Published in final edited form as:

Cancer. 2013 November 15; 119(22): 3992–4002. doi:10.1002/cncr.28303.

Development and Multi-Institutional Validation of an Upgrading Risk Tool for Gleason 6 Prostate Cancer

Matthew Truong, MD¹, Jon A. Slezak, BS¹, Chee Paul Lin, MA², Viacheslav Iremashvili, MD³, Martins Sado, BS¹, Aria A. Razmaria, MD⁴, Glen Leverson, PhD², Mark S. Soloway, MD³, Scott E. Eggener, MD⁴, E. Jason Abel, MD¹, Tracy M. Downs, MD¹, and David F. Jarrard, MD¹

¹Department of Urology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

²Department of Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

³Department of Urology, Miller School of Medicine, University of Miami, Miami, Florida

⁴Department of Urology, University of Chicago Medicine and Biological Sciences, Chicago, Illinois

Abstract

BACKGROUND—Many patients with low-risk prostate cancer (PC) who are diagnosed with Gleason score 6 at biopsy are ultimately found to harbor higher grade PC (Gleason 7) at radical prostatectomy. This finding increases risk of recurrence and cancer-specific mortality. Validated clinical tools that are available preoperatively are needed to improve the ability to recognize likelihood of upgrading in patients with low-risk PC.

METHODS—More than 30 clinicopathologic parameters were assessed in consecutive patients with Gleason 6 PC upon biopsy who underwent radical prostatectomy. A nomogram for predicting upgrading (Gleason 7) on final pathology was generated using multivariable logistic regression in a development cohort of 431 patients. External validation was performed in 2 separate cohorts consisting of 1151 patients and 392 patients. Nomogram performance was assessed using receiver operating characteristic curves, calibration, and decision analysis.

RESULTS—On multivariable analysis, variables predicting upgrading were prostate-specific antigen density using ultrasound (odds ratio [OR] =229, $P=.003$), obesity (OR =1.90, $P=.05$), number of positive cores (OR =1.23, $P=.01$), and maximum core involvement (OR =0.02, $P=.01$).

Corresponding author: David F. Jarrard, MD, Department of Urology, University of Wisconsin School of Medicine and Public Health, 7037 Wisconsin Institutes of Medical Research, 1111 Highland Avenue, Madison, WI 53792; Fax: (608) 262-6453; ; Email: jarrard@urology.wisc.edu

CONFLICT OF INTEREST DISCLOSURE

Dr. Jarrard has been a consultant for Dendreon and Johnson & Johnson, and has grants from National Institutes of Health. Mr. Slezak has received grants from the National Institutes of Health and University of Wisconsin School of Medicine and Public Health. Dr. Eggener has been a consultant for Myriad Genetics and Genomic Health, has received grants from Myriad Genetics, and has received travel/meeting expenses from Myriad Genetics and Genomic Health. Dr. Downs has been a consultant for Photocure. Dr. Soloway has received travel/meeting expenses from Photocure, Astellas, and Sanofi Aventis. All other authors made no disclosure.

FUNDING SOURCES

No specific funding was disclosed.

On internal validation, the bootstrap-corrected predictive accuracy was 0.753. External validation revealed a predictive accuracy of 0.677 and 0.672. The nomogram demonstrated excellent calibration in all 3 cohorts and decision curves demonstrated high net benefit across a wide range of threshold probabilities. The nomogram demonstrated areas under the curve of 0.597 to 0.672 for predicting upgrading in subsets of men with very low-risk PC who meet active surveillance criteria (all $P < .001$), allowing further risk stratification of these individuals.

CONCLUSIONS—A nomogram was developed and externally validated that uses preoperative clinical parameters and biopsy findings to predict the risk of pathological upgrading in Gleason 6 patients. This can be used to further inform patients with lower risk PC who are considering treatment or active surveillance.

Keywords

prostate-specific antigen; low-risk prostate cancer; active surveillance

INTRODUCTION

Increasing numbers of men (35%–45%) in the late era of prostate-specific antigen (PSA) testing have been identified with low-risk prostate cancers (PCs) as defined by the D'Amico criteria (PSA <10 ng/mL, biopsy Gleason score ≤6, and stage ≤T2a).¹ In many of these patients, the cancer is unlikely to affect long-term prognosis and many may be managed expectantly with active surveillance (AS). However, a proportion of patients on AS still harbor significant cancers and can experience worse clinical outcomes.² Indeed, 30% to 50% of patients with low-risk disease experience Gleason score upgrading (GSU) at the time of radical prostatectomy (RP).^{3–5} This is clinically important because the presence of increasing quantities of Gleason pattern 4 results in an increased risk of biochemical disease recurrence, need for adjuvant therapy, and cancer-specific mortality.⁶ Therefore, refining our ability to define patients at low risk of progression after diagnosis remains an important area of research.

