·论著·

基于第二版前列腺影像报告和数据系统评估前列腺病灶的一致性及准确性的探究

刘　伟[1,2]，刘晓航[1]，汤　伟[1]，高洪波[1]，周良平[1]

1. 复旦大学附属肿瘤医院放射诊断科，复旦大学上海医学院肿瘤学系，上海 200032；
2. 上海市影像医学研究所，上海 200032

［摘要］ **背景与目的：**随着前列腺多参数MR成像技术的发展，影像学在前列腺病灶诊断、恶性度评估及疗效评估等方面均表现出一定的临床意义。但扫描序列增多，且在前列腺不同区域、不同序列所占的权重并不一样，给临床工作带来了一定的负担，基于此基础上的第二版前列腺影像报告和数据系统（Prostate Imaging Reporting and Data System version 2，PI-RADS V2）应运而生。为推广PI-RADS V2，本研究就其对前列腺病灶评估的一致性及准确性进行探究。**方法：**回顾性分析98例符合PI-RADS V2评估要求并有病理资料的患者，共141个病灶。邀请两位不同年资的影像学医师独立评估病灶。采用kappa系数评估两名医师PI-RADS V2分类评分的一致性。利用受试者工作特征（receiver operating characteristic，ROC）曲线方法来分析两名医师诊断出Gleason score≥7的病灶的准确性，并分别计算出各ROC曲线的cut-off值来确定PI-RADS V2评分的最佳界值。**结果：**对于PI-RADS V2≥3分的病灶，两名医师对外周带病灶评估一致性较好（非外周带和外周带kappa值分别为0.668和0.769）。对于PI-RADS V2≥4分的病灶，其评估一致性更好（非外周带和外周带kappa值分别为0.710和0.843）。影像学医师1、2诊断出Gleason score≥7的病灶的ROC曲线的曲线下面积（area under curve，AUC）分别为0.816、0.792，且非外周带的AUC均比外周带的高（AUC1：0.886 *vs* 0.791；AUC2：0.791 *vs* 0.730）。另外，两名医师在评估外周带及非外周带病灶良恶性的最佳PI-RADS V2评分界值均为3分。**结论：**PI-RADS V2评价前列腺病灶具有较好的一致性及准确性。PI-RADS V2可能更适合非外周带病灶的评估，尚需进一步验证。

［关键词］ 第二版前列腺影像报告和数据系统；前列腺癌；准确性；一致性

DOI: 10.19401/j.cnki.1007-3639.2018.03.005

中图分类号：R737.25　　**文献标志码：**A　　**文章编号：**1007-3639(2018)03-0197-06

The exploration of interobserver agreement and accuracy of Prostate Imaging Reporting and Data System version 2 for the prostate carcinoma

LIU Wei[1,2], LIU Xiaohang[1], TANG Wei[1], GAO Hongbo[1], ZHOU Liangping[1] (1. Department of Radiology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China; 2. Shanghai Institute of Medical Imaging, Shanghai 200032, China)

Correspondence to: ZHOU Liangping E-mail: zhoulp2006@163.com

［**Abstract**］ **Background and purpose:** With the development of multi-parametric MR imaging techniques for detection of prostate carcinoma, medical imaging has shown the promising value in diagnosis, prediction of aggressiveness and evaluation of responses to different treatments. However, the increase of scanning sequences and the different weight of sequences in different regions of prostate put additional burden on diagnosticians, and thus the Prostate Imaging Reporting and Data System version 2 (PI-RADS V2) is generated. For popularizing PI-RADS V2, we performed this study to investigate interobserver agreement and accuracy in diagnosing prostate lesions. **Methods:** We retrospectively analyzed 98 patients with 141 prostate carcinoma lesions confirmed by biopsy and/or surgery who met the conditions of PI-RADS V2 assessment. Two readers independently assigned a PI-RADS V2 assessment category to the lesions. The Cohen's kappa statistic was used to quantify interobserver agreement. The area under the receiver operating characteristic (ROC) curve (AUC) was calculated to determine reader accuracy for the detection of clinically significant prostate cancer (Gleason score ≥7). Simultaneously, the cutoff value of all ROCs were calculated, which

