

# A systematic review of outcome reporting, definition and measurement heterogeneity in non-muscle invasive bladder cancer effectiveness trials of adjuvant, prophylactic treatment after transurethral resection

Veskimae, Erik; Subbarayan, Selvarini; Campi, Riccardo; Carron, Domitille; Omar, Muhammad Imran; Yuan, Cathy; Dimitropoulos, Konstantinos; Van Hemelrijck, Mieke; Bryan, Rik; N'Dow, James; Babjuk, Marek; Witjes, J Alfred; Sylvester, Richard; MacLennan, Steven

DOI:  
[10.3233/BLC-201510](https://doi.org/10.3233/BLC-201510)

License:  
None: All rights reserved

*Document Version*  
Peer reviewed version

*Citation for published version (Harvard):*  
Veskimae, E, Subbarayan, S, Campi, R, Carron, D, Omar, MI, Yuan, C, Dimitropoulos, K, Van Hemelrijck, M, Bryan, R, N'Dow, J, Babjuk, M, Witjes, JA, Sylvester, R & MacLennan, S 2021, 'A systematic review of outcome reporting, definition and measurement heterogeneity in non-muscle invasive bladder cancer effectiveness trials of adjuvant, prophylactic treatment after transurethral resection', *Bladder Cancer*, vol. 7, no. 2, pp. 221-241.  
<https://doi.org/10.3233/BLC-201510>

[Link to publication on Research at Birmingham portal](#)

**Publisher Rights Statement:**  
The final publication is available at IOS Press through <http://dx.doi.org/10.3233/BLC-201510>

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

Download date: 17. Apr. 2024

## **A systematic review of outcome reporting, definition and measurement heterogeneity in Non-Muscle Invasive Bladder Cancer effectiveness trials of adjuvant, prophylactic treatment after transurethral resection**

Erik Veskimae 1, Selvarani Subbarayan 2, Riccardo Campi 3, Domitille Carron 4, Muhammad Imran Omar 2, Cathy Yuan 5, Konstantinos Dimitropoulos 2,6, Mieke Van Hemelrijck 7, Richard T. Bryan 8, James N'Dow 2,6, Marek Babjuk 9, J. Alfred Witjes 10 Richard Sylvester 11+ Steven MacLennan 2+.

1 Department of Urology, Tampere University Hospital, Tampere, Finland

2 Academic Urology Unit, University of Aberdeen, Aberdeen, UK

3 Department of Department of Urologic Robotic Surgery and Renal Transplantation, University of Florence, Careggi Hospital, Florence, Italy; Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy

4 Inspiring-health GmbH, Munich, Germany

5 Division of Gastroenterology, Department of Medicine, McMaster University, Hamilton, ON, Canada

6 Department of Urology, Aberdeen Royal Infirmary, NHS Grampian, Aberdeen, UK

7 King's College London, Faculty of Life Sciences and Medicine, Translational and Oncology Research (TOUR), London, UK

8 Institute of Cancer & Genomic Sciences, University of Birmingham, Birmingham, UK.

9 Department of Urology, 2nd Faculty of Medicine, Hospital Motol, Charles University, Prague, Czech Republic

10 Department of Urology, Radboud University Medical Center, Nijmegen.

11 EAU NMIBC Guidelines Panel, Brussels, Belgium

+Joint senior authors

### **Abstract**

**Background:** Heterogenous outcome reporting in non-muscle-invasive bladder cancer (NMIBC) effectiveness trials of adjuvant intervention after transurethral resection (TURBT) has been noted in systematic reviews (SRs). This hinders comparing results across trials, combining them in meta-analyses, and evidence-based decision-making for patients and clinicians.

**Objective:** We aimed to systematically review the extent of reporting and definition heterogeneity.

**Methods:** We included randomized controlled trials (RCTs) identified from SRs comparing adjuvant treatments after TURBT or TURBT alone in patients with NMIBC (with or without carcinoma in situ) published between 2000-2020. Abstracts and full texts were screened independently by two reviewers. Data were extracted by one reviewer and checked by another.

**Results:** We screened 807 abstracts; from 15 SRs, 57 RCTs were included. Verbatim outcome names were coded to standard outcome names and organised using the Williamson and Clarke taxonomy. Recurrence (98%), progression (74%), treatment response (in CIS studies) (40%), and adverse events (77%) were frequently reported across studies. However, overall (33%) and cancer-specific (33%) survival, treatment completion (17%) and treatment change (37%) were less often reported. Quality of Life (3%) and economic outcomes (2%) were rarely reported. Heterogeneity was evident

throughout, particularly in the definitions of progression and recurrence, and how CIS patients were handled in the analysis of studies with predominantly papillary patients, highlighting further issues with the definition of recurrence and progression vs treatment response for CIS patients. Data reporting was also inconsistent, with some trials reporting event rates at various time-points and others reporting time-to-event with or without Hazard Ratios. Adverse events were inconsistently reported. QoL data was absent in most trials.

**Conclusions:** Heterogenous outcome reporting is evident in NMIBC effectiveness trials. This has profound implications for meta-analyses, SRs and evidence-based treatment decisions. A core outcome set is required to reduce heterogeneity.