Several nomograms for predicting upgrading at the time of RP have been reported, but demonstrate lower predictive accuracy and poor calibration upon external validation.⁷ This may result, in part, from their use of higher risk populations for generating these nomograms. Many of these models also rely on PSA, which has been repeatedly found to be a poor predictor of GSU.^{4,8–10} Recent studies suggest that PSA density (PSAD) may be a more accurate predictor than PSA.^{2,11} Previous nomograms have also not analyzed potentially important clinical predictors of upgrading, such as prostate volume, smoking, obesity, and family history of PC.⁷ In this study, we performed a comprehensive analysis of a large number of clinical and pathological variables and built a Biopsy-Integrated Algorithm for Determining Gleason 6 Upgrading Risk (BADGR), an accurate and well-calibrated nomogram, for use in patients with Gleason score 6 PC; it represents a superior method of risk stratification compared with single cutpoints used in other AS protocols.

MATERIALS AND METHODS

Three centers were included in this institutional review board-approved study involving modern cohorts of patients. A total of 1024 patients underwent RP at the University of Wisconsin-Madison (UW) from 2005 to 2011. Among these, 644 patients were diagnosed with Gleason 6 PC using transrectal ultrasound (TRUS)-guided biopsy. None of the patients with Gleason 6 PC had neoadjuvant therapy. Only patients with biopsies of at least 10 cores (n =413) were included in nomogram development and internal validation. Although the majority of biopsy slides were re-reviewed by a genitourinary pathologist at UW, the outside pathological read was used for nomogram development to maximize generalizability of this model for use in both academic and community settings. External validation was conducted separately for patients with Gleason 6 PC with biopsies of at least 10 cores who underwent RP at 2 academic centers: the University of Chicago (UC) (n =1151) and the University of Miami (UM) (n =392). Clinical and pathological features of the patient cohorts are provided (Table 1). All outside biopsy slides were re-reviewed by genitourinary pathologists at UC and UM.

Clinical variables were collected including age, body mass index (BMI), PSA, American Urological Association (AUA) symptom score, race, family history of PC, any history of smoking, American Society of Anesthesiologists (ASA) physical status, surgical approach, clinical stage, interval between biopsy and surgery, digital rectal examination estimation of prostate size. Because patients were only included from the modern era, year of diagnosis was not evaluated. TRUS-guided biopsy was performed in all patients with the following variables collected: TRUS estimation of prostate size, biopsy-estimated tumor volume, maximum percentage core involvement, number of positive cores, total number of cores, and laterality. Biopsy-estimated tumor volume was calculated by averaging the percentage involvement of the left and right lobes of the prostate. PSAD was calculated by dividing PSA by TRUS-estimated prostate size. Previously, some studies used pathological weight as a surrogate for TRUS-estimated prostate size.^{7,12} We focused on TRUS-estimated prostate weight because pathological prostate weight is not available at the time of diagnosis and has little relevance to preoperative model development. However, we developed and validated an alternate nomogram (BADGR-P) using PSAD calculated from pathological weight (PSAD-P). The alternate nomogram had slightly lower areas under the curve (AUCs) of 0.721, 0.616, and 0.637 at UW, UC, and UM, respectively, on external validation, but calibration and decision analyses yielded similar results (data available on request).

Statistical Analysis

Univariate logistic regression was performed to explore the relationship of clinical and pathological variables with GSU (defined as Gleason ≥ 7 on final pathology). Only statistically significant clinical variables ($P < .05$) on univariate analysis were entered into the multivariable model. To avoid overfitting, biopsy variables were assessed for multicollinearity by Pearson correlation matrices. Any pairs of variables with correlation coefficient $r > 0.50$ were analyzed separately using alternate logistic regression models (LRMs) in order to determine the more significant predictor out of the pair to be included in the model. In the multivariable LRM, we applied a backward elimination procedure under a

cutoff probability of 0.05 for the likelihood ratio test to identify the best predictors for our model. Internal and external ROC analysis, calibration, and decision analysis were performed in the 413, 1152, and 392 patients. Multiple imputation was applied to the external validation cohorts of UC and UM to address missing biopsy data because it offers distinct advantages over other alternatives by accounting for the uncertainty in the missing values.¹³

To ensure that our results are not biased, internal validation was done on 200 bootstrap resamples, whereas external validation was accomplished on the 2 external cohorts. We examined the predictive accuracy by generating receiver operating characteristic (ROC) curves and calculating the AUC. Furthermore, we generated calibration curves to examine the magnitude of agreement between predicted probabilities and actual observed risk of upgrading.

Traditional performance measures do not incorporate clinical consequences and thus less directly inform clinical practice. Therefore, we employed a decision-analytic method as described by Vickers et al.¹⁴ Decision analysis provides an assessment of clinical benefit of using prediction models without requiring external data on costs, benefits, and preferences. We compared the net benefits of different strategies for GSU prediction and define the strategies as follows: 1) upgrading none, 2) upgrade all, 3) upgrading on the basis of using all predictors, upgrading on the basis of 4) obesity alone, 5) maximum percentage (%) of core involvement alone, 6) number of positive core alone, and 7) PSAD alone. Calibration curves and decisions curves were generated using R package libraries *rms* and *dca*, respectively.

