Pathology & Oncology Research (2020) 26:1367–1375
https://doi.org/10.1007/s12253-019-00632-1

REVIEW

# Prognostic Differences in ISUP Grade Group 4: a Systematic Review and Meta-Analysis

Thomas Chengxuan Lu[1,2,3] · Luke Collins[4] · Penelope Cohen[5] · Alex Jay[6] · Jared M. Campbell[7] · Michael O'Callaghan[6,8,9]

Received: 8 February 2019 / Accepted: 4 March 2019 / Published online: 14 March 2019


## Abstract

The ISUP (Internal Society of Urologic Pathology) recently adopted a five-tiered prognostication system. There is evidence to suggest that the ISUP grade group 4 is a heterogeneous entity regarding prognosis. Our aim was to systematically examine the existing evidence to determine if outcome differences exist within the ISUP grade group 4. A systematic search of the literature for all studies examining the heterogeneity of the ISUP grade group 4 was conducted. Available studies were combined with meta-analysis to evaluate prognostic differences within the ISUP grade group 4 measured by all-cause mortality (ACM) and the prostate cancer-specific mortality (PCSM). Eight studies were identified and utilised a variety of outcome measures to answer the question of heterogeneity within the ISUP grade group 4. Four of these studies examined prognosis using both ACM and PCSM. These were combined into a meta-analysis. The combined group of 5 + 3/3 + 5 had statistically significant higher ACM (hazard ratio [HR] 1.23, 95% confidence internal [Cl] 1.08–1.41) when compared to the 4 + 4 group. There was no difference in the PCSM between the two groups (HR 1.34, 95% CI 0.89–2.01). However, heterogeneity was high for this analysis secondary to a range of methodological differences. Our meta-analysis showed that Gleason grade 3 + 5/5 + 3 had higher ACM than Gleason grade group 4 + 4. Measures of PCSM were statistically insignificant, although heterogeneity was high. Evidence suggests that heterogeneity is likely, although inconclusive. Further studies with consistent methodologies are required to answer this question.

**Keywords** Prostate cancer · Gleason score · Grade group 4 · Biopsy · Prostate cancer specific mortality

✉ Thomas Chengxuan Lu
thomascxlu@gmail.com

1 The George Institute for Global Health, UNSW, Sydney, NSW 2052, Australia

2 St George Hospital, Kogarah, NSW 2207, Australia

3 Westmead Hospital, Westmead, NSW 2145, Australia

4 Royal Adelaide Hospital, Adelaide 5000, Australia

5 SA Pathology, Royal Adelaide Hospital, Adelaide, Australia

6 Urology Unit, Flinders Medical Centre, Bedford Park, Australia

7 Centre for Nanoscale Biophotonics, Graduate School of Biomedical Engineering, UNSW, Sydney, NSW 2052, Australia

8 South Australia Prostate Cancer Clinical Outcomes Collaboration, Repatriation Hospital, Heidelberg Heights, Australia

9 Flinders Centre for Innovation in Cancer, Flinders University, Adelaide, Australia

## Introduction

The Gleason score was first described in 1966 and represents the first systematic approach to the grading of prostate cancer. It is based on the microscopic architectural features of prostatic adenocarcinoma, as observed by an anatomical pathologist. Five glandular architectural patterns were described and each assigned a numeric 1–5. The overall Gleason score was calculated by combining the score of the most predominant pattern and the second most predominant pattern. By combining both the primary and secondary patterns the prognostic power of the system was increased and it allowed clinicians to better risk stratify patients and guide decision-making [1, 2].

There have been several modifications to Gleason's original grading system with the most significant occurring in 2005 [3]. These changes included cribriform architecture being placed into grade 4 and the reporting of needle biopsies using

the primary pattern and the highest-grade present (as opposed to the primary and secondary pattern). More recently the International Society of Urologic Pathology (ISUP) has recommended a new prognostic grading system that has simplified the 9 possible Gleason scores (3 + 3, 3 + 4, 3 + 5, 4 + 3, 4 + 4, 4 + 5, 5 + 3, 5 + 4, 5 + 5) into 5 prognostic grade groups (ISUP 1: 3 + 3, ISUP 2: 3 + 4 ISUP 3: 4 + 3 ISUP 4 3 + 5, 4 + 4 and 5 + 3 and ISUP 5: 4 + 5, 5 + 4 and 5 + 5 [4]. The proposed reasons for these changes have been to simplify risk stratification of patients, recognition of the low incidence of patterns 1 and 2 (meaning a scoring system that begins at Gleason 6) and to improve patient acceptance of active surveillance management strategies for Gleason 6 (ISUP 1). The new system has been published by WHO in Pathology and genetics of tumour of the urinary system and male genital organs and adopted by uropathologists widely [5].

Reflected in the new ISUP grade groups is the differentiation of Gleason score 7, separating 3 + 4 into group 2, and 4 + 3 into group 3. It is clear from the literature that Gleason score 4 + 3 has a less favourable outcome, and appears to be more aggressive with higher rates of with higher rates of biochemical failure, systemic recurrence, and cancer-specific death rates [6–11] when compared to 3 + 4. This has been considered in both the American Urological Association (AUA) and the National Comprehensive Cancer Network (NCCN) guidelines [1, 2, 4].