通信作者：周良平　E-mail: zhoulp2006@163.com

would be regarded as the optimal value to define the prostate lesions as benign or malignant foci. **Results:** When a PI-RADS V2 assessment category ⩾3 was considered positive, the agreement between readers was good for non-peripheral zone lesions (kappa=0.668) and peripheral zone lesions (kappa=0.769). When a PI-RADS V2 assessment category ⩾4 was considered positive, the agreement was better for non-peripheral zone lesions (kappa=0.710) and excellent for peripheral zone lesions (kappa=0.843). The AUCs for readers 1 and 2 were 0.816 and 0.792, and had no significant difference. The AUCs were greater for non-peripheral zone lesions than for peripheral zone lesions (AUC1: 0.886 *vs* 0.791; AUC2: 0.791 *vs* 0.730). Additionally, the cutoff value of all ROC curves was 3. **Conclusion:** Two experienced readers were able to accurately identify patients with clinically significant prostate cancer using PI-RADS V2 with good interobserver agreement. PI-RADS V2 may be more suitable for non-peripheral zone lesions, which need further investigation.

［**Key words**］ Prostate Imaging Reporting and Date System version 2; Prostate carcinoma; Accuracy; Interobserver agreement

前列腺癌（prostate carcinoma，PCa）是中老年男性常见的恶性肿瘤之一，在美国其发病率居首位[1]。因寿命延长、饮食西化等因素，我国PCa的发病率、死亡率逐年上升[2]。血清总体前列腺特异性抗原（prostate-specific antigen，PSA）以4.0 ng/mL为基线时，诊断PCa的灵敏度仅为21%，诊断高级别的PCa的灵敏度也仅为51%[3]，虽然游离PSA百分比能更好地鉴别前列腺病灶[4]，但其价值仍然有限。而MR是诊断PCa最好的影像学方法，随着多参数扫描技术的成熟，其在PCa的定位、诊断及疗效评估等方面的价值越来越得到认可。为了规范前列腺MRI报告模式，欧洲泌尿生殖放射学协会于2012年提出了第一版前列腺影像报告和数据系统（Prostate Imaging Reporting and Data System，PI-RADS），但是该系统为累加评分模式（即病灶在MR不同序列上评分之和，范围为4~20），且前列腺波普耗时较长[6]，外周带病变诊断更依赖于弥散成像（diffusion-weighted imaging，DWI），移行带病变诊断更依赖于T2加权成像（T2-weighted imaging，T2WI）[5-6]，故该系统临床推广受限。在此基础上，采用权重模式的第二版PI-RADS（PI-RADS version 2，PI RADS V2）应运而生[7]。

目前，国内外针对PI-RADS V2的临床研究较少[2,8-11]，其诊断一致性、准确性及定位精准性等方面仍需要充分研究。本研究旨在探究不同医师之间采用PI-RADS V2来评估前列腺病灶的一致性及准确性。

1 资料和方法

1.1 一般资料

收集2016年1月—2017年7月因前列腺病变来复旦大学附属肿瘤医院就诊的患者，从中筛选出在治疗前、穿刺前或穿刺6周后行mp-MR扫描的患者，扫描参数满足PI-RADS V2评估要求，并且这些患者均有超声引导下细针穿刺和（或）手术病理资料，共98例患者。

1.2 MR扫描参数

采用德国西门子公司3.0 T MR扫描仪。患者仰卧，并用绑带固定线圈以减少呼吸伪影。动态对比度增强成像（dynamic contrast enhancement imaging，DCE）采用马根维显造影剂（德国Bayer公司），经肘静脉注入（剂量15 mL，流速2 mL/s），其余扫描参数见表1。

1.3 数据获取

邀请两位系统学习过PI-RADS V2的影像学医师1、2（资历分别为7年、1年）进行独立阅片。根据PI-RADS V2的整体评分标准，各评分对应为PCa的可能性[12]如下：1分，非常低，极不可能；2分，低，不可能；3分，中等，可疑；4分，高，可能；5分，非常高，极有可能。图1、2分别为非外周带及外周带病灶PI-RADS V2分类评分的示例。通过超声引导下细针穿刺和（或）手术获得病理资料。因Gleason score=6的病灶表现为惰性，临床不建议积极干预[13]，因此，我们将此类病灶归类至“良性”病灶中，以便更好地研究需积极干预的病灶（Gleason score⩾7）。

