Diagnostic Performance of Magnetic Resonance Imaging for Preoperative Local Staging of Penile Cancer: A Systematic Review and Meta-Analysis

Rocco Simone Flammia 1,2,4,*, Antonio Tufano 1,3,†, Luca Antonelli 1‡, Arianna Bernardotto 1, Alberto A. Castro Bigalli 3, Zhen Tian 2, Marc C. Smaldone 3, Pierre I. Karakiewicz 2, Valeria Panebianco 4 and Costantino Leonardo 1

1 Department of Maternal-Child and Urological Sciences, Sapienza University Rome, Policlinico Umberto I Hospital, 00162 Rome, Italy; antonio.tufano91@gmail.com (A.T.); Lucaantonelli.md@gmail.com (L.A.); aribernardotto@gmail.com (A.B.); costantino.leonardo@uniroma1.it (C.L.)
2 Cancer Prognostics and Health Outcomes Unit, Division of Urology, University of Montréal Health Center, Montréal, QC H4A 3J1, Canada; zhe.tian24@gmail.com (Z.T.); pierrekarakiewicz@gmail.com (P.I.K.)
3 Department of Urology, Fox Chase Institute, Temple University, Philadelphia, PA 19122, USA; Alberto.castrobigalli@fccc.edu (A.A.C.B.); antoniotufano@yahoo.com (M.C.S.)
4 Department of Radiological Sciences, Oncology and Pathology, Sapienza University of Rome, 00162 Rome, Italy; Valeria.panebianco@uniroma1.it
* Correspondence: roccosimone92@gmail.com
† The authors equally contributed to the manuscript.

Abstract: Invasion of the tunica albuginea (TA) and/or urethra are key factors in determining the feasibility of organ-preserving surgery in penile cancer (PC). Magnetic resonance imaging (MRI) appeared to be a promising technique for preoperative local staging. We performed a systematic review (SR) and pooled meta-analysis to investigate the diagnostic performance of MRI in preoperative local staging of primary PC. An SR up to May 2021 was performed according to the PRISMA statement. The diagnostic performance of MRI was evaluated according to TA invasion, urethra invasion, and pT-stage ≥ 2. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) from eligible studies were pooled and summary receiver operating characteristic (SROC) curves were constructed. Overall, seven qualified studies were deemed suitable. Diagnostic performance of MRI showed an accuracy of 0.89 for TA invasion (sensitivity 0.78, PPV 0.79, specificity 0.86, and NPV 0.93); an accuracy of 0.90 for pT≥0.91, and NPV 0.90); an accuracy of 0.88 for urethra invasion (sensitivity 0.65, PPV 0.46, specificity 0.86, and NPV 0.93); an accuracy of 0.90 for pT ≥ 2 (sensitivity 0.86, PPV 0.84, specificity 0.70, and NPV 0.73). Currently available evidence indicates that MRI might be a one-stop shop for local staging of primary PC and play a central role with regard to conservative surgical management.

Keywords: penile neoplasms; magnetic resonance imaging; organ sparing surgery; penile amputation; radical penectomy

1. Introduction

Penile cancer (PC) is a relatively rare condition representing only 0.2% of newly diagnosed cancers in 2020, according to global cancer statistics [1]. When diagnosed at initial stage, PC can be cured in over 80% of cases [2]. Guidelines recommend using organ-preserving surgery whenever possible, since more aggressive local treatment can be devastating for the patient’s psychological well-being. Invasion of the corpus spongiosum, corpus cavernosum, and urethra are key factors in determining the feasibility of organ-preserving surgery [3]. Physical examination still represents the corner-stone when evaluating local extension of PC and ultra-sound may provide additional support [4]. Nonetheless, magnetic resonance imaging (MRI) appears to be a promising technique for preoperative local staging [5]. Unfortunately, studies investigating the role of MRI in PC...
are limited by small sample size. To address this void, we performed a systematic review (SR) and pooled meta-analysis. We hypothesized that MRI represents a reliable and highly accurate imaging modality for preoperative local staging of primary PC.