The ISUP grade group 4 is equivalent to prostate cancer with a total Gleason score of 8, comprising of 4 + 4, 3 + 5 and 5 + 3. It is still considered a homogenous entity by both the AUA and the NCCN in light of risk and treatment. However, recent literature has raised questions regarding whether Gleason score 8, or ISUP grade group 4 is a heterogeneous entity in terms of prognosis, and hence whether there is merit in the reclassification of grade group 4 into separate grade groups [12, 13].

Consequently, the purpose of this review was to systematically assess the prognostic differences within ISUP group 4 in terms of mortality and biochemical or clinical progression. This information may thus guide current opinion on prognostic heterogeneity within ISUP grade group 4.

## Methods

### Search Strategy and Selection Criteria

A three-step search strategy was utilised in this review. An initial limited search of EMBASE and PubMed was undertaken using key words, to identify any further keywords and index terms. A second search using all identified key words and index terms was then performed in EMBASE and PubMed. Finally, the reference lists of all identified literature were searched for additional suitable studies. The full search protocol is published in PROSPERO [14]. Studies were extracted into Covidence systematic review software (Veritas Health Innovation). Duplicate articles were then excluded, abstracts were screened for relevance and the full texts of those identified as potentially relevant were retrieved. Manual review of reference lists was also performed to identify additional studies.

Eligible studies for the systematic review examined prognostic differences within the ISUP group 4 and were limited to human subjects and publications in the English language. Unpublished data, conference presentations, editorials, case reports and correspondence were excluded. The primary endpoints sought were all-cause mortality, prostate cancer-specific mortality and biochemical recurrence.

### Data Extraction and Critical Appraisal

Articles, texts, table and figures were independently reviewed by two investigators (C.L, L.C). Quality of the studies was examined using the Newcastle-Ottawa scale. This instrument grades risk of bias and quality using a scoring system for methods of selection, comparability and outcome [15]. Discrepancies were resolved by discussion until a consensus was reached. Results were reviewed by all authors.

### Statistical Analysis

Statistical analysis of the prognostic difference between group 4 + 4 and a combined group of 3 + 5/5 + 3 was performed for the available data. Two separate measures of prognostication were utilised; all-cause mortality and prostate cancer-specific survival were measured between these groups.

The hazard ratio (HR) was used as a summary statistic and reported with a 95% confidence interval (CI). The Chi-square and the $I^2$ statistic were used to estimate the percentage of total variation across studies due to heterogeneity. $I^2$ values exceeding 50% is indicative of considerable heterogeneity. A $p$ value <0.1 was considered statically significant, indicative of heterogeneity. Statistical analysis was conducted with R 3.5.0. Random-effect model was utilised for the meta-analysis to account for clinical and methodological variation between studies.

## Results

The initial electronic database search yielded 512 references (Fig. 1). Following the exclusion of duplicates and irrelevant articles through title and abstract screening, 24 potentially relevant publications were retrieved for critical evaluation. Manual review of the reference list did not yield any additional references. After application of the selection criteria, eight

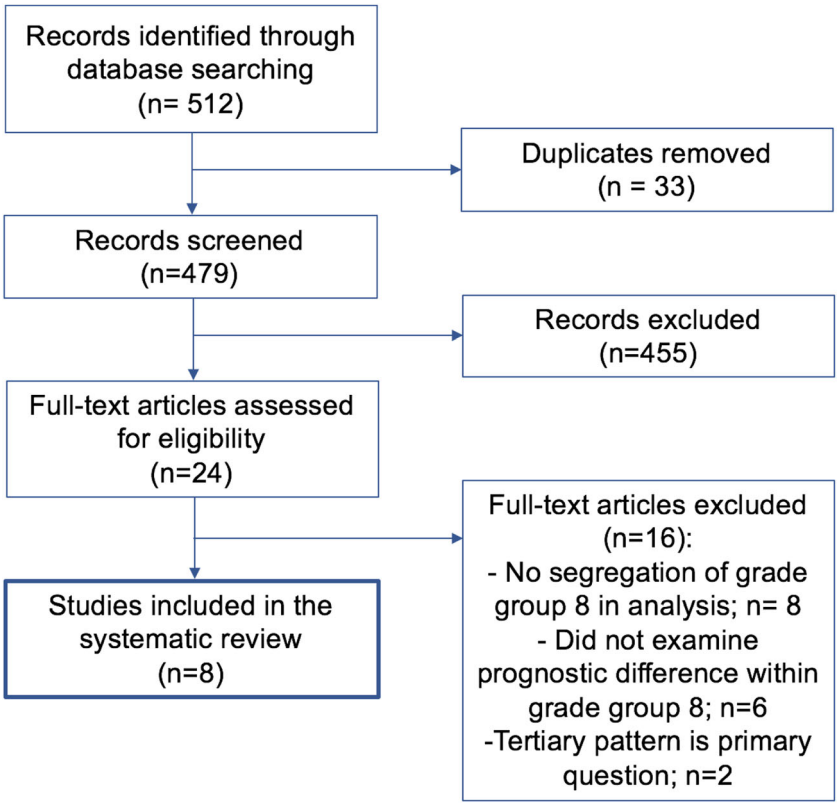


**Fig. 1** Systematic review search strategy

studies were included in our systematic review, and data from four studies were included in our meta-analysis. There were no major discrepancies between reviewers during trial inclusion and data extraction. Minor discrepancies were resolved by consensus agreement.

