<td>80%</td>
<td>20%</td>
</tr>
<tr>
<td>pT2</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>3</td>
<td>-</td>
<td>7</td>
<td>57%</td>
<td>43%</td>
<td>0%</td>
</tr>
<tr>
<td>pT3</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>50%</td>
<td>33.4%</td>
<td>16.6%</td>
</tr>
<tr>
<td>pT4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>100%</td>
<td>-</td>
<td>0%</td>
</tr>
<tr>
<td>Tot/mean</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>3</td>
<td>23</td>
<td>56.4%</td>
<td>45.3%</td>
<td>9.15%</td>
</tr>
</tbody>
</table>

#### Table 2: Accuracy, overstaging and substaging risk for T stages – Late group

<table>
<thead>
<tr>
<th>LATE</th>
<th>uT0/uTis</th>
<th>uT1</th>
<th>uT2</th>
<th>uT3</th>
<th>uT4</th>
<th>Total</th>
<th>Accuracy</th>
<th>Overstaging</th>
<th>Understaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT0/uTis</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>100%</td>
<td>0%</td>
<td>-</td>
</tr>
<tr>
<td>pT1</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>75%</td>
<td>25%</td>
<td>0%</td>
</tr>
<tr>
<td>pT2</td>
<td>1</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>5</td>
<td>60%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>pT3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>66.6%</td>
<td>33.4%</td>
<td>0%</td>
</tr>
<tr>
<td>pT4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tot/mean</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>23</td>
<td>75.4%</td>
<td>19.6%</td>
<td>6.6%</td>
</tr>
</tbody>
</table>

#### Table 3: Accuracy, overstaging and substaging risk for uN stages – Early group

<table>
<thead>
<tr>
<th>EARLY</th>
<th>uN0</th>
<th>uN+</th>
<th>Total</th>
<th>Accuracy</th>
<th>Overstaging</th>
<th>Understaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>pN0</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>81.8%</td>
<td>18.2%</td>
<td>-</td>
</tr>
<tr>
<td>pN+</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>80%</td>
<td>-</td>
<td>20%</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>6</td>
<td>16</td>
<td>81%</td>
<td>18.2%</td>
<td>20%</td>
</tr>
</tbody>
</table>
Differences noted during statistical analysis are evaluated with chi-square test, that shows statistically significant differences in evaluation of T stage ($p = 0.0325; \text{ODDS RATIO} = 0.25$). Otherwise, statistical significance is not reached for N stage ($p = 0.45; \text{odds ratio} = 1.93$) because of the low sample ($n = 29$).

### 4 Discussion

International literature suggests a more than 50 cases LC for T parameter and a more than 75 cases LC for N parameter [6]. Orrom et al. show an increased diagnostic accuracy when ERU is performed by a single expert operator (95%) compared to multiple operators (59.3%). The same observation was made by several authors, who show an increase in diagnostic accuracy from 50% to over 90% [5].

In our experience, the importance of a LC is evident for T staging: accuracy ranges from 56.4% to 75.4% between early and late group. Understaging risk and overstaging risk are reduced (from 9% to 6.6% and from 45.3% to 19.6%, respectively). Technology improvement has two major implications in clinic: a reduced overstaging risk for pT1 (rising from 80% to 25%) and a reduced understaging risk for pT3 (from 16.6% to 0%), the limits for minimally invasive surgical approaches [7] and neoadjuvant therapy indication, respectively.

Analysis shows a reduction of diagnostic accuracy for N staging during the LC (from 81% to 69%). Overstaging risk in late group is 28.6% (vs 18.2% in early group), understaging risk is 33.4% (vs 20% in early group). Results are not statistically significant. Worsening trend was not statistically significant and results during LC get worse maybe because of an higher degree of accuracy and excessively low levels of under- and over-staging observed in early group; in literature, in fact, the accuracy for hanging of N is 68.8%, with the risk of over- and under-staging respectively of 16% and 15% [8-11]. Therefore, values observed in the first 13 scans performed are possibly associated to casualty and LC for the nodal staging seems to be more complex, requiring a higher number of patients compared to our series to be defined finished.

Retrospective analysis of our results allows a critical assessment on main sources of error in staging; Kim described a classification for ERU mistakes, which are divided into five categories [12]:
- Technical errors
- Anatomical changes
- Interpretation mistakes
- False images (ghosting)
- Inevitable mistakes

In our experience, even place, size, location of the tumor can cause a staging mistake: for this reason, our work includes a sixth category: errors related to characteristics of the tumor. In these 6 categories are classified our mistakes.