Finally, we compared the predictive accuracy of BADGR against existing AS eligibility criteria by generating AUC curves. For external data sets, the average AUC from 20 imputed data sets was used. AUCs were statistically compared using the DeLong method. Statistical analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC) software. A *P* value <.05 was considered statistically significant.

RESULTS

Patient Characteristics and Association of Upgrading in Low-Risk Patients

GSU is associated with a poorer prognosis.^{3–5} In our training cohort of 413 patients at UW diagnosed with Gleason 6 PC, factors associated with upgrading was determined by analyzing multiple pathological parameters at initial biopsy (6), clinical factors (13), and pathological parameters at final pathology (15). At final pathological analysis, GSU was significantly associated with smaller pathological prostate size, and an increase in percent positive pathological slides, percent tumor volume, pathological stage, extraprostatic extension, positive margins, and bilateral disease (all *P*<.05) (Table 2). GSU approached significance for increased seminal vesicle involvement (*P*=.14) and was not associated with a positive apical margin (*P*=.2).

GSU at RP was observed in 52% of patients at UW. In our external validation cohorts, 1152 and 392 patients from UC and UM underwent upgrading in 41% and 33% of cases,

respectively (Table 1). The mean age, BMI, and PSA were similar across all cohorts. The percentage of patients with clinical stage T1c ranged from 73% to 84%. The majority of patients (90%, 90%, and 91% at UW, UM, and UC, respectively) met the D'Amico criteria for low-risk PC (PSA <10 ng/mL, biopsy Gleason score ≤6, and stage T2a).

Nomogram Development and Validation Using Preoperative Clinical and Pathological Variables

On univariate analysis, preoperative variables significantly associated with upgrading included obesity (BMI ≥30), smaller TRUS-estimated prostate size, higher number of biopsy cores with cancer, higher percentage of positive cores, higher maximum percentage core involvement, higher tumor volume, bilateral disease, and higher PSAD (all $P<.05$) (Table 3). A multivariable analysis confirmed that PSA density, obesity, number of positive cores, and maximum percent core involvement were independent predictors after backward elimination (Table 4).

PSAD was strongly predictive and yielded an odds ratio of 1.72 for every 0.1 unit increase in PSAD. This translates to a 72% increase in the odds of GSU for every 0.1 unit increase in PSAD. For obesity (BMI ≥30), there is a 90% increased likelihood of upgrading. We found that many clinical variables, including age, PSA, smoking history, family history, clinical stage, AUA symptom score, and ASA class, were not predictive of upgrading in the multivariable model (Table 3). The predictive accuracy of BADGR as demonstrated by an AUC in the ROC analysis for UW was 0.753 ($P<.0001$). The nomogram was then applied to 2 independent external cohorts from UC and UM, and these generated AUCs of 0.677 and 0.672 (all $P<.0001$) on external validation (Fig. 1).

Calibration and Decision Analysis

In addition to AUC, calibration is an important indicator of nomogram performance. Calibration analysis is used to measure how far predictions are from actual outcomes. The bias-corrected calibration plots showed virtually no departure from ideal predictions (mean absolute error of 2.6%; <1% mean squared error of the difference) (Fig. 2A). The bias-corrected calibrated values were generated from internal validation based on 200 bootstrap resamples. Low error rates (<5%) were also detected in the external validation cohorts.

On decision analysis, the nomogram demonstrated a superior net benefit for all of the threshold probabilities of GSU when compared with individual variables (Fig. 2B). In addition, high positive net benefit was observed across a wide range of threshold probabilities in all 3 cohorts, suggesting that this nomogram will be beneficial to most low-risk patients. A graphical representation of BADGR (Fig. 3) and an online risk calculator (https://www.urology.wisc.edu/research/researchers-labs/jarrard/prostate_cancer_predictor) have been provided for clinical use.

Comparative Analysis of BADGR Versus Existing AS Eligibility Criteria for the Prediction of Upgrading

Current AS protocols differ in their ability to predict pathologically insignificant cancer at RP of which upgrading is an important component.¹⁵ We compared the ability of the

BADGR nomogram and several existing AS protocols to predict upgrading in the RP specimens from our 3 low-risk cohorts.^{16–19} Although all AS protocols showed some ability to predict GSU (AUC 0.505–0.622), BADGR demonstrated a significantly higher predictive accuracy for all cohorts ranging from an AUC of 0.672 to 0.753 (all $P < .0001$) (Table 5).

BADGR Further Risk Stratifies Men Who Qualify for AS

We then assessed whether BADGR can further risk stratify men who qualify for AS using different criteria. We applied criteria for various AS protocols to our population of 1956 patients who underwent RP at UW, UC, and UM and identified from 569 to 1199 men dependent on the protocol (Table 6). In these subsets of very low-risk men, we then determined whether BADGR could further identify the men who will experience Gleason upgrading. BADGR was able to predict upgrading in patients who met various AS criteria with AUCs of 0.597 to 0.672 (all $P < .001$) (Table 6). Therefore, this nomogram has the ability to provide further stratification of patients with low-risk PC who are eligible for AS.