## Study Characteristics

All eight studies were retrospective observational studies (Table 1). The number of participants enrolled in the studies ranged from 423 to 40,533. Significant variation was noted regarding the inclusion criteria and patient population, as well as the outcomes measured.

Four of the studies examined all-cause mortality (ACM) [12, 16–18] and another examined prostate cancer-specific mortality (PCSM) only [13]. A sixth article examined multiple outcomes including prostate cancer-specific survival and cancer status [19]. There were a further two studies of biochemical recurrence (BCR) [20, 21].

Biopsy was the predominant form of pathological specimen analysed and was included in 7 of the 8 studies. Two of the eight studies included both biopsy and radical prostatectomy (RP) [13, 21]. One study utilised RP alone [20]. Four of the reviewed studies combined Gleason 5 + 3/3 + 5 into a single group for evaluation, and three segregated the 5 + 3 and 3 + 5 groups. One study completely excluded 5 + 3 due to small sample size [22]. All studies were conducted in countries which would be considered as high income by the World Bank classification.

Only four studies were included in the meta-analysis due to methodological and outcome measure differences [17, 19–21]. Studies included all examined ACM and PCSM between the Gleason scorings 4 + 4 and 5 + 3/3 + 5 in biopsy specimens. One study which measured PCSM could not be included, as the comparator groups 5 + 3 and 3 + 5 were measured separately. However, this study had the largest number of participants [18]. A final study [24] which examined multiple aspects of prostate cancer survival between the 3 + 5 and the 4 + 4 groups, lacked adequate reporting of statistical data and hazard ratios which precluded its use in the meta-analysis. The two studies of BCR were not combined in meta-analysis due to methodological differences and incomplete reporting of outcome data.

## Assessment of Quality of Evidence

Study quality was assessed using the Newcastle-Ottawa scale (Table 2). All studies were graded as moderate to high quality, where scores of 0–3, 4–6 and 7–9 were defined as low, moderate and high quality respectively. Common study limitations included variable length of follow-up and failure to clearly describe loss to follow-up.

**Table 1** Study characteristics

| | Lu 2017 [16] | Huynh 2015 [12] | Rusthoven 2014 [17] | Rusthoven 2015 [18] | Van den Bergh 2016 [21] | Gandaglia 2017 [20] | Mahal 2015 [13] | Harding-Jackson 2016 [19] |
|---|---|---|---|---|---|---|---|---|
| Location, dates | Australia 1998–2015 | USA, 1998–2012 | USA, 2006–2008 | USA, 2004–2006 | Australia, Netherlands 2003–2015 | Italy, USA, France 1990–2014 | USA, 2004–2011 | USA 2005–2013 |
| Database | SA-PCCOC | Chicago Prostate Cancer Center | SEER | SEER | Two centres: Epworth, Melbourne & NKI, Amsterdam | Three tertiary referral institutions: IRCCS Ospedale San Raffaele, Milan; Mayo Clinic, Rochester, NY; L'Institut Mutualiste Montsouris, Paris | SEER | Medical College of Wisconsin/Froedtert Hospital and University of Miami Miller School of Medicine |
| Included participants & criteria | 4080 men with non-metastatic PC with biopsy GS 7–9 | 462 men with T1c-T3 PC treated with brachytherapy ± EBRT or salvage ADT and GS 8 at biopsy | 4654 men with metastatic PC M1 with biopsy GS 6–10 | 26,885 men with non-metastatic PC and EBRT alone (combined modality excluded), GS 6–10 | Men post-radical prostatectomy with GS 6–10; 3416 with biopsy specimens and 3479 with RP specimens | 1089 men who underwent radical prostatectomy with pelvic LN dissection and GS 8 at final pathology. No neoadjuvant therapy | 40,533 men with N0 M0 Gleason score 8 or 9 at biopsy or surgery | 423 men with highest biopsy GS of 8 |
| Number excluded or lost | Unclear, excluded if missing data on age or date of diagnosis, Gleason patterns | Nil | 20 GS ≤5 | 1791, missing PSA, primary-secondary patterns, cause of death | Nil | 207 with missing follow-up or BCR information | 6626, missing T stage or PSA data | 5, GS 5 + 3 = 8 excluded due to small size |
| Study design | | | | | | | | |
| Study type | Retrospective | Retrospective; prospectively assembled database | Retrospective | Retrospective | Retrospective | Retrospective | Retrospective | Retrospective |
| Aim | Differences in ISUP 4 | Differences in ISUP 4 | Prognosis of GS in metastatic PC | Evaluate prognosis of GS in ERBT | Validation of ISUP tiering in biopsy and RP | Validation of ISUP Grade 4 grouping | Differences in Gleason score 8 | GS 4 + 4 vs 3 + 5 |
| Gleason grouping | GS 3 + 4, 4 + 3, 4 + 4, 3 + 5/5 + 3, 9 | GS 4 + 4, 3 + 5/5 + 3 | GS 6, 7, 8 (4 + 4, GP 5 present (i.e. 3 + 5/5 + 3)), 9, 10 | GS 6,7,8 (4 + 4, GP 5 present), 9, 10 | GS 6, 7, 4 + 4, 5 + 3, 4 + 5, 5 + 4, 5 + 5 | GS 4 + 4, 3 + 5, 5 + 3 | GS 4 + 4, 3 + 5, 5 + 3, GS 9 | GS 4 + 4, 3 + 5 |
| Number of GS8 | *N* = 740; 4 + 4 (*n* = 664), 3 + 5 (*n* = 55) 5 + 3 (n = 21) | *N* = 462, 4 + 4 (*n* = 421), 3 + 5/5 + 3 (*n* = 41) | *N* = 1073, 3 + 5/5 + 3 (*n* = 167), 4 + 4 (*n* = 906) | *N* = 2904, 3 + 5/5 + 3 (*n* = 359), 4 + 4 (*n* = 2545) | *N* = 529, Biopsy [*n* = 318; 3 + 5 (*n* = 50), 4 + 4 (*n* = 252), 5 + 3 (*n* = 16)], RP [*n* = 211; 3 + 5 (*n* = 62), 4 + 4 (*n* = 134), 5 + 3 (*n* = 15) | *N* = 1089; 3 + 5 (*n* = 295), 4 + 4 (*n* = 651), 5 + 3 (*n* = 143) | *N* = 40,533; 4 + 4 (n = 21,503), 3 + 5 (*n* = 2668), 5 + 3 (*n* = 892) | *N* = 179; 4 + 4 (*n* = 121), 3 + 5 (*n* = 58) |
| Outcome measures | All-cause mortality, prostate cancer specific mortality | All-cause mortality (ACM), PCSM | Overall survival, Prostate cancer specific survival at 4 yrs | Overall survival, Prostate cancer specific survival | Biochemical recurrence (PSA ≥0 .2ng/ml) | Biochemical recurrence (2 consecutive PSA ≥0 .2ng/ml), Clinical recurrence | Prostate cancer specific mortality (PCSM) | Prostate cancer specific survival, treatment choice, cancer status and vital status |
| Follow-up (months) | Median 60 | Median 91 | 48 | Median 72 | Median 20.4 | Median 83 | Median 36 | Median 33.4 |
| Pathology | | | | | | | | |
| Specimen | Biopsy | Biopsy | Biopsy | Biopsy | Biopsy "first and worst" average of all cores; RP highest grade individual nodule | Radical Prostatectomy | Highest GS of biopsy or RP, 26914 biopsies (66.4%) and 13,619 RP (33.6%) | Biopsy |