2. Materials and Methods

2.1. Search Strategy

In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [6], a systematic review of the literature was performed up to May 2021 by searching PubMed, Embase, Cochrane Central Search Library, and Web of Science with no restriction of time or language. We relied on the following search strategy: (penile neoplasms [MeSH Terms]) OR (neoplasms, penis) OR (penis neoplasms) OR (neoplasm, penis) OR (penis neoplasm) OR (neoplasms, penile) OR (neoplasm, penile) OR (penile neoplasm) OR (cancer of penis) OR (penis cancers) OR (cancer of the penis) OR (penis cancer) OR (cancer, penis) OR (cancers, penis) OR (penile cancer) OR (cancer, penile) OR (cancers, penile) OR (penile cancers) AND (magnetic resonance imaging).

This SR has been registered on Prospero (CRD42021257640).

2.2. Inclusion and Exclusion Criteria

The population, intervention, comparison, outcome, and study design principle (PICOS) was adopted to define study eligibility. We considered eligible for this SR both prospective and retrospective cohort studies (study design), reporting diagnostic accuracy in local staging (outcome) of preoperative pelvis MRI (intervention) vs. final histopathological features (comparison) in PC patients (population).

We excluded studies that did not provide clinical or pathological TNM staging, studies with overlapping patients, and case reports.

Two researchers (A.T. and L.A.) independently and in duplicate explored online databases, applying the above criteria, then, all authors independently reviewed the full text of the remaining articles to determine their final inclusion. Possible conflicts were resolved by discussion or with an independent arbiter (R.S.F).

2.3. Methodological Quality Assessment

Two independent reviewers assessed the methodological quality including risk of bias and applicability of each study according to the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [7]. Assessment of study bias and applicability was evaluated as low, high, or unclear (Supplementary Table S1). Disagreements were resolved through a discussion and, if necessary, arbitration by another reviewer.

2.4. Extractable Data

The diagnostic performance was evaluated based on three different endpoints. Of those, invasion of tunica albuginea (TA) and invasion of urethra were provided within the eligible studies. Additionally, based on data reported by the authors, we defined a new endpoint, namely pathological T-stage ≥ 2 (pT ≥ 2). According to different endpoints, true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values were extracted, whenever possible, from each study and reported in 2 × 2 contingency tables. Specifically, for the endpoint consisting in pT ≥ 2: clinical T-stage ≤ 1 (cT ≤ 1) on MRI and pT ≥ 2 at final histology was defined as FN; cT ≥ 2 on MRI and pT ≤ 1 at final histology was defined as FP; finally, TP or TN were defined, in case of agreement between cT ≥ 2 and pT ≥ 2 or cT ≤ 1 and pT ≤ 1, respectively. Subsequently, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for each endpoint. All the included studies reported pathological T-stage based on TNM 6th or 7th edition [8,9].

Other descriptive variables were extracted: study design, year of study, number of patients, median age, histology, tumor size, number of readers, use of intra-cavernous PGE1 injection, MRI machine technical features, and inter-reader agreement (K value).
2.5. Statistical Analysis

To address diagnostic performance of MRI according to the three different endpoints, sensitivity, specificity, PPV, and NPV from different studies were pooled and summary receiver operating characteristic (SROC) curves were constructed. We applied a random effects model [10] accounting for both heterogeneity of data and small sample size of the eligible studies. The meta-analysis was performed using Meta Disc v. 1.4 (Unit of Clinical Biostatistics, Madrid, Spain) [11]. All statistical tests were two-sided, and statistical significance was defined as $p < 0.05$.

2.6. Sensitivity Analyses

We also conducted a sensitivity analysis to evaluate diagnostic performance of MRI, when PGE1 intra-cavernous injection was used, since this approach is recommended whenever possible. Sensitivity analysis was applicable only to TA invasion and in pT-stage $\geq 2$ endpoints since eligible studies addressing urethra invasion did not use PGE1 intra-cavernous injection.

3. Results

3.1. Characteristics of Included Studies

The initial literature search yielded a total of 468 articles. After removing duplicate records, 347 articles were included for title and abstract screening. Subsequently 148 studies were identified for further full-text evaluation. Finally, seven qualified studies were deemed suitable for this meta-analysis (Figure 1).