DISCUSSION

Development of a reliable tool for predicting GSU risk for patients with Gleason 6 PC is a priority given the large percentage of patients identified with this score. Identifying those patients at significant risk for upgrading could alter the management of patients who are considering active surveillance, especially in young patients. In this multi-institutional study, we generated a predictive nomogram with excellent calibration. In this study, we avoid the use of postoperative prostate weight, a problem with earlier studies,^{7,12,20} thus incorporating only data that is available prior to treatment. To extend the applicability of our nomogram, we compared BADGR in our 3 patient cohorts to existing AS eligibility criteria for their ability to predict GSU. BADGR was a better predictor of GSU. BADGR uses similar variables, but assesses risk based on a continuous scale, which may explain its superiority for predicting upgrading.

We evaluated a large number of pathological and clinical variables, many of which have not been previously studied. Family history, smoking history, AUA symptoms score, ASA class, or clinical stage did not predict GSU. Obesity, prostate size, and PSA density were the only clinical variables that independently predicted GSU. The use of obesity is one unique aspect of BADGR, which remained statistically significant in all LRMs. The correlation between metabolic syndrome and more aggressive PC characteristics has been noted in epidemiological studies.²¹ Obesity has been associated with GSU,²² a significant relationship that was also evident in our analyses.

The calibration curves generated for BADGR showed minimal departure from ideal predictions with an error rate of 2.6%. The result is consistent across internal and external validation. This well-calibrated nomogram is based on a parsimonious model (following elimination of nonpredictive or weakly predictive variables), affirming the accuracy of the predictions for GSU. Even at low probability thresholds, which represent patients of very low-risk disease, the nomogram still demonstrated excellent calibration. Furthermore, decision analysis demonstrated a positive net benefit when variables are used alone and the highest net benefit when variables are used in combination. Despite variations in biopsy

technique among different centers, the nomogram was predictive of upgrading across all institutions.

Smaller prostate size is a known independent predictor of GSU.^{12,20} Smaller prostates may reflect a low intraprostatic androgen status that can be associated with more aggressive PC.²³ In addition, larger prostates produce more PSA, resulting in increased detection of clinically insignificant cancers.⁹ We believe the strength of increased PSA density as a predictor of upgrading is likely related to a combination of higher PSA levels produced by larger-volume cancers, as well as the ascertainment bias associated with increased PSA detection of low-risk PC from larger prostates.²⁴ Unlike previously published upgrading models, BADGR is the first to incorporate PSA density, which has been recently suggested to be a stronger predictor of GSU than PSA.¹¹ Some criticism against TRUS estimation of prostate size has arisen due to its variability.²⁵ However, more modern data suggests a linear correlation between pathologic weight and TRUS.²³ To address this issue, we also analyzed the nomogram using pathologic weight and found TRUS estimation to be comparable to pathologic weight, although TRUS outperformed pathological weight in terms of discrimination (data not shown). We find that PSAD calculated from TRUS-estimated prostate size was one of the strongest independent predictors of upgrading in all cohorts.

Chun and colleagues previously developed a nomogram that predicts upgrading from any Gleason score. External validation using patients from all risk categories generated AUCs of 74.9 to 79.0.^{26,27} However, when applied externally to a modern cohort of low-risk Gleason 6 patients, the AUC was reduced to 0.550.⁷ This suggests that patients with Gleason 6 cancer have different predictive factors and need to be evaluated separately compared to higher risk groups. Upgrading nomograms in patients with Gleason 6 cancer developed by Capitanio et al,²⁸ Kulkarni et al,²⁹ and Moussa et al³⁰ all generated lower AUCs with lower calibration and negative net benefit upon external validation.⁷ BADGR differs from previously published nomograms in that our model accounts for preoperative prostate size and obesity, which are important clinical predictors of GSU.

Our study has several limitations. The study was performed at several large, referral-based academic centers and requires validation in community-based cohorts. For the UW cohort, we used outside pathology reads in an effort to make the applicability of this nomogram more widespread. Subjects in this study also underwent prostatectomy and therefore may not have been representative of all low-risk patients treated with AS or radiation. The use of GSU as an endpoint has limitations because Gleason score is a powerful but imperfect predictor of biochemical recurrence or survival. Increasing quantities of Gleason pattern 4 are associated with increased risk of biochemical failure and cancer-specific mortality.⁶

Better models are needed that improve our ability to predict patients who are at the very low-risk end of the spectrum and are candidates for AS. An important criterion for AS at many centers is absence of Gleason pattern 4/5 features.^{16,19} We compared the BADGR nomogram to existing AS protocols for their ability to predict GSU. BADGR demonstrated a better ability to predict the presence of Gleason pattern 4/5 in all cohorts tested. One reason may be that previously described AS protocols are based on strict cutpoints in eligibility criteria and do not allow for further stratification after criteria are met. We also demonstrate

that BADGR provides further risk stratification among very low-risk patients who qualify for AS (Table 6). We speculate that this nomogram can be used to reduce the frequency of PSA screening and surveillance biopsies in some extremely low-risk men enrolled in active surveillance. The ultimate impact of further stratification of these low-risk patients with regard to cancer-specific mortality is not yet clear; however, its true benefit may lie in reducing AS monitoring intensity. Subsequent studies will focus on validating this predictive nomogram in prospective AS cohorts.