**Table 1** (continued)

| | Lu 2017 [16] | Huynh 2015 [12] | Rusthoven 2014 [17] | Rusthoven 2015 [18] | Van den Bergh 2016 [21] | Gandaglia 2017 [20] | Mahal 2015 [13] | Harding-Jackson 2016 [19] |
|---|---|---|---|---|---|---|---|---|
| Number of pathologists | Various specialist uro-pathologists and non-specialist pathologists | 1 pathologist expert in genitourinary cancers | Unclear | Unclear | Unclear, local specialist uro-pathologists | Unclear, dedicated uro-pathologists | Unclear | 2 urological pathologists |
| Statistical Analysis | Cox regression and Fine and Gray regression | Cox and Fine-Gray regression; Competing risks regression | Cox regression | Cox regression | Kaplan-Meier curves with log-rank test; Cox regression for confounders | Kaplan-Meier analyses with log-rank test; Multivariable Cox regression | Multivariable Fine and Gray competing risks regression | Log-rank test |
| Adjusted variables | PSA, clinical stage, intention for cure | Age, pathologic T-stage (T1–3), log PSA, treatment (BT, ADT, EBRT), treatment propensity score | Age (<70 y, ≥70 y), T stage (T1–4), PSA (0–49.9, 50–97.9, ≥98 ng/ml), | Age (<70, ≥70 y), PSA (<10, 10–19.9, ≥20 ng/ml), T stage (T1–4), year, race, marital status, regional tumour registry | Margin status, pathological T and N stage | Age, year of surgery, pathologic stage, positive surgical margins, lymph node invasion, adjuvant treatments | Age, race, treatment received (no definitive, RT, RP), PSA level, T-stage | |

## All-Cause Mortality

The ACM was significantly higher in the combined group of 5 + 3/3 + 5 compared to the 4 + 4 group (hazard ratio [HR] of 1.24, 95% confidence interval [CI] 1.03–1.49; Fig. 2). ACM variation was not significant ($p = 0.16$) and the percentage attributable to between study heterogeneity was moderate ($I^2 = 36.20\%$).

## Prostate Cancer-Specific Mortality

Regarding the analysis for PCSM, there was no difference between the 5 + 3/3 + 5 groups compared to the 4 + 4 groups (HR 1.34, 95% CI 0.89–2.01; Fig. 3). Chi square analysis suggested a significant level of heterogeneity ($p = 0.04$) between the results of the studies in the PCSM analysis with $I^2$ analysis indicating a large percentage of the heterogeneity was not attributable to chance variation ($I^2 = 70.04$). Sensitivity analysis with removal of the largest contributor to statistical heterogeneity [12] did not produce a significant result, with persistent high heterogeneity reflected by $I^2$. The largest study which measured only PCSM, that could not be combined into meta-analysis did demonstrate heterogeneity in the ISUP grade group 4 [13], with worse prognosis for the 5 + 3 group.