### 4.1 Technical errors

More frequent error in T parameter staging [12].

#### 4.1.1 Sensor Type

The probe used is a mechanical probe with a rotating crystal inside. Factors that may modify the quality of ultrasound images, and therefore its interpretation, are the
acquisition speed (number of scans per minute, slicing) and degree of probe mobility.

4.1.2 Sensor preparation technique

- Degree of balloon filling: overfilling compresses rectal wall layers, leading to the obliteration of various interfaces. An underfilling of the balloon may affect acoustic contact between probe and lesion.
- Anoscope length (Figure 1): it influences balloon filling type. If it is too long, the balloon takes on a spherical shape; if the anoscope is too short, balloon takes on a cylindrical shape, with a different contact between surface and balloon.
- Cavitation bubbles within water inserted into the balloon: to minimize presence, it is necessary to boil water before filling the balloon.
- Air leaks in the system, at balloon base: air bubbles between probe and rectal wall determine loss of signal.

4.1.3 Sampling of images

- Gain: adjustments of the gains allows to change intensity of ultrasound used; increasing intensity of the echoes enhances the echogenicity of examined structures.
- Slicing: the higher is number of images acquired in time unit, the higher is spatial resolution of the image.
- Shooting angle: US incidence angle should be 90° to optimize the spatial resolution, and to reduce refraction.

4.2 Anatomical changes

- Previous polypectomy or biopsy: inflammation that results causes an infarction of rectal wall which obliterates anatomical planes.
- Presence of peritumoral inflammatory exudate: it appears as hypoechoic as tumor, so switch between edema and tumor is not very definable.
- Presence of bleeding (hematoma post-biopsy), also hypoechoic like tumor mass.
- Neoadjuvant chemoradiotherapy: it reduces ERU specificity to only 46% for T parameter [9]. Reduction in specificity is related to edema and fibrosis that follows radiation therapy. Edema, that is hypoechoic, can not be distinguished from neoplasia, leading to an overstaging; however, an experienced operator may underestimate real neoplastic infiltration. It can be interpreted as inflammatory exudate. The desmoplastic reaction of the tumor resulting from radiation is visible; it reduces the gain of ultrasound probe, similar to a hyper-echoic spicula. Histopathological examination shows that in these cases residual tumor, when present, is always within the fibrotic reaction, not outside or separate from this [10]. An hypoechoic flange around the tumor may persist for 12 months or more after radiotherapy [11].

4.3 Interpretation mistakes (most frequent error for N parameter [12].)

4.3.1 Hypoechoic images (eg: female urethra, vessels, seminal vesicles, small intestine)

- These can be interpreted as metastatic lymph nodes.

4.3.2 Ileum

- In Douglas can be seen as a continuation of rectal wall.
Generally, keep in mind that understaging a tumor has severe clinical consequences and is almost always due to wrong interpretation of the images by operator: for this reason, most operators tend to give an increased stage to uncertain backdrop.

4.4 False images (ghosting)

4.4.1 Diffraction

- When the probe is not perpendicular to the rectal wall. Areas in which is more difficult to maintain a 90 degree angle with rectal wall are: post-anal area, just distal to rectumRectosigmoid junctionHouston
valves

4.4.2 Reverberation

- (Figure 2) occurs when ultrasonic beam is massively and repeatedly reflected (by air, metal clips, or feces with air) against the surface of transducer. This creates multiple echoes in images, equidistant from each other, which continue to reflect on tissue and on probe, repeatedly, causing formation of images equal to themselves, equidistant from each other.

4.4.3 Comet tail images

- Usually due to presence of air on the uneven surface of polypoid lesions that causes an echoes reverberation with the formation of hyperechoic strip, similar to a comet tail.

4.4.4 Refraction

- It is determined when the ray of ultrasound impact the surface with an angle different from 90° angle.

4.4.5 Thinning

- Fulcrum of the balloon creates a thinning of the wall and a shadow which can lead to a false image of tumor infiltration: reposition of balloon is necessary to get the right picture of rectal wall [13].

4.4.6 Mirror images

- Mirror image looks like a virtual object similar to the real image on the opposite side of structures, resulting however, hypoechoic and more distorted than original structure for US absorption.

4.5 Inevitable mistakes

- Images that don’t observe malignancy criteria;
- Hypoechoic images greater than 6 mm similar to tumor can be interpreted as metastatic lymph nodes.

4.6 Features of cancer

4.6.1 Dimension

- A massive neoplasia increases attenuation of US beam in deeper tissues, with increased risk of overstaging T parameter.

4.6.2 Location

4.6.2.1 Level
Tumors of upper rectum are poorly visualized with ERU because of the length of probe and anoscope. According to Sentovich [14], ERU accuracy is 50% for lesions at 6 cm or more from anus. Probability of error here rises because of difficulty to maintain the probe at 90° degrees to rectal wall, especially posteriorly. Moreover at rectosigmoid junction, it is difficult to apply a proper compression against rectal wall and balloon can not be stretched enough to get in touch with rectal wall to of rectal bulb.
Artifacts caused by other adjacent intestinal loops can cause errors in reports.

4.6.2.2 Houston valves
Balloon can compress the Houston valve on rectal wall, increasing thickness. This creates an image difficult to interpret with potential confusion between a Houston valve and a layer of rectal wall.

4.6.2.3 Lateral wall
Tumors of right wall require examination conduction in right lateral decubitus to improve interface between lesion and probe.

4.6.3 Stage of cancer

- T1 / T2:
  - Layers object of study are closer to the probe, compared to deeper layers: interpretation of images may be more accurate using higher emission frequencies compared to those used for evaluation of deeper layers.
  - Tissue is soft, so it can be easily compressed; this can lead to an overstaging induced by distension of the balloon that obliterates interface between submucosa and muscularis propria.

- T3/T4:
  - Layers of rectal wall invaded by neoplasia are located further away from the probe: examination thus requires to use lower frequencies compared to T1/T2 stages.
  - Tissue is stiff: staging error can be induced by insufficient relaxation of the balloon

4.6.4 Evaluation of iliac lymph nodes

- Not possible with rigid probes. Flexible probes are able to detect metastatic lymph nodes in iliac region: up to 28% of the distal N + tumors are associated with iliac lymph node metastasis, and 6% of patients have lymph node metastasis only in iliac area. Presence of metastasis in this area is considered a distal metastasis (M1), changing therapeutic procedure for patient.

Our experience shows that most frequent staging error of N parameter is often related to an oval hypoechoic image increased in volume, but always less than 5 mm in diameter, which is considered erroneously as a reactive lymph node.

5 Conclusion

ERU is the most accurate technique for rectal cancer staging, with an accuracy of 72-97%. Most errors occur differentiating T1 stage from T2 stage and T2 stage from T3 stage lesions; this distinction is clinically relevant because the treatment options are different. The sensitivity of ERU for detection of lymph node involvement is relatively low (59-89%) [15]. This low sensitivity is partly due to the lack of specific imaging criteria for lymph node involvement; node size and appearance are currently used. The results of tumor staging with ERU are strongly dependent on the experience of the operator. ERU tumor staging accuracy depends on the experience of the operator. ERU limits are: inability to stenotic or upper rectum tumor analysis; patient discomfort; operator dependence; limited ultrasound penetration depth; false images produced by biopsy outcomes, inflammation, hematoma or tumor shape [17-19]; margin evaluation accuracy could be affected by inadequate bowel preparation or bulky tumor that lie outside the probe focal length [20].

ERU learning curve is denied by Badger et al [22], supporting that experience doesn’t affect the staging accuracy, contrarily other authors [23,24]: Orrom et al [25] report an increased diagnostic accuracy from 59.3% and 95% after three years of practice. Different studies support this data, showing an increased accuracy from 50% to over 90% with practice [26,27], so ERU centralization could improve staging accuracy [21]. Furthermore, in order to send the patient at surgery only in case of need, being able to rely on surgical [28-48] and diagnostic procedures [49-53] increasingly fine and accurate, the course of study with an accurate learning curve turns out to be fundamental [54-66].

We can conclude that ultrasound evaluation of ano-rectal and perirectal tissues is technically challenging and requires a long LC. Our LC can not be closed down, at least for N parameter, which is still too high the risk of over- and under-staging. It is important, therefore, to continue to compare our ultrasound reports with the pathological report to learn from our mistakes, always maintaining a high level of criticism in studying ultrasound image. Awareness of the technical and anatomic factors that produce over- or underestimation of depth of tumor invasion will allow more accurate tumor staging and thus facilitate clinical management.
Conflict of interest statement: Authors state no conflict of interest.

References