**Patient Summary:** ~~Patients with non-muscle-invasive bladder cancer benefit from adjuvant instillation therapies.~~ This systematic review found inconsistencies in outcome definitions and reporting, pointing out the urgent need for a core outcome set to help improve evidence-based treatment decisions.

## 1. Introduction

### Description of the condition

Bladder cancer is the 6th commonest male, and 17th commonest female cancer globally, with the highest incidence rates being observed in Europe and North America.<sup>1</sup> The disease is categorised into two broad stage groupings, non-muscle invasive (NMIBC) and muscle-invasive (MIBC) bladder cancer. Most cases (75-85%) present as NMIBC and these patients typically have a higher long-term survival and a lower cancer specific mortality compared to those with MIBC.<sup>2</sup>

NMIBC is defined as tumour(s) confined to the mucosa or invading the lamina propria.<sup>3</sup> Using the TNM staging system, they are classified as Ta-T1 or Tis (or Cis) N0 M0.<sup>4</sup> NMIBC tumours may be graded using the WHO 1973 or WHO 2004 grading systems – both indicating worse prognosis with increasing grade. Most patients diagnosed with NMIBC is initially treated conservatively (sparing the bladder) with curative intent by transurethral resection of bladder tumour (TURBT). NMIBC is seen as a chronic disease requiring frequent follow-up and repeated TURBTs, making it the most expensive of all cancers to treat from diagnosis to death<sup>5-8</sup> with additional productivity losses and informal care costs.<sup>9</sup> Cumulative costs of care are especially high in intermediate- and high-risk NMIBC due to higher risk of progression to MIBC requiring definitive treatment.<sup>7</sup>

Given the high recurrence rates and the risk of progression to MIBC, NMIBC treatment usually involves adjuvant intravesical instillations with chemotherapy or immunotherapy. The timing, treatment duration, and choice of agent for intravesical therapy is guided by a risk categorisation system which is based upon clinical and pathological factors.<sup>3</sup> For instance, evidence from high quality systematic reviews and meta-analyses shows that a single immediate post-operative instillation of chemotherapy (IPOIC) is well tolerated and clinically effective in reducing recurrences in low risk patients.<sup>10-12</sup> The European Association of Urology (EAU)<sup>3</sup> and the National Institute for Clinical and Healthcare Excellence (NICE)<sup>13</sup> both recommend that eligible patients receive IPOIC. It is considered cost effective for the NHS.<sup>13</sup> Intermediate risk patients may also be given repeated chemotherapy instillations, but their timing and frequency remains undefined<sup>14</sup>. It is recommended that high risk patients are treated with intravesical bacillus Calmette-Guerin (BCG) immunotherapy or be considered for immediate cystectomy.<sup>3</sup> Five-year recurrence and progression rates for patients with stage Ta-T1 bladder cancer treated with 1 to 3 years maintenance BCG are 28-51% and 7-20%, respectively.<sup>15</sup>

### Why it is important to do this review

Inconsistent outcome reporting (different outcomes in different trials) and variability in outcome reporting (same outcomes reported, but different definitions used) become acutely evident when many bladder cancer trials are included in systematic reviews of intervention effectiveness.<sup>16–18</sup> Outcome reporting heterogeneity has been highlighted as a concern within evidence-based medicine generally,<sup>19–22</sup> and has been emphasised as an area for improvement in NMIBC trials by the International Bladder Cancer Group.<sup>23</sup> Heterogeneous outcome reporting and the potential for selective outcome reporting bias in NMIBC trials hinder comparing and contrasting the results of individual trials as well as the publication of unbiased systematic reviews and meta-analyses of the evidence base. As a consequence, making evidence-based recommendations in clinical practice guidelines, translating them into health care policy, and decision-making by clinicians and patients are all hampered.

[Developing core outcome set \(COS\) a solution to reduce outcome heterogeneity, selective outcome reporting bias, and helps to ensure that all trials contribute useable information to the evidence base.](#) A COS is an agreed standardised collection of outcomes which should be measured and reported, as a minimum, in all trials for a specific clinical area.<sup>22</sup> ~~Developing a COS is a solution to reduce outcome heterogeneity, selective outcome reporting bias, and helps to ensure that all trials contribute useable information to the evidence base.~~ Our group has registered a bladder cancer COS development project (B-COS) with the Core Outcome Measures for Effectiveness Trials initiative COS register (<http://www.comet-initiative.org/studies/details/1135>), with the intent to create separate COS for three broad categories of disease: NMIBC, MIBC, and metastatic BC. Within each COS we define the scope with regards to the applicable populations and treatments. After defining the scope of a COS, the next step is to identify existing knowledge regarding outcomes. To meet this requirement, we have aimed to systematically review the outcomes reported in NMIBC effectiveness trials. Our systematic review protocol was registered with PROSPERO ([https://www.crd.york.ac.uk/prospero/display\\_record.php?RecordID=91820](https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=91820)). The reviews for the other parts of the project will be reported separately [as will the subsequent phases of the COS development projects, involving qualitative interview studies with patients, and consensus studies with key stakeholders such as patients and healthcare professionals using Delphi methods to come to consensus on the core outcomes to be measured in future bladder cancer effectiveness trials and audits.](#)

## 2. Methods

### Aims and objectives

The aim was to systematically review outcomes reported in NMIBC effectiveness trials of adjuvant, prophylactic treatment after TURBT.