## Biochemical Recurrence and Clinical Recurrence

Two papers examined the possible differences in BCR rates within the ISUP grade group 4 [20, 21]. These were not able to be combined into a meta-analysis due to incomplete reporting of data. One paper reported significant overlap between 3 + 5 and 4 + 3 in both the BCR in biopsy and RP populations [21], whilst there was significant favourable outcome in the 3 + 5 group in comparison to the 4 + 4 group in both the biopsy and RP population. The other study examined the ISUP grade group 4 separating Gleason 3 + 5 and 5 + 3 [20]. Men with 4 + 4 had a significantly increased BCR when compared to the 3 + 5 group. There was no difference in BCR found between men with 3 + 5 and 5 + 3 disease. This paper also reported that clinical recurrence was significantly lower in 3 + 5 disease.

## Other Outcome Measures

A variety of other outcome measures were utilised to differentiate between ISUP grade group 4 prostate cancers. One study examined difference between 4 + 4 and the 3 + 5 group [24]. Outcome measures included the persistent cancer rates (after treatment) and a cancer relapse rate which favoured the 4 + 4 group. The same study also reported the cancer specific survival at 36 months without reporting any hazard ratios and did not show any statistically significant differences in outcomes. The same study also reported the relationship of radical prostatectomy and fraction of positive biopsy core samples with no differences between the groups.

**Table 2** Newcastle – Ottawa scale of quality assessment scale

| Study | Representative sample of relevant population | Comparability score | Outcome | Overall score (maximum 9) |
|---|---|---|---|---|
| Mahal et al. 2015 [13] | ★★★★ | ★★ | ★★ | 8 |
| Huynh et al. 2015 [12] | ★★★★ | ★★ | ★★★ | 9 |
| Rusthoven et al. [17] | ★★★ | ★ | ★★★ | 7 |
| Rusthoven et al. [18] | ★★★ | ★ | ★★★ | 7 |
| Harding-Jackson et al. [19] | ★★★★ | – | ★★ | 6 |
| Lu et al. [16] | ★★★★ | ★★ | ★★ | 8 |
| Van den Bergh et al. [21] | ★★★ | ★ | ★★ | 6 |
| Gandaglia et al. [20] | ★★★★ | ★★ | ★★ | 8 |

## Discussion

The initial validation studies of ISUP grading for prostate cancer combined Gleason 8 into one prognostic group [5]. Subsequent literature has raised the suggestion of outcome differences within ISUP grade group 4. The evidence in this review finds a significant effect on overall mortality from ISUP grade group 4 prostate cancer conferred by the presence of Gleason grade 5. This finding did not correspond to differences in PCSM but there was significant statistical heterogeneity for the measure of PCSM making comparison challenging.

Six of the eight studies examined mortality outcomes in either ACM and PCSM (Table 1). Four of the six studies examining mortality outcomes combined Gleason scores 5 + 3 and 3 + 5 into one comparator group using Gleason grade 5 as a poor prognostic indicator to distinguish between 4 + 4 and 5 + 3/3 + 5 [22, 23]. Using a single comparator group of 5 + 3 and 3 + 5 assumes that the outcomes of Gleason 5 + 3/3 + 5 are comparable. However, evidence from the literature suggest that outcomes might be different particularly in the case of BCR. Of the remaining two studies not included in the meta-analysis, Mahal et al. reported higher mortality for Gleason 5 + 3 compared to 4 + 4 with an almost doubled PCSM similar to the outcome of Gleason grade group 9 [13]. This was the largest of the studies but due to separation of 5 + 3/3 + 5 groups, it could not be directly compared to the others. In this study, Gleason 3 + 5 mortality outcomes were not different to those of 4 + 4. The second study not included by Harding et al. [19] compared only 4 + 4 and 3 + 5 with the exclusion of 5 + 3 due to small sample size and demonstrated no significant difference between mortality outcomes.