![Study flow chart](image)

**Figure 1.** Study flow chart.

Characteristics of eligible studies are summarized in Table 1. Of those, four were prospective and three were retrospective. Studies were conducted in Italy ($n = 3$), UK ($n = 2$), Netherlands ($n = 1$), and Brazil ($n = 1$) and published between 2003 and 2016.

| Author | Year | Institution | Study Design | Patients | Age | Tumor Size (cm) | MRI System | Use of PGE1 | Number of Reader/Experience | Histopathological Type | Agreement (K Value) |
|--------|------|-------------|--------------|----------|-----|----------------|------------|-------------|---------------------------|---------------------|-------------------|
| Lont et al. | 2003 | Single | Retrospective | 28 | NA | NA | 1.5 T-Siemens Gmbh, Germany | NA | yes | 1.5 | SCC (100%) | NA |
| Scardino et al. | 2004 | Single | Prospective | 9 | Mean 55.1; Range (47–68) | NA | 1 T (Signa Horizon Lx, General Electric, Milwaukee, WI, USA) | yes | NA | 8 SCC (88.9%); 1 S (11.1%) | NA |
| Kayes et al. | 2006 | Single | Retrospective | 55 | Mean 61; Range (32–85) | 1.5 T-Siemens (Erhlangen Germany) | yes | 2 | SCC (100%) | 0.75 |
| Petralia et al. | 2008 | Single | Prospective | 13 | Mean 55.4; Range (47–49) | 1 T (Signa Horizon Lx, General Electric, Milwaukee, WI, USA) | yes | 1 | SCC (92.3%); 1 S (7.7%) | NA |
| Hanchanalhe et al. | 2015 | Single | Retrospective | 100 | Median 65; Range (39–88) | NA | 56 yes, 44 no | 2 | NA | 0.54 for tunica albuginea; 0.43 for urethra |
| Author          | Year | Institution | Study Design | Patients | Age        | Tumor Size (cm) | MRI System | Use of PGE-1 | Number of Reader/Experience | Histopathological Type | Agreement (K Value) |
|-----------------|------|-------------|--------------|----------|------------|----------------|-------------|--------------|----------------------------|----------------------|---------------------|
| Lont et al.     | 2003 | Single      | Retrospective | 28       | NA         | NA             | 1.5 T-Siemens Gmbh, Germany | NA           | 1                         | 28 SCC (100%)          | NA                  |
| Scardino et al. | 2004 | Single      | Prospective  | 9        | Mean 55.1; Range (47–68) | NA             | 1 T (Signa Horizon Lx, General Electric, Milwaukee, WI, USA) | yes           | NA                        | 8 SCC (88.9%); 1 Sarcoma (11.1%) | NA                  |
| Kayes et al.    | 2006 | Single      | Retrospective | 55       | Mean 61; Range (32–85) | NA             | 1.5 T-Siemens (Erlangen Germany) | yes           | 2                         | 55 SCC (100%)         | 0.75                |
| Petralia et al. | 2008 | Single      | Prospective  | 13       | Mean 55.4; Range (47–49) | NA             | 1 T (Signa Horizon Lx, General Electric, Milwaukee, WI, USA) | yes           | 1                         | 12 SCC (92.3%); 1 Sarcoma (7.7%) | NA                  |
| Hanchanale et al.| 2015 | Single      | Retrospective | 100      | Median 65; Range (39–88) | NA             | NA          | 56 yes, 44 no | 2                         | NA                   | 0.54 for tunica albuginea 0.43 for urethra |
| Bozzini et al.  | 2016 | Multiple    | Prospective  | 200      | Mean 67.35; S.D (15.45); Range (51–82) | Mean 3.12; S.D.(1.21); Range (1.2–6.1) | 1.5 T-Siemens Gmbh, Germany | yes           | 1                         | 200 SCC (100%)        | NA                  |
| Lucchesi et al. | 2016 | Multiple    | Prospective  | 25       | Mean 58.9; Range (43–89) | Mean 3 cm; Range (1–7.4 cm) | 1.5 T (Signa, GE Medical Systems, Milwaukee, WI; or Achieva, Philips Healthcare, Best, The Netherlands); 3T (Intera, Philips Healthcare) | no            | 2                         | 25 SCC (100%)         | 0.52                |
A total of 430 patients were analyzed in this meta-analysis. The mean patient age ranged from 55 to 67 years. Mean tumor size was reported only in two studies, ranging from 1 to 7 cm. The most common histopathological type was squamous cell carcinoma (SCC), as reported in six of the eligible studies and ranging from 88.9% to 100% of patients. Overall, six out of seven studies reported the type of magnetic field used, ranging from 1 to 3 Tesla. Moreover, three studies relied on two radiologist readers, three had a single reader, and one study did not provide any information. Finally, PGE1 injection was used in five out of seven studies.