The objectives were to systematically review:

1. Outcomes reported
2. Outcome definitions (including time points)
3. Outcome assessment methods

### Eligibility Criteria

#### Types of studies

We included phase III randomised controlled trials (RCTs) comparing different adjuvant instillation treatments after TURBT or trials with TURBT alone as a control arm. We limited to RCTs included in systematic reviews of intervention effectiveness as a pragmatic [and efficient](#) way to identify studies

and overview potentially important outcomes. [This is a strategy that has been used in published systematic reviews of outcome reporting heterogeneity where the aim is to overview outcome reporting heterogeneity rather than to find every outcome previously reported.](#)<sup>22,24</sup> All pre phase III trials and all non-randomised designs were excluded. Studies reported only as abstracts were excluded a priori because it was unlikely that all outcomes would be reported in the abstract, and that they would also not provide enough information on the definition and measurement of outcomes reported.

#### Types of participants

We included studies with adult ( $\geq 18$  years) males and females with histologically confirmed urothelial NMIBC, stage Ta or T1 N0 M0, with or without *carcinoma in situ (CIS)*, and all tumour grades (using any grading system). Studies including paediatric patients and patients with MIBC, clinical N+ or M+ were excluded unless outcomes were separately reported and defined for NMIBC patients.

#### Types of interventions and comparators

We included RCTs comparing any type of intravesical adjuvant prophylactic treatments after TURBT and RCTs comparing intravesical treatment after TURBT versus TURBT alone. Studies of oral vitamins or mineral supplements were excluded.

#### Types of Outcomes

We report on all outcomes related to clinical effectiveness including, for example, outcomes related to recurrence, progression, survival and cause of death, local and systemic adverse events and quality of life/patient reported outcomes. Outcome definitions, timepoints, and assessment methods are also reported.

We do not report any estimates of treatment effect for any individual trials and there was no attempt to synthesise aggregated quantitative data.

#### Literature Search

The literature search was undertaken by an experienced information specialist (CY) using the search criteria specified in Appendix 1. Medline, Embase and Cochrane Database of Systematic Reviews (CDSR) were searched for relevant systematic reviews. We also hand-searched the reference sections of relevant international clinical practice guidelines. We restricted to systematic reviews and RCTs published after 2000 to reflect outcomes reported in the current clinical practice. We excluded non-English studies as a pragmatic consideration due to resource restrictions.

An update search was done on 15<sup>th</sup> January 2020.

#### **Data collection and analysis**

##### Selection of studies

Following de-duplication, at least two review authors (DC, SM, SS, IO, EV, RC) independently screened the titles and abstracts of identified systematic reviews for eligibility. The full texts of all potentially eligible publications were retrieved and screened independently by two review authors (DC, SM, SS, IO, EV, RC) using a standardised form, linking together multiple records of the same

study in the process. Any disagreements were resolved by discussion or by consulting a senior review author (RS). Once the list of systematic reviews meeting the inclusion criteria were finalised, a second screening process was initiated whereby the studies included in the systematic reviews were screened against our inclusion criteria. Where lists of studies excluded from the systematic reviews were available, we also screened these in case the studies had been excluded for not reporting on outcomes of interest. In such instances the trial may still have met inclusion criteria for our review. The study selection process is described in the PRISMA flow diagram (Figure 1).<sup>25</sup>

#### Data extraction and management

A standardised data extraction form was developed and piloted. One review author extracted data and a second review author checked data extractions for accuracy (DC, SM, SS, IO, EV, RC). Any disagreements were resolved by discussion or by consulting a third review author.

Data that were extracted included: the study design; countries and institutions where the data were collected; dates defining start and end of patient recruitment and follow-up; how intervention comparator groups were formed; participant demographic and clinical characteristics; eligibility criteria for participants; the numbers of participants who were included in the study, assigned to each intervention comparator group; description of interventions; study funding sources; and ethical approval. All primary and secondary effectiveness outcomes reported, their definitions, and any outcome measurement instruments used were extracted verbatim.

#### Assessment of risk of bias in included studies

Risk of bias assessment is not necessary for systematic reviews undertaken for COS development. Some outcomes may be at risk of detection bias depending on whether they are relatively subjective or objective. Although these aspects were extracted under the 'definition' or 'measurement' fields in the data extraction form, this is out of the scope of this phase of our project. They will be investigated in a subsequent phase whereby we will assess the psychometric properties of the various outcome measurements and seek consensus on the most appropriate and feasible definitions and measurements.<sup>26,27</sup>

#### Data synthesis

Verbatim outcome names were recoded to common names. This was done by categorising outcomes referring to the same underlying constructs under a common term. For example, "survival rates", "overall survival", "number of deaths at median follow up" and "mortality rate" all refer to the concept of 'overall survival' and were coded as such. The outcome and domain coding process was inductive and iterative. Coded outcomes were further grouped in broader domains using the standardised Williamson and Clarke Taxonomy (W/C Taxonomy).<sup>28</sup>

### **3.0 Evidence synthesis**

#### ***Characteristics of the included studies***

Our initial search for relevant systematic reviews yielded 807 abstracts, of which 639 remained after removing duplicates. In total, 100 full-text SRs were assessed and 19 SRs, including 14 meta-analyses, were included. Four SRs included only previously identified RCTs and these SRs were not utilised further (Supplemental table 1). From 15 SRs published between years 2010-2018, 106 full-texts of RCTs were screened and 57 eligible RCTs were finally included (see PRISMA flow diagram, Fig. 1).