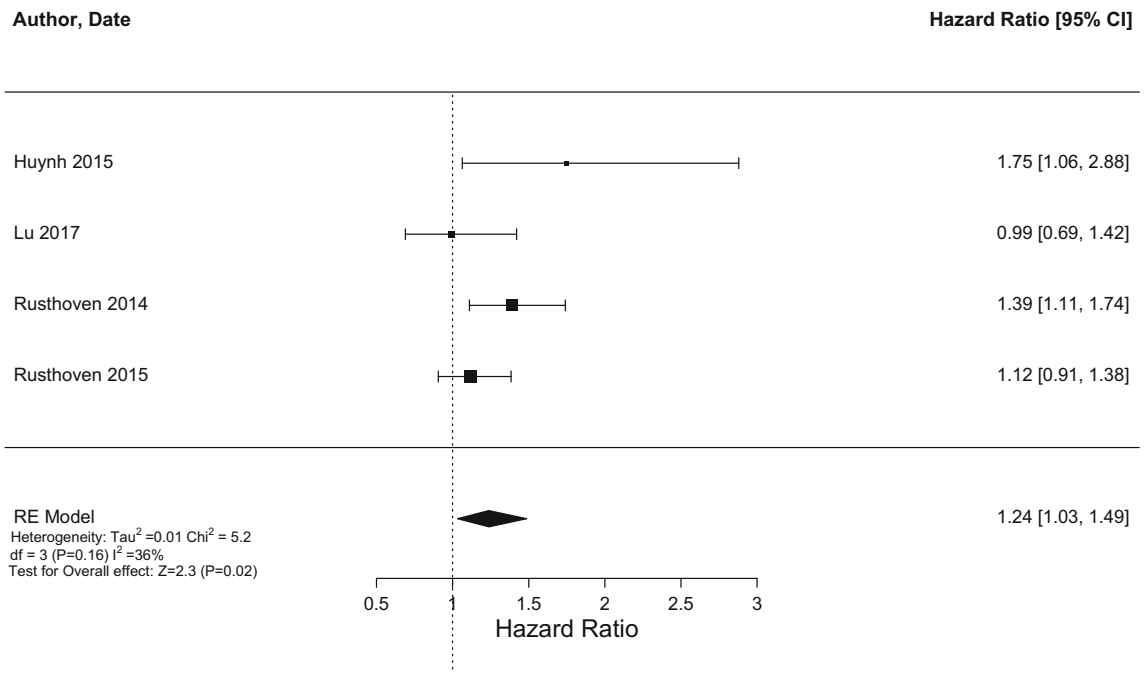


**Fig. 2** Forest plot of ACM

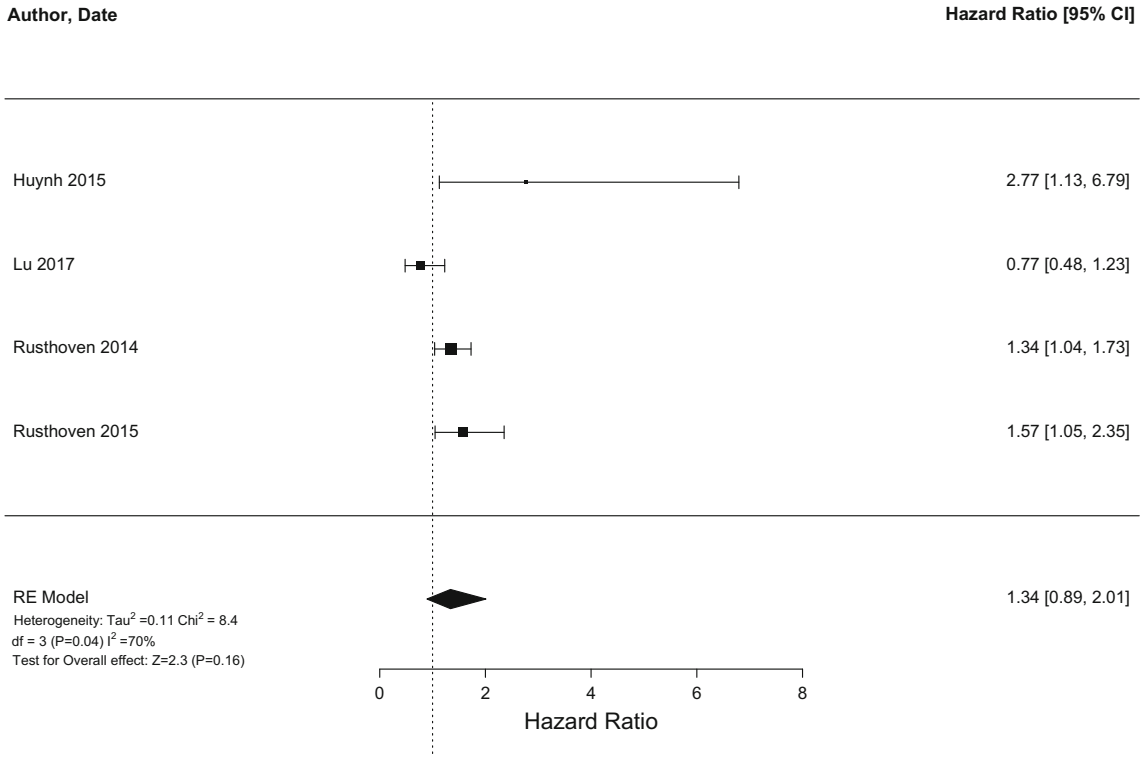


**Fig. 3** Forest plot of PCSM

Two papers examined BCR following radical prostatectomy (RP) using PSA ≥ 0 .2ng/ml as an alternative measure of prognosis [20, 21]. Van den Bergh et al. demonstrated that BCR-free survival rates were significantly more favourable for 3 + 5 than 4 + 4, comparable to the outcomes of Gleason 4 + 3, in their cohort of patients post-RP [24]. Similarly, Gandaglia et al. noted better BCR-free survival and clinical recurrence-free survival rates for men with 3 + 5 as compared to their counterparts with 4 + 4 [23]. The utility of BCR following RP as a surrogate of prognosis is somewhat limited by the heterogeneity and variable natural history of this group of men [24–26].

There are a number of methodological variances that limited this meta-analysis. The study cohorts included men with a range of disease stages and subsequent treatments. The different treatment characteristics of the included papers comprised of brachytherapy with or without further treatment, metastatic disease regardless of treatment, non-metastatic disease with ERBT alone and finally non-metastatic disease without treatment preference. Although there is clear clinical heterogeneity, there was value in combining these sub-groups as the ISUP grade groups are a prognostic indicator for all prostate cancers. Additionally, given the lack of randomisation inherent with retrospective observational studies the associations we see with mortality may be the result of confounding, rather than a true effect of ISUP group 4 heterogeneity.