表 1 PCa的mp–MR扫描参数

Tab. 1 The imaging protocol of mp-MR for PCa

Parameter	T1WI	T2WI	DWI	DCE
Sequence	GRE	TSE	Resolve	VIBE
TR *t*/ms	231	9040	7400	3.56
TE *t*/ms	2.46	89	62/99	1.39
Flip angle	70	160	180	9
Metrics	320×320	320×320	200×180	350×350
Field of visio A/mm^2	350×350	200×200	200×200	350×350
Number of excitation	2	2	1/2	1
Slice thickness *d*/mm	5	3.5	3.5	3
B values/(s·mm^{-2})	-	-	0/1 500	-

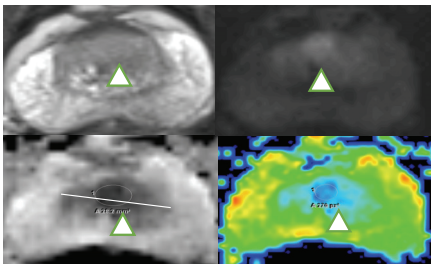

图1 非外周带PI–RADS评分为5的病灶示例

Fig. 1 An example of PCa in non-peripheral zone with PI-RADS of 5

A patient with PSA of 25.64 ng/mL was 64 years old. The lesion was located at central gland with the longest diameter of 23mm. It had low signal intension on T2WI and ADC map, and high signal intension on DWI, of which PI-RADS score was 5 and finally Gleason score was 8 according to surgical pathology

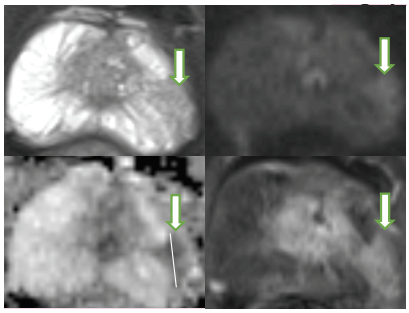

图2 外周带PI–RADS评分为4的病灶示例

Fig. 2 An example of PCa in peripheral zone with PI-RADS of 4

A patient with PSA of 16.32 ng/mL is 58 years old. The lesion is located at left peripheral zone of prostate with the longest diameter of 12 mm. It is low signal intension on T2WI and ADC map, high signal intension on DWI, and has obvious enhancement on DCE, of which PI-RADS score is 4 and finally Gleason score is 7 on basis of biopsy pathology

1.4 统计学处理

所用统计软件为SPSS 20.0及Medcalc 15.2.2。用kappa系数来评估两位影像学医师的诊断一致性：kappa≤0.40，差；0.40<kappa≤0.60，中等；0.60<kappa≤0.80，较好；0.80<kappa≤1.00，极好。利用受试者工作特征（receiver operating characteristic，ROC）曲线的方法来比较两位医师依据PI-RADS V2对前列腺病灶的诊断效能，并探究PI-RADS V2区分病灶良恶性的最佳界值，P<0.05为差异有统计学意义。另外，分析PI–RADS V2≥3分、≥4分为阳性时的灵敏度及特异度，以进一步研究PI-RADS V2。

2 结 果

患者中位年龄为68岁（38~83岁），仅有良性病灶的患者中位PSA为9.44 ng/mL（2.53~52.54 ng/mL），有PCa的患者中位PSA为15.08 ng/mL（4.67~278.00 ng/mL），患者的病理资料见表2。98例患者共有141个病灶（定义病灶：在T2WI上为异常信号，并且得到两名医师的认可）。

对于PI-RADS V2≥3分的病灶，两名医师对外周带病灶评估一致性较好（非外周带、外周带kappa=0.668、0.769），对于PI-RADS V2≥4分的病灶，其评估一致性更好（非外周带和外周带kappa值分别为0.710和0.843，表3）。

两名医师基于PI-RADS V2诊断前列腺病灶的准确性均较好［曲线下面积（area under curve，AUC）：0.730~0.886］，且对同一区域病灶的诊断水平差异无统计学意义（P > 0.05），PI-RADS V2区分良恶性的最佳界值为3分（图3）。PI-RADS V2≥3分、≥4分为阳性时的灵敏度及特异度见表4。

表 2 患者的病理资料

Tab. 2 The pathological data of patients

Item	Peripheral zone	Non-peripheral zone
Benign lesions[#]	26	39
Malignant lesions	46	30
Gleason score		
7	23	13
8	12	9
9	9	6
10	2	2

[#]: Benign lesions here include prostate carcinoma with Gleason score of not more than 6, which is suggested to active surveillance rather than active intervention.