An overview of the included studies' populations, stage and grade, instillation treatments and number of outcome domains reported is shown in Table 1. Overall, 32 studies included patients with papillary only tumors, while 25 studies included a mixed population of patients with CIS with/without papillary tumors. There were 11 "single-instillation" trials, 12 "single instillation followed by induction course" trials, 27 "maintenance instillation" trials and 7 trials comparing instillations with different schedules.

In all studies, patients were followed up at regular intervals in the same and largely accepted manner: urinary cytology, cystoscopy and if necessary, by taking biopsies from the urinary bladder.<sup>3</sup>

### ***Heterogeneity in outcome reporting, detection, and definitions***

The outcomes were organised into the 10 domains in the W/C taxonomy [27]: "recurrence", "progression", "treatment response" (for CIS), "cancer-specific survival", "overall survival", "adverse events", "completion/adherence", "treatment failure/change of treatment", "quality of life" and "health economics" (Table 2).

As seen in Table 2, tumor related outcomes such as recurrence (98%), progression (74%), treatment response (in CIS studies) (40%), and adverse events (77%) were frequently reported across studies. However, overall (33%) and cancer specific (33%) survival, treatment completion (17%) and treatment change (37%) were less often reported. Quality of Life (3%) and economic outcomes (2%) were rarely reported.

### ***Tumor related outcomes***

The heterogeneity in the definition and reporting of recurrence and progression in studies that recruited patients with papillary tumors only, and also treatment response in patients with CIS with or without papillary tumors, are shown in Tables 3 and 4, respectively.

#### ***Recurrence***

Recurrence was reported in 56 (98%) of 57 trials (Tables 1,3,4), with 35 different verbatim names (Table 5), often related to the definition. The definition of recurrence was missing in 8/56 (14%) studies and in the others, variations of the percent of recurrences at a given time point or as a time to event outcome were used, but no consistent way of defining and measuring recurrence was used overall. Furthermore, in studies that reported both progression and recurrence, progression as the first event was regarded as a recurrence event in 12 studies and in 34 others it was not.

#### ***Progression***

Of 57 studies, 42 (74 %) reported bladder cancer progression. Definition for progression was given in 41/42 (97%) studies with a large variability in definition. A common threshold for "progression" was  $\geq pT2$  in 16 (38%) studies, with 2 of them also classifying CIS as a progression. As an example of inconsistency in verbatims used, "progression to MIBC" was used in the definition in 31/42 (74%) studies, with 22 of those further including metastases. Ta->T1 and T1->MIBC were considered progression in 4/42 (9%) studies (Tables 3 and 4).

#### ***Treatment response***

Treatment response in patients with CIS was reported in 10 (40%) of 25 studies (Table 2). There was heterogeneity in what time-point was considered to assess the response to treatment. de Reijke et al defined and reported "complete response", "partial response", "no change" and "progression".<sup>29</sup> The rest of the studies reported only complete response to treatment.

The time-point to assess complete response varied largely, ranging from 3 months from enrollment up to 12 months.

Eight different outcomes were included in the “Treatment response (for CIS)” domain (Tables 4 and 5).

Treatment relapse after complete response was described in three trials (Table 4).

### **Death**

A survival outcome was reported in 44/57 (40%) of studies; equally common were cancer-specific survival and overall survival, each reported in 19 (33%) studies. Ten and eleven different verbatim names were used to report overall survival and cancer-specific survival, respectively (Tables 2, 5).

### **Adverse events**

Adverse events (AEs) were heterogeneously defined. In 12 of the 44 studies (27%) reporting AEs, there was no definition of an AE, and overall 24 different definitions/instruments were used. Studies reporting AEs used unique systems to categorise the type of AE or grade the severity of the AEs, and made no reference to a standardised reporting system. Across 10 studies, 3 standardised AE reporting instruments were used, but these did not include some of the most relevant AEs for intravesical instillations:

- NCI-CTCAE (Common Terminology Criteria of Adverse Events),
- WHO toxicity grading scale,
- WHO-ART (1979 WHO Adverse Reaction Terminology)

Adverse events were further grouped in numerous ways, e.g. local or systemic toxicity, constitutional symptoms, laboratory abnormality, death, and treatment interruption due to AEs (Table 5). Detailed lists of how AEs were described and reported are provided in Supplementary table 2.

In 25 of the 44 studies (57%) specific AEs were not listed; instead, authors reported either only local toxicities, or major/severe/more common side-effects or AEs that resulted in treatment interruption. Five of these 25 studies did not report the list of individual toxicities at all; instead, authors presented only the frequency and percentage [n (%)] of any AEs which occurred. Furthermore, poor treatment compliance related to AEs was not consistently reported.