Conclusions

BADGR predicts GSU robustly in several external, contemporary cohorts that varied in their rates of upgrading. In addition, BADGR has the benefit of assessing risk factors on a continuous scale rather than using strict cut-points. This validated nomogram can be used at the time of diagnosis to aid patients with low-risk PC who are considering treatment or observation. Although this nomogram improves risk stratification, we anticipate that incorporating biomarkers into surveillance biopsies will further improve our ability to identify individuals with low-risk PC.

References

1. Miller DC, Gruber SB, Hollenbeck BK, Montie JE, Wei JT. Incidence of initial local therapy among men with lower-risk prostate cancer in the United States. *J Natl Cancer Inst.* 2006; 98:1134–1141. [PubMed: 16912266]
2. El Hajj A, Ploussard G, de la Taille A, et al. Analysis of outcomes after radical prostatectomy in patients eligible for active surveillance (PRIAS). *BJU Int.* 2013; 111:53–59. [PubMed: 22726582]
3. Chun FK, Steuber T, Erbersdobler A, et al. Development and internal validation of a nomogram predicting the probability of prostate cancer Gleason sum upgrading between biopsy and radical prostatectomy pathology. *Eur Urol.* 2006; 49:820–826. [PubMed: 16439050]
4. Pinthus JH, Witkos M, Fleshner NE, et al. Prostate cancers scored as Gleason 6 on prostate biopsy are frequently Gleason 7 tumors at radical prostatectomy: implication on outcome. *J Urol.* 2006; 176:979–984. [PubMed: 16890675]
5. Epstein JI, Feng Z, Trock BJ, Pierorazio PM. Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. *Eur Urol.* 2012; 61:1019–1024. [PubMed: 22336380]
6. Eggener SE, Scardino PT, Walsh PC, et al. Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J Urol.* 2011; 185:869–875. [PubMed: 21239008]
7. Iremashvili V, Manoharan M, Pelaez L, Rosenberg DL, Soloway MS. Clinically significant Gleason sum upgrade: external validation and head-to-head comparison of the existing nomograms. *Cancer.* 2012; 118:378–385. [PubMed: 21732339]
8. Dong F, Jones JS, Stephenson AJ, Magi-Galluzzi C, Reuther AM, Klein EA. Prostate cancer volume at biopsy predicts clinically significant upgrading. *J Urol.* 2008; 179:896–900. [PubMed: 18207180]
9. Kassouf W, Nakanishi H, Ochiai A, Babaian KN, Troncoso P, Babaian RJ. Effect of prostate volume on tumor grade in patients undergoing radical prostatectomy in the era of extended prostatic biopsies. *J Urol.* 2007; 178:111–114. [PubMed: 17499289]
10. Tilki D, Schlenker B, John M, et al. Clinical and pathologic predictors of Gleason sum upgrading in patients after radical prostatectomy: results from a single institution series. *Urol Oncol.* 2011; 29:508–514. [PubMed: 19837614]
11. Oh JJ, Hong SK, Lee JK, et al. Prostate-specific antigen vs prostate-specific antigen density as a predictor of upgrading in men diagnosed with Gleason 6 prostate cancer by contemporary multicore prostate biopsy. *BJU Int.* 2012; 110:E494–E499. [PubMed: 22540236]