表 3 采用PI-RADS V2诊断PCa病灶的一致性

Tab. 3 The interobserver agreement of diagnosing PCa foci based on PI-RADS V2

Kappa value	Non-peripheral zone	Peripheral zone	All lesions
PI-RADS V2≥2	0.544	0.640	0.612
PI-RADS V2≥3	0.668	0.769	0.747
PI-RADS V2≥4	0.710	0.843	0.840

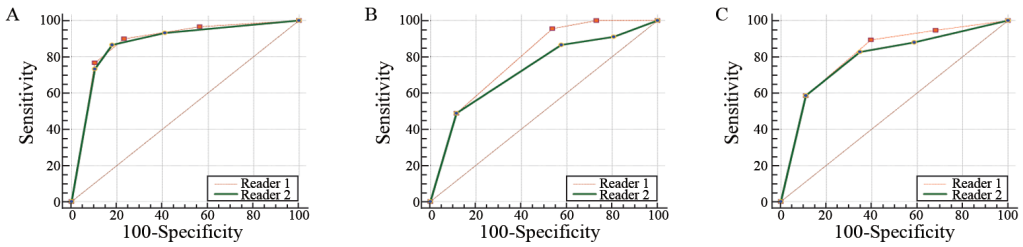

图 3 两名医师在PI-RADS V2基础上诊断各区PCa的ROC曲线分析比较

Fig. 3 ROC curves for two readers utilizing PI-RADS V2 to predict the presence of PCa confirmed by biopsy and/or surgery in different regions

The areas under ROC curves of depicting the diagnostic ability in non-peripheral, peripheral and total regions for radiologist 1 and 2 are as the following: 0.886 *vs* 0.791, *P*=0.828 6; 0.791 *vs* 0.730, *P*=0.230 1; 0.816 *vs* 0.792, *P*=0.443 1. A: Evaluation of non-peripheral zone lesions; B: Evaluation of peripheral zone lesions; C: Evaluation of the overall lesions

表 4 采用PI-RADS V2诊断PCa病灶的灵敏度及特异度

Tab. 4 The sensitivity and specificity of diagnosing PCa foci based on PI-RADS V2

Item	PI-RADS V2≥3		PI-RADS V2≥4	
	Sensitivity/%	Specificity/%	Sensitivity/%	Specificity/%
Radiologist 1				
Non-peripheral zone	96.67	43.59	90.00	76.92
Peripheral zone	100.00	26.92	95.56	46.15
All prostate	94.67	31.75	89.33	60.32
Radiologist 2				
Non-peripheral zone	93.33	58.97	86.67	82.05
Peripheral zone	91.11	19.23	86.67	42.31
All prostate	88.00	41.27	82.67	65.08

3 讨 论

本研究结果显示，两名医师对PI-RADS V2≥3分的病灶评估一致性较好（kappa=0.747），与Kasel-Seibert等[14]（kappa=0.68）及Polanec等[15]（kappa=0.71）的研究结果一致。但Baldisserotto等[16]和Muller等[17]的研究提示，PI-RADS V2评分仅呈中等程度一致性（kappa值分别为0.53和0.46）。另外，对于PI-RADS V2≥4分的病灶评估一致性极好（kappa=0.840），与Park等[18]的研究结果

相近（kappa=0.801），但明显好于Rosenkrantz等[19]的研究（kappa=0.552）。我们详细分析了造成这种情形的原因。首先，所邀请的影像学医师数量对此影响较大。Kasel-Seibert、Polanec、Park及本研究均邀请两名医师作为评估医师，这些研究的结果相似，其诊断一致性均比较好。Muller邀请了5位医师，Rosenkrantz邀请了6位医师，他们所得到的kappa系数值均相对较低。即使是经过系统的训练，但医师在进行PI-RADS V2分类评分时的主观性差异仍无法消除[8]，评估人数越多，累积差异越大，一致性越小。另外，不同机构所用的设备及成像方案并非完全一致，这可能对评估的结果也有一定影响，在Rosenkrantz等[19]的研究中，所邀请的6位医师并非来自同一个机构，这些机构的前列腺成像方案不尽相同，因而在研究中所评估的病灶图像与平时临床工作中所见的图像有一定的差异，这可能也在一定程度上影响了一致性。除此之外，所邀请的医师的资历对此也有一定的影响，Muller等[17]的研究中，5位评估医师的资历从6个月到12年不等，其kappa值仅为0.46，但这一因素的影响程度并不大，在Kasel-Seibert等[14]及本研究中均有一位医师的年资在1年或1年以下，但这两个研究的kappa值均较高。并且，通过系统学习PI-RADS V2后，低年资的医师也可具有较好的诊断能力[20]，本研究中，两名医师的诊断效能差异无统计学意义（P>0.05），这也是推广PI-RADS V2的价值所在，即短期内可提高初级医师对前列腺病灶的评估能力。

Glazer等[8]、Rosenkrantz等[19]及Purysko等[20]的研究表明，PI-RADS V2评估外周带病灶的一致性优于移行带，与本研究结果一致。另外，对于PI-RADS V2评分越高的病灶，其评估的一致性也越高，这也与之前的研究一致[19-20]。对于评分≥4分的病灶，PI-RADS V2具有非常理想的诊断一致性。并且对于这些病灶，两名医师均表现出理想的诊断敏感性及特异性。与PI-RADS V2≥3分定义为阳性时相比，其诊断灵敏度变化不大，但其特异性明显升高，与之前的研究类似[20]，这也是PI-RADS V2≥4分的病灶更需积极穿刺的原因[19]。

本研究中，无论是PI-RADS V2≥3分还是≥4分定义为阳性，对整个腺体的诊断灵敏度均极好，特异度中等。Zhang等[21]对13篇与PI-RADS V2有关的文献进行Meta分析，与本研究的结果类似（灵敏度为85%，特异度为71%）。这可能与Gleason评分为6的病灶定义为“良性”有一定关系。此外，PI-RADS V2在非外周带的诊断特异性比外周带高（表4），与Zhang等[21]的研究结果一致，并且非外周带病灶的诊断准确性比外周带高，故PI-RADS V2可能更适合非外周带区域病灶的评估[15]。但不同医师对外周带的PI-RADS V2分类评分一致性更好，故PI-RADS V2是否更适合非外周带尚需进一步验证。

但本研究仍存在一些不足：① 阅片者所评估的病灶已提前确定，因此无法探究病灶定位的一致性，不利于对较为隐匿的病灶进行研究；② 所邀请的医师来自同一机构，且人数仅为两人；③ 病理资料并非都来源于手术切除的前列腺组织，有些是来源于超声引导的细针穿刺，其Gleason评分可能并不能代表整个病灶的评分。

综上所述，不同医师利用PI-RADS V2评估前列腺病灶具有较好的一致性，外周带病灶的诊断一致性略高于非外周带病灶。另外，鉴于其在非外周带区域较理想的诊断灵敏度及特异度，PI-RADS V2可能更适用于非外周带区。

［参 考 文 献］

［1］ AMERICAN CANCER SOCIETY. Cancer facts & figures 2017［J/OL］. Atlanta, GA: American Cancer Society, 2017. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2017.html.

［2］ 齐金蕾, 王黎君, 周脉耕, 等. 1990-2013年中国男性前列腺癌疾病负担分析［J］. 中华流行病学杂志, 2016, 37(6): 778-782.

［3］ MEIGS J B, BARRY M J, OESTERLING J E, et al. Interpreting results of prostate-specific antigen testing for early detection of prostate cancer［J］. J Gen Intern Med, 1996, 11(9): 505-512.

［4］ OMAR J, JAAFAR Z, ABDULLAR M R. A pilot study on percent free prostate specific antigen as an additional tool in prostate cancer screening［J］. Malays J Med Sci, 2009, 16(1): 44-47.

［5］ BAUR A D, MAXEINER A, FRANIEL T, et al. Evaluation of the prostate imaging reporting and data system for the detection

of prostate cancer by the results of targeted biopsy of the prostate [J]. Invest Radiol, 2014, 49(6): 411–420.

[6] SCHIMMOLLER L, QUENTIN M, ARSOV C, et al. Predictive power of the ESUR scoring system for prostate cancer diagnosis verified with targeted MR–guided in–bore biopsy [J]. Eur J Radiol, 2014, 83(12): 2103–2108.

[7] WEINREB J C, BARENTSZ J O, CHOYKE P L, et al. PI–RADS Prostate Imaging–Reporting and Data System: 2015, version 2 [J]. Eur Urol, 2016, 69(1): 16–40.

[8] GLAZER D I, MAYO–SMITH W W, SAINANI N I, et al. Interreader agreement of Prostate Imaging Reporting and Data System version 2 using an in–bore MRI–guided prostate biopsy cohort: a single institution's initial experience [J]. Am J Roentgenol, 2017, 209(3): W1–W7.

[9] PARK S Y, SHIN S J, JUNG D C, et al. PI–RADS version 2: preoperative role in the detection of normal–sized pelvic lymph node metastasis in prostate cancer [J]. Eur J Radiol, 2017, 91: 22–28.

[10] MATSUOKA Y, ISHIOKA J, TANAKA H, et al. Impact of the Prostate Imaging Reporting and Data System, Version 2, on MRI diagnosis for extracapsular extension of prostate cancer [J]. Am J Roentgenol, 2017, 209(2): W1–W9.

[11] 赵承琳, 高 歌, 李飞宇, 等. 探讨高低场强多参数MRI对PI–RADS(第2版)定性诊断前列腺临床显著癌的价值 [J]. 放射学实践, 2017, 32(4): 414–417.

[12] HASSANZADEH E, GLAZER D I, DUNNE R M, et al. Prostate imaging reporting and data system version 2 (PI–RADS V2): a pictorial review [J]. Abdom Radiol (NY), 2017, 42(1): 278–289.

[13] HUMPHREY P A, MOCH H, CUBILLA A L, et al. The 2016 WHO classification of tumours of the urinary system and male genital organs–part B: prostate and bladder tumours [J]. Eur Urol, 2016, 70(1): 106–119.

[14] KASEL–SEIBERT M, LEHMANN T, ASCHENBACH R, et al. Assessment of PI–RADS V2 for the detection of prostate cancer [J]. Eur J Radiol, 2016, 85(4): 726–731.

[15] POLANEC S, HELBICH T H, BICKEL H, et al. Head–to–head comparison of PI–RADS V2 and PI–RADS V1 [J]. Eur J Radiol, 2016, 85(6): 1125–1131.

[16] BALDISSEROTTO M, NETO E J, CARVALHAL G, et al. Validation of PI–RADS v.2 for prostate cancer diagnosis with MRI at 3T using an external phased–array coil [J]. J Magn Reson Imaging, 2016, 44(5): 1354–1359.

[17] MULLER B G, SHIH J H, SANKINENI S, et al. Prostate cancer: interobserver agreement and accuracy with the revised Prostate Imaging Reporting and Data System at multiparametric MR imaging [J]. Radiology, 2015, 277(3): 741–750.

[18] PARK S Y, JUNG D C, OH Y T, et al. Prostate cancer: PI–RADS Version 2 helps preoperatively predict clinically significant cancers [J]. Radiology, 2016, 280(1): 108–116.

[19] ROSENKRANTZ A B, GINOCCHIO L A, CORNFELD D, et al. Interobserver reproducibility of the PI–RADS version 2 lexicon: a multicenter study of six experienced prostate radiologists [J]. Radiology, 2016, 280(3): 793–804.

[20] PURYSKO A S, BITTENCOURT L K, BULLEN J A, et al. Accuracy and Interobserver Agreement for Prostate Imaging Reporting and Data System, version 2, for the characterization of lesions identified on multiparametric MRI of the prostate [J]. Am J Roentgenol, 2017, 209(2): 339–349.

[21] ZHANG L, TANG M, CHEN S, et al. A meta–analysis of use of Prostate Imaging Reporting and Data System version 2 (PI–RADS V2) with multiparametric MR imaging for the detection of prostate cancer [J]. Eur Radiol, 2017, 27(12): 5204–5214.

[22] WEINREB J C, BARENTSZ J O, CHOYKE P L, et al. PI–RADS Prostate Imaging–Reporting and Data System: 2015, version 2 [J]. Eur Urol, 2016, 69(1): 16–40.

（收稿日期：2017–12–15 修回日期：2018–02–05）