### **Completion/adherence**

Adherence to completion of all planned instillations was at least partially reported in 10/57 (17%) studies: six studies concerning maintenance instillations, two “induction course” studies, and two studies comparing induction to maintenance. None of the single-instillation studies reported completion rates. Four studies gave a comprehensive overview of the reasons for treatment discontinuation. The author definitions for instillation treatment completion are reported in Table 5.

### **Treatment failure/change of treatment**

21/57 studies (37%) reported treatment failure and/or the need to change from instillations to a different treatment. 21 studies specified the treatment that was given after instillations were discontinued:

- radical cystectomy (RC) (14/21 studies)
- RC and/or radiotherapy (RT) (4/12 studies)



- TURBT (1/21)
- RC, TURBT+RT, chemotherapy (1/21)
- “non-allowed instillations” (1/21)

### Global quality of life

Two studies measured and reported patient experience during the instillations; Koga et al by measuring QoL, and Huang et al by evaluating instillation related pain/irritation.<sup>30,31</sup>

In the study by Koga et al, QoL was assessed according to the Japanese version of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30) v2.0. QoL was assessed before induction therapy, after the 5th instillation of induction therapy, 4 weeks after the completion of induction therapy, and 14 months after randomization.<sup>30</sup>

Huang et al evaluated the effect of hyaluronic acid in reducing pirarubicin instillation related side-effects. A visual analog scale (VAS) was used daily to evaluate pain.<sup>31</sup>

### Resource use (health economics)

Only one study evaluated the costs related to the treatment. Berrum-Svennung et al randomized BC patients to one instillation of epirubicin or placebo after TURBT and evaluated cancer recurrences. They also calculated the cost of delivering a single instillation during the initial treatment and as first recurrences occurred.<sup>32</sup>

### Discussion

This is the first study to systematically and comprehensively overview the extent of outcome reporting, measurement, and definition heterogeneity in the setting of adjuvant treatments for NMIBC.

Recurrence was frequently reported in the included RCTs; yet, some studies did not define it. In those that did, there was variability in the names that were used ~~(for example “disease-free interval”, “disease-free survival”, “time to recurrence”, and “recurrence-free survival”)~~, the definitions and the reporting, ~~with some reporting hazard ratios or median time to event, or event rates at inconsistent time points (e.g. 1,2,3,5,10 years)~~. Most concerning, however, was the variation in how progression was handled in the analysis of recurrence. In studies where progression as the first event was counted as a recurrence, the measure provided is qualitatively and quantitatively different from those where recurrence was more narrowly defined as the re-appearance of a non-muscle invasive tumour. To overlook this subtlety runs the risk of not comparing like with like across studies, or statistically pooling aggregated results in a potentially misleading way.

Progression was also frequently reported, but again the definitions were inconsistent across trials, ~~with some using progression to MIBC as the threshold, while others considered an increase in stage from Ta-T1 and/or an increase in grade as progression~~. Worsening of the disease leads to a change in treatment strategy, and that was also inconsistently reported. It is also crucial to point out whether assessment of progression has been made based on imaging (e.g. CT or MRI), TURBT or radical cystectomy. As only four studies gave a comprehensive overview of reasons to change the treatment strategy, there is a high risk of getting misleading results. If prior to progression, patients die due to an unrelated cause, or undergo cystectomy (for example due to recurrent high grade T1 disease),

then the progression rates at specific time points will be different according to whether the death and cystectomy have been counted as a competing risk (cumulative incidence function) or simply as censored (Kaplan-Meier curve). Equally important is to highlight how patients are followed for the efficacy outcomes in case the treatment has been stopped due to side-effects. There may also be a difference in outcomes according to whether the results are reported in all randomized patients (intent to treat analysis) or only in eligible patients who have been treated according to the protocol (per protocol analysis).

Treatment response in patients with CIS in specific was evaluated and reported in only 40% of studies. The rest of the studies recruiting patients with CIS considered CIS as papillary tumors, and reported only recurrence or/and progression. However, CIS additional diagnostic challenges and may have a very different disease course than papillary tumors do: as such, separate approaches to measure and define their outcomes should be applied.<sup>23</sup>

The most heterogenous outcome was AEs, evident in the many categorizations and instruments used to record AEs, and in the system level subgroupings chosen by trialists. Unfortunately, many of these were not optimal for instillation-related AEs. Whilst in some instances it may be possible for systematic reviewers to recode lists of AEs (if they are provided) to a common standardized toxicity classification system, this is a poor excuse for lack of standardization in primary trials and needlessly adds time and complexity to the critical interpretation of the evidence base. Poor treatment compliance reporting is likely to confound other cancer related outcomes such as recurrence, progression and overall survival.