12. Davies JD, Aghazadeh MA, Phillips S, et al. Prostate size as a predictor of Gleason score upgrading in patients with low risk prostate cancer. *J Urol*. 2011; 186:2221–2227. [PubMed: 22014803]
13. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996; 91:473–489.
14. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006; 26:565–574. [PubMed: 17099194]
15. Iremashvili V, Pelaez L, Manoharan M, Jorda M, Rosenberg DL, Soloway MS. Pathologic prostate cancer characteristics in patients eligible for active surveillance: a head-to-head comparison of contemporary protocols. *Eur Urol*. 2012; 62:462–468. [PubMed: 22445138]
16. Tosoian JJ, Trock BJ, Landis P, et al. Active surveillance program for prostate cancer: an update of the Johns Hopkins experience. *J Clin Oncol*. 2011; 29:2185–2190. [PubMed: 21464416]
17. van den Bergh RC, Roemeling S, Roobol MJ, et al. Outcomes of men with screen-detected prostate cancer eligible for active surveillance who were managed expectantly. *Eur Urol*. 2009; 55:1–8. [PubMed: 18805628]
18. Whitson JM, Porten SP, Hilton JF, et al. The relationship between prostate specific antigen change and biopsy progression in patients on active surveillance for prostate cancer. *J Urol*. 2011; 185:1656–1660. [PubMed: 21419438]
19. Soloway MS, Soloway CT, Eldefrawy A, Acosta K, Kava B, Manoharan M. Careful selection and close monitoring of low-risk prostate cancer patients on active surveillance minimizes the need for treatment. *Eur Urol*. 2010; 58:831–835. [PubMed: 20800964]
20. Gershman B, Dahl DM, Olumi AF, Young RH, McDougal WS, Wu CL. Smaller prostate gland size and older age predict Gleason score upgrading. *Urol Oncol*. 2011; doi: 10.1016/j.urolonc.2011.11.032
21. Su LJ, Arab L, Steck SE, et al. Obesity and prostate cancer aggressiveness among African and Caucasian Americans in a population-based study. *Cancer Epidemiol Biomarkers Prev*. 2011; 20:844–853. [PubMed: 21467239]
22. Freedland SJ, Kane CJ, Amling CL, et al. Upgrading and downgrading of prostate needle biopsy specimens: risk factors and clinical implications. *Urology*. 2007; 69:495–499. [PubMed: 17382152]
23. Freedland SJ, Isaacs WB, Platz EA, et al. Prostate size and risk of high-grade, advanced prostate cancer and biochemical progression after radical prostatectomy: a search database study. *J Clin Oncol*. 2005; 23:7546–7554. [PubMed: 16234520]
24. Liu JJ, Brooks JD, Ferrari M, Nolley R, Presti JC. Small prostate size and high grade disease--biology or artifact? *J Urol*. 2011; 185:2108–2111. [PubMed: 21496855]
25. Loeb S, Han M, Roehl KA, Antenor JA, Catalona WJ. Accuracy of prostate weight estimation by digital rectal examination versus trans-rectal ultrasonography. *J Urol*. 2005; 173:63–65. [PubMed: 15592029]
26. Capitanio U, Karakiewicz PI, Jeldres C, et al. The probability of Gleason score upgrading between biopsy and radical prostatectomy can be accurately predicted. *Int J Urol*. 2009; 16:526–529. [PubMed: 19389085]
27. Imamoto T, Suzuki H, Utsumi T, et al. External validation of a nomogram predicting the probability of prostate cancer Gleason sum upgrading between biopsy and radical prostatectomy pathology among Japanese patients. *Urology*. 2010; 76:404–410. [PubMed: 19716590]
28. Capitanio U, Karakiewicz PI, Valiquette L, et al. Biopsy core number represents one of foremost predictors of clinically significant gleason sum upgrading in patients with low-risk prostate cancer. *Urology*. 2009; 73:1087–1091. [PubMed: 19195695]
29. Kulkarni GS, Lockwood G, Evans A, et al. Clinical predictors of Gleason score upgrading: implications for patients considering watchful waiting, active surveillance, or brachytherapy. *Cancer*. 2007; 109:2432–2438. [PubMed: 17497649]
30. Moussa AS, Kattan MW, Berglund R, Yu C, Fareed K, Jones JS. A nomogram for predicting upgrading in patients with low- and intermediate-grade prostate cancer in the era of extended prostate sampling. *BJU Int*. 2010; 105:352–358. [PubMed: 19681898]

Figure 1.

Receiver operating characteristic curves are shown for the Biopsy-Integrated Algorithm for Determining Gleason 6 Upgrading Risk (BADGR). BADGR discriminated with areas under the curve (AUCs) of 0.753, 0.677, and 0.672 at University of Wisconsin (UW), University of Chicago (UC), and University of Miami (UM), respectively (all $P < 0.0001$).

Figure 2. Calibration and decision curves for the Biopsy-Integrated Algorithm for Determining Gleason 6 Upgrading Risk (BADGR) are shown. (A) Calibration demonstrated virtually no departure from ideal predictions in all cohorts across all probabilities of upgrading from

Gleason 6 prostate cancer. (B) Decision analysis demonstrated high net benefit across a wide range of threshold probabilities. PSAD indicates prostate-specific antigen density.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Graphical representation of the Biopsy-Integrated Algorithm for Determining Gleason 6 Upgrading Risk (BADGR). BADGR is validated and calibrated for use in predicting upgrading in patients with Gleason 6 PC with least 10 cores on transrectal ultrasound-guided biopsy and no history of neoadjuvant therapy. Each individual variable is assigned points using the top scale. The total points score is then converted to the predicted probability of Gleason upgrading using the bottom scale. Abbreviations: BMI, body mass index; PSAD, prostate-specific antigen density.

TABLE 1

Gleason 6 Cohort Characteristics

Variable	University of Wisconsin	University of Chicago	University of Miami
Years	2004–2012	2003–2012	2000–2012
No. of patients	413	1156	392
Age, mean (SD)	59.0 (6.5)	58.6 (6.7)	59.7 (3.6)
BMI, mean (SD)	28.8 (4.2)	28.3 (4.6)	26.8 (3.6)
PSA, mean (SD)	6.6 (5.6)	5.6 (4.0)	6.1 (4.1)
No. of cores obtained, mean (SD)	11.8 (2.1)	12.3 (2.5)	11.4 (1.7)
Upgrading (%)	52.5	40.6	33.3
Clinical stage T1c (%)	84	83	73

Abbreviations: BMI, body mass index; PSA, prostate-specific antigen; SD, standard deviation.

TABLE 2
Association of Gleason 6 Upgrading With Pathological Variables After Radical Prostatectomy at University of Wisconsin

Variable	No Upgrade	Upgrade	OR	95% CI	P
Prostate weight (g)					
Median (range)	45 (18–167)	39.4 (20–157)	0.983	0.973–0.993	.001
Mean (SD)	48.7 (21)	43.2 (17)			
Prostate dimensions (cm):					
Largest					
Median (range)	5 (1–8)	4.8 (3–8)	0.862	0.695–1.068	.174
Mean (SD)	4.9 (0.8)	4.9 (0.8)			
2nd largest					
Median (range)	4.2 (2–7)	4.2 (3–7)	0.871	0.676–1.121	.282
Mean (SD)	4.3 (0.7)	4.2 (0.6)			
3rd largest					
Median (range)	3.5 (1–6)	3.6 (2–6)	1.020	0.914–1.138	.728
Mean (SD)	3.5 (0.7)	3.6 (2.2)			
Slides positive (%)					
Median (range)	29 (0–100)	43 (0–100)	10.674	4.699–24.247	<.0001
Mean (SD)	32 (21)	43 (21)			
Tumor volume (%)					
Median (range)	5 (1–50)	10 (1–70)	1.054	1.034–1.075	<.0001
Mean (SD)	7.2 (8)	12.7 (12)			
Pathological stage					
Stage 2	282 (97%)	225 (86%)	5.013	2.365–10.625	<.0001
Stage 3	9 (3%)	36 (14%)			
Lymph node dissection performed					
No	221 (73%)	209 (80%)	0.656	0.441–0.977	.038
Yes	80 (27%)	52 (20%)			
Extraprostatic extension					
No	283 (97%)	227 (87%)	4.414	2.095–9.301	<.0001
Yes	7 (3%)	33 (13%)			

Variable	No Upgrade	Upgrade	OR	95% CI	P
Seminal vesicle involvement					
No	289 (99%)	255 (98%)	3.400	0.680–68.995	.136
Yes	2 (1%)	6 (2%)			
Margins					
Any					
No	258 (89%)	207 (79%)	2.040	1.275–3.264	.003
Yes	33 (11%)	54 (21%)			
Apex					
No	260 (89%)	224 (86%)	1.385	0.832–2.306	.210
Yes	31 (11%)	37 (14%)			
Bladder neck					
No	284 (98%)	246 (94%)	2.474	0.993–6.166	.052
Yes	7 (2%)	15 (6%)			
Peripheral					
No	260 (89%)	216 (83%)	1.747	1.069–2.857	.026
Yes	31 (11%)	45 (17%)			
Laterality (pathological)					
Unilateral	53 (18.2%)	25 (9.6%)	2.111	1.270–3.510	.004
Bilateral	237 (81%)	236 (90%)			

Abbreviations: CI, confidence interval; OR, odds ratio; SD, standard deviation.

TABLE 3
Univariate Analysis of Clinical and Pathological Variables at Diagnosis Associated With Gleason 6 Upgrading

Variable	No Upgrade	Upgrade	OR	95% CI	P
Clinical Variables					
Age	58.6 (6.5)	59.3 (6.5)	1.017	0.987–1.048	.259
PSA	6.18 (6.3)	7.06 (5.1)	1.033	0.990–1.077	.137
Obese (BMI ≥ 30)					
No	139 (51.3%)	132 (48.7%)	1.677	1.107–2.541	.015
Yes	54 (38.6%)	86 (61.4%)			
AUA symptom score	9.5 (6.7)	9.0 (5.6)	0.986	0.945–1.028	.499
White race					
No	4 (30.8%)	9 (69.2%)	0.491	0.149–1.621	.243
Yes	190 (47.5%)	210 (52.5%)			
Family history					
1st degree					
No	115 (44.7%)	142 (55.3%)	0.772	0.504–1.183	.235
Yes	65 (51.2%)	62 (48.8%)			
Any degree					
No	108 (44.3%)	136 (55.7%)	0.761	0.502–1.153	.198
Yes	72 (51.1%)	69 (48.9%)			
Smoking history					
No	107 (49.1%)	111 (50.9%)	1.391	0.928–2.084	.110
Yes	70 (40.9%)	101 (59.1%)			
ASA class					
I	0 (0%)	1 (100%)	1.067	0.618–1.841	.816
II	13 (68.4%)	6 (31.6%)			
III	135 (46.1%)	158 (53.9%)			
IV	16 (55.2%)	14 (44.8%)			
Clinical stage					
T1c	169 (48.4%)	180 (51.6%)	1.106	0.711–1.721	.654
T2a	20 (51.3%)	19 (48.7%)			

Variable	No Upgrade	Upgrade	OR	95% CI	P
T2b	1 (33.3%)	2 (66.7%)			
T2c	1 (25%)	3 (75%)			
DRE-estimated size (cc)	35.2 (14)	33.4 (12)	0.988	0.966–1.011	.306
TRUS-estimated size (cc)	41.3 (18)	35.7 (16)	0.980	0.966–0.993	.004
PSA density	0.164 (0.16)	0.228 (0.20)	15.370	2.228–106	.006
Pathological Variables					
Number of positive cores	2.4 (1.6)	3.3 (2.5)	1.289	1.099–1.512	.002
Number of negative cores	9.3 (2.4)	8.6 (3.2)	0.911	0.845–0.981	.014
% of cores positive	20.9 (14.1)	28.2 (21.0)	1.016	1.006–1.026	.002
Maximum core involvement (%)	18.5 (20)	31.6 (26)	1.024	1.013–1.036	<.0001
Estimated tumor volume (%)	4.58 (5.9)	9.27 (11.7)	1.078	1.042–1.117	<.0001
Laterality					
Unilateral	133 (54.1%)	113 (45.9%)	2.044	1.355–3.084	.001
Bilateral	57 (36.5%)	99 (63.5%)			

Abbreviations: ASA, American Society of Anesthesiologists; AUA, American Urological Association; CI, confidence interval; DRE, digital rectal examination; OR, odds ratio; PSA, prostate-specific antigen; TRUS, transrectal ultrasound.

TABLE 4
Multivariate Model for Predicting Gleason Score Upgrading From Gleason 6 (BADGR)

Variable	B coefficient	SE	Wald	OR	CI	P
PSAD	5.434	1.810	9.013	1.72 ^a	1.214–2.452	.003
Obesity	0.644	0.325	3.933	1.90	1.008–3.599	.047
No. of positive cores	0.208	0.083	6.224	1.23	1.046–1.450	.013
Maximum core involvement (%)	0.020	0.008	6.873	1.02	1.005–1.035	.009
Constant	–1.877					

^aPer 0.1 unit increase in PSAD.

Abbreviations: CI, confidence interval; OR, odds ratio; PSAD, prostate-specific antigen density; SE, standard error.

TABLE 5

Comparison of Predictive Accuracies of BADGR Versus Existing Active Surveillance Criteria for Predicting Upgrading in Gleason 6 Biopsy Patients Who Underwent Prostatectomy at UW, UC, and UM

Eligibility Criteria Used	Prostate-Specific Antigen	No. of Positive Cores	Maximum % Core Involvement	Clinical Stage	Obesity	UW (n = 413)			UC (n = 1151)			UM (n = 392)		
						AUC	95% CI		AUC	95% CI		AUC	95% CI	
Carter (JHU, Epstein criteria) ¹⁶	PSAD 0.15	2	50	T1c	-	0.598	(0.553, 0.643)		0.584	(0.556, 0.610)		0.581	(0.543, 0.619)	
Schroder (PRIAS) ¹⁷	PSA 10 and PSAD 0.2	2	-	T1c-T2	-	0.577	(0.529, 0.624)		0.584	(0.555, 0.613)		0.622	(0.575, 0.669)	
Carroll (UCSF) ¹⁸	PSA 10	33% (of at least 6)	50	T1c-T2	-	0.540	(0.495, 0.586)		0.550	(0.521, 0.579)		0.505	(0.455, 0.556)	
Soloway (UM) ¹⁹	PSA 15	2	20	T1c-T2	-	0.585	(0.537, 0.633)		0.578	(0.549, 0.607)		0.582	(0.532, 0.632)	
BADGR	Continuous	Continuous	Continuous	-	Binary	0.753	(0.690, 0.819)		0.677	(0.646, 0.711)		0.672	(0.619, 0.723)	
<i>P</i>						<.0001 ^a , <.0001 ^b , <.0001 ^c , <.0001 ^d , <.0001 ^a , <.0001 ^b , <.0001 ^c , <.0001 ^d , <.0001 ^a , <.0001 ^b , <.0001 ^c , <.0001 ^d								

Abbreviations: AUC, area under the curve; BADGR, Biopsy-Integrated Algorithm for Determining Gleason 6 Upgrading Risk; CI, confidence interval; JHU, Johns Hopkins University; PRIAS, Prostate Cancer Research International: Active Surveillance; UCSF, University of California, San Francisco; UW, University of Wisconsin; UC, University of Chicago; UM, University of Miami.

^aBADGR versus Epstein.

^bBADGR versus PRIAS.

^cBADGR versus UCSF.

^dBADGR versus UM.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 6

BADGR Predicts Upgrading in Radical Prostatectomy Patients Who Qualify for Active Surveillance Using Previous Protocols

Eligibility Criteria Used	<i>n</i>	Not Upgraded	Upgraded	AUC	<i>P</i>
Carter (JHU, Epstein criteria) ¹⁶	569	413 (72.58%)	156 (27.4%)	0.597	0.0009
Schroder (PRIAS) ¹⁷	836	578 (69.14%)	258 (30.9%)	0.620	<.0001
Carroll (UCSF) ¹⁸	1199	728 (60.72%)	471 (39.3%)	0.646	<.0001
Soloway (UM) ¹⁹	991	647 (65.29%)	344 (34.7%)	0.672	<.0001