The modifications to the Gleason grading system in 2005 recommended that cribriform gland pattern be placed into grade group 4 and that biopsy reporting be based on the primary and highest grades (as opposed to primary and secondary). This effectively means that secondary pattern reported can be less than 5% of tumour volume and if high grade, may lead to a significant upgrading of cases [32]. All eight studies spanned this period of transition, and only one study accounted for this [23].

Although the individual studies were generally of moderate to high quality, they each had different study designs, with some of these adjusting for patient factors including treatment method or curative intent but others not. The lack of a standardised approach to the available Gleason scoring outcome studies limits the comparability of the pool of data available. Given the results of this systematic review, it seems likely that heterogeneity does exist within ISUP Grade group 4. There is currently, however insufficient evidence to support any changes to ISUP Group 4 based on the presence of Gleason grade 5. Further high-quality, ideally prospective studies using standardised pathological methods for sample handling, Gleason scoring, and outcome reporting are needed to validate these findings and assess the true size of the risk associated with Grade 5.

## Conclusion

This meta-analysis showed that patients with 5 + 3/3 + 5 Gleason had higher ACM than those with Gleason 4 + 4. Heterogeneity was greater in studies examining PCSM and

no conclusion could be reached. Although evidence suggests that there is heterogeneity within ISUP Grade group 4, large methodological differences between current studies limits a definitive conclusion. Further studies need to be produced with consistent methodologies examining a range of outcome measures.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that there is no conflict of interests in this work.

## References


1. Sanda MG, Cadeddu JA, Kirkby E, Chen RC, Crispino T, Fontanarosa J, Freedland SJ, Greene K, Klotz LH, Makarov DV, Nelson JB, Rodrigues G, Sandler HM, Taplin ME, Treadwell JR (2017) Clinically localized prostate cancer: AUA/ASTRO/SUO guideline. Part I: risk stratification, shared decision making, and care options. J Urol 199:683–690. https://doi.org/10.1016/j.juro.2017.11.095
2. Sanda MG, Cadeddu JA, Kirkby E, Chen RC, Crispino T, Fontanarosa J, Freedland SJ, Greene K, Klotz LH, Makarov DV, Nelson JB, Rodrigues G, Sandler HM, Taplin ME, Treadwell JR (2018) Clinically localized prostate cancer: AUA/ASTRO/SUO guideline. Part II: recommended approaches and details of specific care options. J Urol 199(4):990–997. https://doi.org/10.1016/j.juro.2018.01.002
3. Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL (2005) The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol 29(9):1228–1242
4. (NCCN) NCCN (2018) Clinical practice guidelines on prostate cancer
5. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, Vickers AJ, Parwani AV, Reuter VE, Fine SW, Eastham JA, Wiklund P, Han M, Reddy CA, Ciezki JP, Nyberg T, Klein EA (2016) A contemporary prostate Cancer grading system: a validated alternative to the Gleason score. Eur Urol 69(3):428–435. https://doi.org/10.1016/j.eururo.2015.06.046
6. Burdick MJ, Reddy CA, Ulchaker J, Angermeier K, Altman A, Chehade N, Mahadevan A, Kupelian PA, Klein EA, Ciezki JP (2009) Comparison of biochemical relapse-free survival between primary Gleason score 3 and primary Gleason score 4 for biopsy Gleason score 7 prostate cancer. Int J Radiat Oncol Biol Phys 73(5):1439–1445. https://doi.org/10.1016/j.ijrobp.2008.07.033
7. Kang DE, Fitzsimons NJ, Presti JC Jr, Kane CJ, Terris MK, Aronson WJ, Amling CL, Freedland SJ, Group SDS (2007) Risk stratification of men with Gleason score 7 to 10 tumors by primary and secondary Gleason score: results from the SEARCH database. Urology 70(2):277–282. https://doi.org/10.1016/j.urology.2007.03.059
8. Koontz BF, Tsivian M, Mouraviev V, Sun L, Vujaskovic Z, Moul J, Lee WR (2012) Impact of primary Gleason grade on risk stratification for Gleason score 7 prostate cancers. Int J Radiat Oncol Biol Phys 82(1):200–203. https://doi.org/10.1016/j.ijrobp.2010.11.023
9. Stark JR, Perner S, Stampfer MJ, Sinnott JA, Finn S, Eisenstein AS, Ma J, Fiorentino M, Kurth T, Loda M, Giovannucci EL, Rubin MA, Mucci LA (2009) Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? J Clin Oncol 27(21):3459–3464. https://doi.org/10.1200/jco.2008.20.4669
10. Tollefson MK, Leibovich BC, Slezak JM, Zincke H, Blute ML (2006) Long-term prognostic significance of primary Gleason pattern in patients with Gleason score 7 prostate cancer: impact on prostate cancer specific survival. J Urol 175(2):547–551. https://doi.org/10.1016/s0022-5347(05)00152-7
11. Wright JL, Salinas CA, Lin DW, Kolb S, Koopmeiners J, Feng Z, Stanford JL (2009) Prostate cancer specific mortality and Gleason 7 disease differences in prostate cancer outcomes between cases with Gleason 4 + 3 and Gleason 3 + 4 tumors in a population based cohort. J Urol 182(6):2702–2707. https://doi.org/10.1016/j.juro.2009.08.026
12. Huynh MA, Chen MH, Wu J, Braccioforte MH, Moran BJ, D'Amico AV (2016) Gleason score 3 + 5 or 5 + 3 versus 4 + 4 prostate cancer: the risk of death. Eur Urol 69(6):976–979. https://doi.org/10.1016/j.eururo.2015.08.054
13. Mahal BA, Muralidhar V, Chen YW, Choueiri TK, Hoffman KE, Hu JC, Sweeney CJ, Yu JB, Feng FY, Trinh QD, Nguyen PL (2016) Gleason score 5 + 3 = 8 prostate cancer: much more like Gleason score 9? BJU Int 118(1):95–101. https://doi.org/10.1111/bju.13239
14. Thomas Chengxuan Lu MOC (2017) Are there prognostic differences within the International Society of Urological Pathology Grade Group 4 comprised of Gleason pattern 4+4, 3+5, 5+3? A systematic review. http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017058923. Accessed 27 May 2017
15. Stang A (2010) Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 25(9):603–605. https://doi.org/10.1007/s10654-010-9491-z
16. Lu TC, Moretti K, Beckmann K, Cohen P, O'Callaghan M (2017) ISUP group 4 - a homogenous group of prostate cancers? Pathol Oncol Res 24:921–925. https://doi.org/10.1007/s12253-017-0331-2
17. Rusthoven CG, Carlson JA, Waxweiler TV, Yeh N, Raben D, Flaig TW, Kavanagh BD (2014) The prognostic significance of Gleason scores in metastatic prostate cancer. Urol Oncol 32(5):707–713. https://doi.org/10.1016/j.urolonc.2014.01.004
18. Rusthoven CG, Waxweiler TV, DeWitt PE, Flaig TW, Raben D, Kavanagh BD (2015) Gleason stratifications prognostic for survival in men receiving definitive external beam radiation therapy for localized prostate cancer. Urol Oncol 33(2):71 e11–71 e79. https://doi.org/10.1016/j.urolonc.2014.07.010
19. Harding-Jackson N, Kryvenko ON, Whittington EE, Eastwood DC, Tjionas GA, Jorda M, Iczkowski KA (2016) Outcome of Gleason 3 + 5 = 8 prostate cancer diagnosed on needle biopsy: prognostic comparison with Gleason 4 + 4 = 8. J Urol 196(4):1076–1081. https://doi.org/10.1016/j.juro.2016.05.105
20. Gandaglia G, Karnes RJ, Sivaraman A, Moschini M, Fossati N, Zaffuto E, Dell'Oglio P, Cathelineau X, Montorsi F, Sanchez-Salas R, Briganti A (2017) Are all grade group 4 prostate cancers created equal? Implications for the applicability of the novel grade grouping. Urol Oncol 35(7):461 e467–461 e414. https://doi.org/10.1016/j.urolonc.2017.02.012
21. van den Bergh RC, van der Kwast TH, de Jong J, Zargar H, Ryan AJ, Costello AJ, Murphy DG, van der Poel HG (2016) Validation of the novel International Society of Urological Pathology 2014 five-tier Gleason grade grouping: biochemical recurrence rates for 3+5 disease may be overestimated. BJU Int 118(4):502–505. https://doi.org/10.1111/bju.13478
22. Jackson W, Hamstra DA, Johnson S, Zhou J, Foster B, Foster C, Li D, Song Y, Palapattu GS, Kunju LP, Mehra R, Feng FY (2013) Gleason pattern 5 is the strongest pathologic predictor of recurrence, metastasis, and prostate cancer-specific death in patients

receiving salvage radiation therapy following radical prostatectomy. Cancer 119(18):3287–3294. https://doi.org/10.1002/cncr.28215

23. Sabolch A, Feng FY, Daignault-Newton S, Halverson S, Blas K, Phelps L, Olson KB, Sandler HM, Hamstra DA (2011) Gleason pattern 5 is the greatest risk factor for clinical failure and death from prostate cancer after dose-escalated radiation therapy and hormonal ablation. Int J Radiat Oncol Biol Phys 81(4):e351–e360. https://doi.org/10.1016/j.ijrobp.2011.01.063
24. Antonarakis ES, Feng Z, Trock BJ, Humphreys EB, Carducci MA, Partin AW, Walsh PC, Eisenberger MA (2012) The natural history of metastatic progression in men with prostate-specific antigen recurrence after radical prostatectomy: long-term follow-up. BJU Int 109(1):32–39. https://doi.org/10.1111/j.1464-410X.2011.10422.x
25. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, Partin AW (2005) Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. Jama 294(4):433–439. https://doi.org/10.1001/jama.294.4.433
26. Punnen S, Cooperberg MR, D'Amico AV, Karakiewicz PI, Moul JW, Scher HI, Schlomm T, Freedland SJ (2013) Management of biochemical recurrence after primary treatment of prostate cancer: a systematic review of the literature. Eur Urol 64(6):905–915. https://doi.org/10.1016/j.eururo.2013.05.025