Perhaps the most alarming finding is that QoL is conspicuously missing. Instillation treatments are demanding for patients and it would be very important to understand all the consequences (both oncological and QoL-related) for patients before the decision about treatment is made. A recent investigation of QoL in bladder cancer patients compared to a matched sample of older adults without bladder cancer in a US population found significant declines in health-related QoL (HRQoL) scores over time in the physical, mental and social components of the SF-36.<sup>33</sup> The EORTC Quality of Life Group also developed an externally validated QLQ-BLS24 questionnaire for NMIBC.<sup>34</sup> In a systematic review, Mason and colleagues used the COSMIN checklist to evaluate the psychometric properties of PROMs used in bladder cancer populations, of which two of the 15 included PROMs were NMIBC-specific (QLQ-BLS24 and CAVICAVEMNI).<sup>35,36</sup> Of note, they found that no existing PROM stood out as the most appropriate measure of QoL in any bladder cancer populations and although further validation studies are required generic PROMs, cancer-generic PROMs and bladder cancer-specific PROMs will currently provide the most robust picture. This is a very important study to a subsequent phase of our COS development as most existing cancer COS have included QoL and it is anticipated NMIBC patients will also prioritise this, encompassing urinary, bowel and sexual function, as a critically important outcome domains.

Without having included NMIBC patients in a qualitative study of their experiences of bladder cancer and its treatments, it cannot yet be known which outcomes are of most importance to them, or if they are adequately captured in current trials, but it is discouraging that so few trials routinely include patient reported outcome measures (PROMs).

Health economics was considered in only one RCT, which calculated costs of single instillation.<sup>32</sup> Bladder cancer, especially NMIBC, contributes significantly to healthcare costs due to intense surveillance strategies and its potential to recur and progress.<sup>8,37</sup> This should be considered when treatments and outcomes are compared.

Kamat et al provided recommendations on NMIBC intervention trial designs, eligibility criteria, and 'clinically meaningful' effect size thresholds for outcomes.<sup>23</sup> Likewise, Lamm et al suggested a change in definition for progression in NMIBC.<sup>38</sup> These initiatives are important to bear in mind for

subsequent phases of our project. Once the outcomes considered core by all stakeholders (e.g. patients, urologists, oncologists, nurses, payers, methodologists) are known (i.e. *what* to measure)<sup>22</sup> then we will turn attention to definitions and measurement tools (i.e. *how* to measure)<sup>39</sup> whilst again including key stakeholders. Importantly, these initiatives, in conjunction with ours, show that there is an acknowledgement of problems with the evidence base and a desire to do improvements.

## Limitations

The decision to exclude phase I and II trials (phases before determining the therapeutic effect of the drug) and to exclude all non-randomised designs may have limited the chance to capture longer-term and patient reported outcomes relating to function and QoL. However, in subsequent phases of the project, such as in Delphi survey and consensus meetings, participants will have an opportunity to propose 'new' outcomes not already considered for prioritisation, therefore we consider that the risk of having missed outcomes is minimal, and that we have carried out a pragmatic trade-off against the resource implication of including all study designs.

## Conclusions

We have shown that there is inconsistency in outcome reporting and variation in definitions in randomized trials comparing adjuvant treatments in NMIBC patients. This situation makes comparing the results of individual studies difficult, and makes their statistical combination challenging, impossible, or inappropriate; hence, providing summaries of the evidence which are, at best, unwieldy and at worst misleading, making evidence-based treatment recommendations difficult. A core outcome set, incorporating the views of a variety of stakeholders such as urologists, oncologists, methodologists and, most importantly, patients, is urgently required.

Acknowledgements: The authors have no acknowledgements

Funding: The authors report no funding

Author Contributions:

Erik Veskimae - performance of work; interpretation or analysis of data; writing the article.

Selvarani Subbarayan – performance of work; interpretation or analysis of data; writing the article.

Riccardo Campi - performance of work; interpretation or analysis of data; writing the article.

Domitille Carron - performance of work; interpretation or analysis of data; writing the article.

Muhammad Imran Omar – conception; performance of work; interpretation or analysis of data; writing the article.

Cathy Yuan- performance of work; interpretation or analysis of data; writing the article.

Konstantinos Dimitropoulos - performance of work; interpretation or analysis of data; writing the article.

Mieke Van Hemelrijck - interpretation or analysis of data; writing the article.

Richard T. Bryan - interpretation or analysis of data; writing the article.

James N'Dow - interpretation or analysis of data; writing the article.

Marek Babjuk - interpretation or analysis of data; writing the article.

J. Alfred Witjes - interpretation or analysis of data; writing the article.

Richard Sylvester – conception; performance of work; interpretation or analysis of data; writing the article.

Steven MacLennan – conception; performance of work; interpretation or analysis of data; writing the article.

#### Conflict of interest statements

Erik Veskimäe – Has no conflict of interest to report

Selvarani Subbarayan – Has no conflict of interest to report

Riccardo Campi - Has no conflict of interest to report

Domitille Carron - Has no conflict of interest to report

Muhammad Imran Omar – Has no conflict of interest to report

Cathy Yuan- Has no conflict of interest to report

Konstantinos Dimitropoulos - Has no conflict of interest to report

Mieke Van Hemelrijck - Has no conflict of interest to report

Richard T. Bryan - Reports other from Janssen EMEA, grants from UroGen Pharma, grants from QED Therapeutics, outside the submitted work