3.2. Diagnostic Accuracy of MRI in Detecting Invasion of Tunica Albuginea (TA)

Four studies [12–15] evaluated the accuracy of MRI in detecting invasion of TA. Here, pooled sensitivity and specificity were 0.78 (95% CI: 0.69–0.86) and 0.91 (95% CI: 0.86–0.94), respectively. Additionally, pooled PPV and NPV were 0.79 (95% CI: 0.70–0.87) and 0.90 (95% CI: 0.86–0.94), respectively. The SROC curve depicted diagnostic accuracy of 89% (Figure 2).

![Figure 2](image)

Figure 2. Pooled sensitivity, specificity, and SROC curve of MRI in detecting invasion of TA.

After sensitivity analysis (Supplementary Figure S2a,b), we observed pooled sensitivity of 0.75 (95% CI: 0.63–0.84), specificity of 0.99 (95% CI: 0.95–1.00), PPV of 0.96 (95% CI: 0.87–1.00), and NPV of 0.90 (95% CI: 0.84–0.94).

3.3. Diagnostic Accuracy of MRI in Detecting Invasion of the Urethra

Only two studies [15,16] evaluated the accuracy of MRI in detecting urethral invasion. Here, pooled sensitivity and specificity were 0.65 (95% CI: 0.41–0.85) and 0.86 (95% CI: 0.78–0.92), respectively. Additionally, pooled PPV and NPV were 0.46 (95% CI: 0.28–0.66) and 0.93 (95% CI: 0.86–0.97), respectively. The SROC curve depicts diagnostic accuracy of 88% (Figure 3).

![Figure 3](image)

Figure 3. Pooled sensitivity, specificity, and SROC curve of MRI in detecting urethral invasion.
3.4. Diagnostic Accuracy of MRI in Detecting Pathologic T-Stage ≥ 2

Four studies [12,14,17,18] evaluated the accuracy of MRI in detecting pT ≥ 2. Here, pooled sensitivity and specificity were 0.86 (95% CI: 0.73–0.95) and 0.70 (95% CI: 0.47–0.87), respectively. Additionally, pooled PPV and NPV were 0.84 (95% CI: 0.71–0.94) and 0.73 (95% CI: 0.50–0.89), respectively. The SROC curve depicted diagnostic accuracy of 90% (Figure 4).

Figure 4. Pooled sensitivity, specificity, and SROC curve of MRI in pathologic T-stage ≥ 2.

After sensitivity analysis (Supplementary Figure S3a,b), we observed pooled sensitivity of 0.96 (95% CI: 0.80–1.00), specificity of 0.71 (95% CI: 0.44–0.90), PPV of 0.83 (95% CI: 0.64–0.94), and NPV of 0.92 (95% CI: 0.64–1.00). The SROC curve depicted diagnostic accuracy of 98% (Supplementary Figure S4).

4. Discussion

Surgical treatment is often curative for localized PC and preoperative staging is essential to refine decision-making strategy. Based on contemporary literature, pelvic MRI appears to be a promising imaging tool to assess tumor size and local extension [19]. In consequence, our aim was to systematically review the role of MRI as a reliable and accurate imaging modality for preoperative local staging of primary PC. Our analyses led to several noteworthy findings.

First, diagnostic performance of MRI in detecting TA invasion showed an accuracy of 0.89 with a sensitivity of 0.78, a PPV of 0.79, a specificity of 0.91, and an NPV of 0.90. In consequence, MRI may represent a reliable tool to exclude TA invasion. This result is mainly imputable to the study from Bozzini et al. [13], which relied on the greatest sample size (n = 200) and reported an interesting high specificity of 0.99 (95% CI: 0.95–1.00).

Second, diagnostic performance of MRI in detecting urethral invasion showed an accuracy of 0.88 with a sensitivity of 0.65, a PPV of 0.46, a specificity of 0.86, and an NPV of 0.93. In consequence, MRI may represent a reliable tool to exclude urethral invasion. Here, the study from Hanchanale et al. [15] was the most informative (n = 100) and reported an interesting high NPV of 0.92 in urethral invasion.

Third, we tested the diagnostic performance of MRI in detecting pT ≥ 2. This endpoint was calculated basing on extractable data from eligible studies (four out of six), as reported in Section 2.4. We observed a diagnostic accuracy of 0.90 with a sensitivity of 0.86, a PPV of 0.84, a specificity of 0.70, and an NPV of 0.73. These results showed that MRI accurately identifies patients with local extension beyond subepithelial connective tissue.

Taken together, our analyses supported the role of MRI in preoperative staging of localized PC. Specifically, MRI adequately performs in excluding TA and urethral invasion, as well as in detecting extension beyond subepithelial connective tissue. These observations are of utmost importance in preoperative patient counselling, when planning organ-sparing surgery in primary PC. As stated by EAU guidelines, in patients with invasion of both
corpus cavernosum and urethra, partial or total penectomy with perineal urethrostomy is strongly recommended, since these patients yield an increased risk of local recurrence (35%) and a 5-year mortality of 30% [20]. In consequence, the ability of MRI to exclude both TA and urethral invasion may allow clinicians to refine surgical management [21]. Despite our encouraging findings, one could argue that MRI is an expensive technique. However, over the last decade, MRI has been increasingly adopted in preoperative staging of genito–urinary cancers and is currently widely available in most centers [22,23]. Nonetheless, the use of this imaging modality is still limited by relative or absolute contraindications such as cardiac implantable electronic device, metallic intraocular foreign bodies, implantable neurostimulation systems, coronary and peripheral artery stents, inferior vena cava filters, and penile protheses [24].

To best of our knowledge, we are the first to systematically review and meta-analyze the literature addressing the diagnostic performance of MRI in preoperative local staging of PC patients. In consequence, our results cannot be directly compared to others. However, our study is not devoid of limitations. First, only seven studies were identified and three of them were retrospective. Nonetheless, it is not negligible that PC is a relatively rare condition. Second, there was heterogeneity in the MRI magnetic field strength among eligible studies. This could lead to biased interpretation of T1- and T2-weighted sequences, which are the most useful in determining local extent of penile cancer. Moreover, only three studies relied on more than one reader to interpret images [12,14,15]. Nonetheless, these studies reported a kappa inter-observer agreement ranging from moderate to good [25]. Lastly, one could argue that MRI without induced erection with PGE1 could affect the pooled diagnostic performance. To deal with this issue, we performed a sensitivity analysis, that did not markedly change our findings.

5. Conclusions

Currently available evidence indicates that MRI might be one-stop shop for local staging of primary PC and play a central role with regards to conservative surgical treatments. However, further well-designed, multicenter, prospective studies are needed to definitely confirm this impression.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/app11157090/s1, Table S1: Quality assessment of the studies according to QUADAS-2; Figure S2a,b: Sensitivity analysis of MRI in detecting invasion of TA; Figure S3a,b: Sensitivity analysis of MRI in detecting pT\(\geq 2\); Figure S4: SROC curve of MRI in detecting pT\(\geq 2\).

Author Contributions: Conceptualization, R.S.F. and A.T.; methodology, L.A.; software, A.T.; validation, L.A. and C.L.; formal analysis, Z.T.; investigation, C.L.; resources, A.T.; data curation, A.B.; writing—original draft preparation, A.T.; writing—review and editing, R.S.F.; visualization, C.L.; supervision, P.I.K.; project administration, V.P.; funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J. Clin. 2021, 71, 209–249. [CrossRef] [PubMed]
2. American Cancer Society. Survival Rates for Penile Cancer; Updated January 26; American Cancer Society: Atlanta, GA, USA, 2021.
3. EAU Guidelines. In Presented at the EAU Annual Congress Amsterdam; EAU Guidelines Office: Arnhem, The Netherlands, 2020; ISBN 978-94-92671-07-3.