James N'Dow - Has no conflict of interest to report

Marek Babjuk - Has no conflict of interest to report

J. Alfred Witjes - Has no conflict of interest to report

Richard Sylvester – Has no conflict of interest to report

Steven MacLennan – Has no conflict of interest to report

#### References

1. Richters A, Aben KKH, Kiemenev LALM. The global burden of urinary bladder cancer: an update. *World J Urol.* 2020;38(8):1895-1904. doi:10.1007/s00345-019-02984-4
2. van Rhijn BWG, Burger M, Lotan Y, et al. Recurrence and Progression of Disease in Non-Muscle-Invasive Bladder Cancer: From Epidemiology to Treatment Strategy. *Eur Urol.* 2009;56(3):430-442. doi:10.1016/j.eururo.2009.06.028
3. Babjuk M, Burger M, Compérat EM, et al. European Association of Urology Guidelines on Non-muscle-invasive Bladder Cancer (TaT1 and Carcinoma In Situ) - 2019 Update. *Eur Urol.* 2019;76(5):639-657. doi:10.1016/j.eururo.2019.08.016

4. Brierley J, Gospodarowicz M, Wittekind C. *TNM Classification of Malignant Tumours, 8th Edition*. Wiley-Blackwell; 2017.
5. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer: A comprehensive review of the published literature. *Pharmacoeconomics*. 2003;21(18):1315-1330. doi:10.1007/BF03262330
6. Riley GF, Potosky AL, Lubitz JD, Kessler LG. Medicare payments from diagnosis to death for elderly cancer patients by stage at diagnosis. *Med Care*. 1995;33(8):828-841. doi:10.1097/00005650-199508000-00007
7. Mossanen M, Wang Y, Szymaniak J, et al. Evaluating the cost of surveillance for non-muscle-invasive bladder cancer: an analysis based on risk categories. *World J Urol*. 2019;37(10):2059-2065. doi:10.1007/s00345-018-2550-x
8. Svatek RS, Hollenbeck BK, Holmäng S, et al. The Economics of Bladder Cancer: Costs and Considerations of Caring for This Disease. *Eur Urol*. 2014;66(2):253-262. doi:10.1016/j.eururo.2014.01.006
9. Leal J, Luengo-Fernandez R, Sullivan R, Witjes JA. Economic Burden of Bladder Cancer Across the European Union. *Eur Urol*. 2016;69(3):438-447. doi:10.1016/j.eururo.2015.10.024
10. Sylvester RJ, Oosterlinck W, Holmang S, et al. Systematic Review and Individual Patient Data Meta-analysis of Randomized Trials Comparing a Single Immediate Instillation of Chemotherapy After Transurethral Resection with Transurethral Resection Alone in Patients with Stage pTa-pT1 Urothelial Carcinoma of the Bladder: Which Patients Benefit from the Instillation? *Eur Urol*. 2016;69(2):231-244. doi:10.1016/j.eururo.2015.05.050
11. Perlis N, Zlotta AR, Beyene J, Finelli A, Fleshner NE, Kulkarni GS. Immediate post-transurethral resection of bladder tumor intravesical chemotherapy prevents non-muscle-invasive bladder cancer recurrences: an updated meta-analysis on 2548 patients and quality-of-evidence review. *Eur Urol*. 2013;64(3):421-430. doi:10.1016/j.eururo.2013.06.009
12. Abern MR, Owusu RA, Anderson MR, Rampersaud EN, Inman BA. Perioperative intravesical chemotherapy in non-muscle-invasive bladder cancer: a systematic review and meta-analysis. *J Natl Compr Cancer Netw JNCCN*. 2013;11(4):477-484. doi:10.6004/jnccn.2013.0060
13. Bladder cancer: diagnosis and management of bladder cancer: © NICE (2015) Bladder cancer: diagnosis and management of bladder cancer. *BJU Int*. 2017;120(6):755-765. doi:10.1111/bju.14045
14. Sylvester RJ, Oosterlinck W, Witjes JA. The Schedule and Duration of Intravesical Chemotherapy in Patients with Non–Muscle-Invasive Bladder Cancer: A Systematic Review of the Published Results of Randomized Clinical Trials. *Eur Urol*. 2008;53(4):709-719. doi:10.1016/j.eururo.2008.01.015
15. Cambier S, Sylvester RJ, Collette L, et al. EORTC Nomograms and Risk Groups for Predicting Recurrence, Progression, and Disease-specific and Overall Survival in Non-Muscle-invasive Stage Ta-T1 Urothelial Bladder Cancer Patients Treated with 1-3 Years of Maintenance Bacillus Calmette-Guérin. *Eur Urol*. 2016;69(1):60-69. doi:10.1016/j.eururo.2015.06.045
16. Hernández V, Espinos EL, Dunn J, et al. Oncological and functional outcomes of sexual function–preserving cystectomy compared with standard radical cystectomy in men: A systematic

review. *Urol Oncol Semin Orig Investig*. 2017;35(9):539.e17-539.e29. doi:10.1016/j.urolonc.2017.04.013

17. Veskimäe E, Neuzillet Y, Rouanne M, et al. Systematic review of the oncological and functional outcomes of pelvic organ-preserving radical cystectomy (RC) compared with standard RC in women who undergo curative surgery and orthotopic neobladder substitution for bladder cancer. *BJU Int*. 2017;120(1):12-24. doi:10.1111/bju.13819
18. Bruins HM, Veskimäe E, Hernandez V, et al. The impact of the extent of lymphadenectomy on oncologic outcomes in patients undergoing radical cystectomy for bladder cancer: a systematic review. *Eur Urol*. 2014;66(6):1065-1077. doi:10.1016/j.eururo.2014.05.031
19. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012;13:132. doi:10.1186/1745-6215-13-132
20. Williamson P, Altman D, Blazeby J, Clarke M, Gargon E. Driving up the Quality and Relevance of Research Through the Use of Agreed Core Outcomes. *J Health Serv Res Policy*. 2012;17(1):1-2. doi:10.1258/jhsrp.2011.011131
21. Tunis SR, Clarke M, Gorst SL, et al. Improving the relevance and consistency of outcomes in comparative effectiveness research. *J Comp Eff Res*. 2016;5(2):193-205. doi:10.2217/ce-2015-0007
22. Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials*. 2017;18(Suppl 3):280. doi:10.1186/s13063-017-1978-4
23. Kamat AM, Sylvester RJ, Böhle A, et al. Definitions, End Points, and Clinical Trial Designs for Non-Muscle-Invasive Bladder Cancer: Recommendations From the International Bladder Cancer Group. *J Clin Oncol Off J Am Soc Clin Oncol*. 2016;34(16):1935-1944. doi:10.1200/JCO.2015.64.4070
24. Bruce I, Harman N, Williamson P, et al. The management of Otitis Media with Effusion in children with cleft palate (mOMEnt): a feasibility study and economic evaluation. *Health Technol Assess*. 2015;19(68):1-374. doi:10.3310/hta19680
25. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097
26. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. *Trials*. 2016;17(1):449. doi:10.1186/s13063-016-1555-2
27. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549. doi:10.1007/s11136-010-9606-8
28. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol*. 2018;96:84-92. doi:10.1016/j.jclinepi.2017.12.020
29. de REIJKE TM, Kurth KH, Sylvester RJ, et al. BACILLUS CALMETTE-GUERIN VERSUS EPIRUBICIN FOR PRIMARY, SECONDARY OR CONCURRENT CARCINOMA IN SITU OF THE BLADDER: RESULTS OF A EUROPEAN ORGANIZATION FOR THE RESEARCH AND TREATMENT OF CANCER—GENITO-URINARY GROUP PHASE III TRIAL (30906). *J Urol*. 2005;173(2):405-409.

doi:10.1097/01.ju.0000150425.09317.67

30. Koga H, Ozono S, Tsushima T, et al. Maintenance intravesical bacillus Calmette-Guérin instillation for Ta, T1 cancer and carcinoma in situ of the bladder: Randomized controlled trial by the BCG Tokyo Strain Study Group: Maintenance intravesical BCG. *Int J Urol*. 2010;17(9):759-766. doi:10.1111/j.1442-2042.2010.02584.x
31. Huang W, Wang F, Wu C, Hu W. Efficacy and safety of pirarubicin combined with hyaluronic acid for non-muscle invasive bladder cancer after transurethral resection: a prospective, randomized study. *Int Urol Nephrol*. 2015;47(4):631-636. doi:10.1007/s11255-015-0940-1
32. Berrum-Svennung I, Granfors T, Jahnsen S, Boman H, Holmäng S. A Single Instillation of Epirubicin After Transurethral Resection of Bladder Tumors Prevents Only Small Recurrences. *J Urol*. 2008;179(1):101-106. doi:10.1016/j.juro.2007.08.166
33. Smith AB, Jaeger B, Pinheiro LC, et al. Impact of bladder cancer on health-related quality of life. *BJU Int*. 2018;121(4):549-557. doi:10.1111/bju.14047
34. Blazeby JM, Hall E, Aaronson NK, et al. Validation and Reliability Testing of the EORTC QLQ-NMIBC24 Questionnaire Module to Assess Patient-reported Outcomes in Non-Muscle-invasive Bladder Cancer. *Eur Urol*. 2014;66(6):1148-1156. doi:10.1016/j.eururo.2014.02.034
35. Mason SJ, Catto JWF, Downing A, Bottomley SE, Glaser AW, Wright P. Evaluating patient-reported outcome measures (PROMs) for bladder cancer: a systematic review using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. *BJU Int*. 2018;122(5):760-773. doi:10.1111/bju.14368
36. Abáigar-Pedraza I, Megías-Garrigós J, Sánchez-Payá J. Cuestionario de calidad de vida para pacientes con cáncer de vejiga no músculo invasivo. *Actas Urol Esp*. 2016;40(4):251-257. doi:10.1016/j.acuro.2015.12.001
37. Cox E, Saramago P, Kelly J, et al. Effects of Bladder Cancer on UK Healthcare Costs and Patient Health-Related Quality of Life: Evidence From the BOXIT Trial. *Clin Genitourin Cancer*. 2020;18(4):e418-e442. doi:10.1016/j.clgc.2019.12.004
38. Lamm D, Persad R, Brausi M, et al. Defining progression in nonmuscle invasive bladder cancer: it is time for a new, standard definition. *J Urol*. 2014;191(1):20-27. doi:10.1016/j.juro.2013.07.102
39. COSMIN handbook.  
[https://fac.ksu.edu.sa/sites/default/files/cosmin\\_checklist\\_manual\\_v9.pdf](https://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf).