2016

From 'D' to 'I': A Critique of the Current United States Preventive Services Task Force Recommendation for Testicular Cancer Screening

Michael J. Rovito  
*University of Central Florida*

Janna Manjelievskaia  
*University of the Sciences in Philadelphia*

James E. Leone  
*Bridgewater State University, james.leone@bridgew.edu*

Michael J. Lutz  
*Michigan Institute of Urology*

Ajay Nangia  
*University of Kansas Medical Center*

Follow this and additional works at: [http://vc.bridgew.edu/mahpls_fac](http://vc.bridgew.edu/mahpls_fac)

Part of the [Community Health and Preventive Medicine Commons](http://vc.bridgew.edu/mahpls_fac)

Virtual Commons Citation  
Available at: [http://vc.bridgew.edu/mahpls_fac/109](http://vc.bridgew.edu/mahpls_fac/109)

This item is available as part of Virtual Commons, the open-access institutional repository of Bridgewater State University, Bridgewater, Massachusetts.
From ‘D’ to ‘I’: A critique of the current United States preventive services task force recommendation for testicular cancer screening

Michael J. Rovito, Janna Manjelievskai, James E. Leone, Michael J. Lutz, Ajay Nangia

In 2004, the United States Preventive Services Task Force (USPSTF) gave testicular cancer (TCA) screening a ‘D’ recommendation, discouraging the use of this preventive service. The USPSTF suggested that screening, inclusive of testicular self-examination (TSE) and clinician examination, does not reduce TCA mortality rates and that the high risk of false positives could serve as a detriment to patient quality of life. Others suggest that TCA screening is ineffective at detecting early-stage cases of TCA and readily highlights a lack of empirical evidence demonstrating said efficacy. These assertions, however, stand in stark contrast to the widely held support of TCA screening among practicing public health professionals, advocacy groups, and clinicians.

In this present study, a review was conducted of the methods and processes used by the USPSTF in their 2011 reaffirmation of the ‘D’ grade recommendation. The evidence base and commentary offered as to why TSE, as part of the overall recommendation for TCA screening, was given a ‘D’ grade were analyzed for logical reasoning and methodological rigor.

Considering the methodological flaws and the veritable lack of evidence needed to grant a conclusive recommendation, the question is raised if the current ‘D’ grade for TCA screening (i.e. discourage the use of said service) should be changed to an ‘I’ statement (i.e. the balance of benefits and harms is indeterminate). Therefore the purpose of this paper is to present the evidence of TCA screening in the context of efficacy and prevention in order for the field to reassess its relative value.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

The United States Preventive Services Task Force (USPSTF, 2004; 2011) suggests that there is a lack of available evidence demonstrating how routine testicular cancer (TCa) screening (including both testicular self-examination [TSE] and clinician examination) has greater yield and/or accuracy for detecting TCa at more curable stages (USPSTF, 2011). The Task Force also claims that, generally, TCa is >90% curable and that TSE screening is unlikely to offer meaningful health benefits. One adverse outcome readily offered as evidence that TSE screening (i.e. TSE) should not be recommended is the potential onset of anxiety associated with a false-positive result. Essentially, according to the USPSTF, among others (e.g. Lin and Sharangpani, 2010), TSE and clinician examinations have limited value. This position, however, is grounded in limited evidence and fails to take into account the potential benefits of TCa screening.

Incidence rates of TCa are rising among the 15 to 54 year-old demographic, but primarily affect those under the age of 40 (Kennett et al., 2014). Howlader et al. (2013) indicates that TCa cases have been rising ~1% each year in the past decade. As there are few known risk factors for TCa outside of age (Znaor et al., 2014), cryptorchidism (Lip et al., 2013), or family history of the disease (Kharazmi et al., 2015), it is wise to operate under the assumption that all males are at-risk for developing the disease. Those males who lay claim to one or more of the aforementioned risk-factors could be labeled ‘high-risk’, but the relationship between TCa manifestation and said factors remain spurious at best due to the lack of research conducted highlighting those associations.

It is the collective wisdom of these authors that all males receive multifactorial benefits from regular TSE performance, inclusive of decreased mortality from the disease (see Rovito et al., 2015). The issue, however, for both sides of this debate, but more so serving as the onus for the anti-TSE camp, that when speaking exclusively about TCa mortality reduction, there are zero studies conducted among asymptomatic males demonstrating the harms and/or benefits received from testicular examination, either clinician or self-examination. In essence, the evidence is insufficient to point one way or the other. Hence, the current D-grade recommendation, according to the definitions used by the Task Force themselves, is erroneously granted to TSE. Due to the dearth of evidence between physical examination of testicles as a preventive measure to decrease TCa mortality, it is more appropriate to grant an I-statement recommendation.

The following discussion will highlight the spurious nature of the data used by the USPSTF to discourage TSE screening, as well as the inconsistencies in methodological rigor used to create its current ‘D’ grade recommendation. As the criteria used to create ratings are limited in scope with some having little, if any, relevance in the decision-making process to determine TSE screening’s worth, these authors advocate for a reassessment of the current methodology used when creating recommendations in the absence of solid evidence. Finally, these authors question the appropriateness of the USPSTF (2011)’ ‘D’ grade for TSE screening and lend support for Rovito (2016) argument for the inclusion of TSE in a standard of care as the potential harms associated with the ‘D’ grade are a cause for concern for male lifespan health.

1.1. Overview of the recommendation process and current support

Recommendations made by the USPSTF (2011) are based on explicit criteria. An independent panel of experts in primary care and prevention systematically review the available evidence of effectiveness on a particular topic and develop recommendations for clinical preventive services accordingly (Agency of Healthcare Research and Quality, 2012). Other experts in the field outside of primary care are invited to provide peer review of existing evidence summaries and draft recommendations (Siu et al., 2015). USPSTF panels tend to be conservative in their recommendation statements, relying solely on available scientific evidence. Their approach differs from other bodies that develop clinical practice guidelines (CPGs), which may rely on expert opinion and clinical judgment in the absence of randomized controlled trials (RCTs) (Goolsby, 2002).

One of the main challenges in developing recommendations is deciding which position to take when the evidence is inadequate and lacking, as is the case for TSE screening, and more specifically, TSE. RCTs are generally regarded as the strongest evidence base for providing an intervention by the USPSTF (2011). Yet, even the Task Force acknowledges that this standard of evidence is unattainable for a majority of clinical preventive services. Recognizing this limitation, non-RCT study designs also are included in the evidence base used by the USPSTF (Petitti et al., 2009). The USPSTF (2011) considers indirect evidence in such cases where a ‘chain of evidence’ is created within an analytic framework to inform the recommendation (Petitti et al., 2009).

Some have advocated that the Task Force provide ‘clinical options’ (especially if the harms and costs with performing a particular service are minimal) or that services, which have not been adequately studied, should not be recommended (Woolf and Atkins, 2001). Some suggest that a neutral stance should be taken (meaning not recommending for or against a service) until better evidence is available or that those who are deemed high-risk should be informed of the benefits of performing regular TSE (Woolf and Atkins, 2001; American Urological Association [AUA], 2014). Others explicitly state (i.e. the Society for Adolescent Health and Medicine, 2012) their support for TSE screening or suggest that identified high-risk males (i.e. Caucasian race, being between the ages of 15–40, family history of the disease, and/or the occurrence of cryptorchidism) should ‘seriously’ consider performing monthly exams (ACS, 2015).

1.2. Past and present USPSTF recommendations for TCa screening

Calongo (2005) states that in 1996, the USPSTF found that existing evidence to recommend either for or against routine screening for TCa was insufficient as it pertains to asymptomatic men. TSE screening was given a ‘C’ grade, indicating the reviewing body was not in a position to recommend promoting the behavior or not, thus leaving the decision to the patient and provider. The reviewers included a caveat that for males who were deemed high-risk for TCa, discussions about screening (either TSE or physician exams) can be carried out. In other words, there was insufficient evidence to fully commit to a more positive or more negative TSE recommendation. If, however, a practitioner identifies a male at high risk for developing TCa, then a conversation about TSE is permissible, which is much akin to the AUA’s (2014) current position.

The USPSTF (2013) grade definitions have since changed, where a ‘C’ grade currently indicates that the service should be offered to select patients depending on individual circumstances. Consequently, the USPSTF reassessed the ‘C’ grade for TCa screening in 2004 and came to the conclusion that they found at least fair evidence that it (i.e. TSE) is ineffective and that the harms outweigh the benefits. They specifically argued that no evidence has been produced from appropriate study designs (i.e. RCTs) demonstrating a significant decrease in TCa mortality stemming from the promotion of screening among asymptomatic males. They gave TCa screening a ‘D’ grade, which is defined as “moderate or high certainty that the service has no net benefit and that the harms outweigh the benefits” (USPSTF, 2013).

A 2008 reaffirmation request brought about Lin and Sharangpani’s (2010) rubberstamping of the TCa screening ‘D’ grade. The authors based their decision upon the high cure rates of TCa (even in later stages of the disease), the potential for false-positive anxiety, the lack of evidence demonstrating TCa screening’s effectiveness in reducing mortality, and the potential of increasing costs due to confirmatory procedures (i.e. ultrasound, biopsies, etc.). In 2011, the USPSTF reissued the ‘D’ grade for the provision of TCa screening by self- or clinician examination.
2. Inconsistencies in logic, language, and methodology

Upon careful consideration of the arguments provided by the USPSTF (2011), these authors suggest that the Task Force’s rationale is flawed, and therefore, insufficient to offer a conclusive recommendation concerning TCa screening. The following is an overview of the logic provided by the Task Force and subsequent gaps in its foundational arguments.

2.1. Overview of the TCa screening recommendation search criteria

The USPSTF presented its final evidence review for TCa screening in 2011. The reaffirmation update was based on a search for new, high-quality studies that may have potentially changed the previous recommendation. However, in conducting their literature search, Lin and Sharangpani (2010), discovered that no such studies met the inclusion criteria to address the primary questions assessing the benefits and harms of screening asymptomatic men for TCa. Of the 113 potentially relevant articles that were initially identified, 99 were excluded on title, leaving 14 articles that underwent an abstract review. The most common reason for exclusion was that testing or interventions were not performed in asymptomatic populations. The remaining three articles (two of which were not about screening) reached the full-text review stage (Lin and Sharangpani, 2010). This lack of evidence was also highlighted by Ilic and Misso (2011) who suggest that no published RCTs exist evaluating the effectiveness TCa screening. In summary, the Task Force failed to identify any sources that could lend conclusive evidence to support or not support TCa screening as defined by their own ‘D’ rating criteria. The current recommendation against TCa screening is essentially based on three arbitrarily chosen citations with little rationale to explain the process.

This lack of transparency and accountability has been noted in other recommendations from the USPSTF, such as prostate cancer screening (Kaffenberger and Penson, 2014). Critics of the USPSTF evidence review and recommendation process in regard to prostate cancer point to critical misinterpretations of key evidence regarding benefits of screening and selective publications that overstate the harms (Cooperberg 2014). With zero studies that explicitly point to either the harms or benefits of TCa screening, the current ‘D’ grade is an unjustified overstatement of the potential harms, with little consideration to the various benefits of self- or clinician screening.

2.2. A lack of evidence

The USPSTF (2011) indicates that no evidence exists indicating that TCa screening would improve health outcomes. As indicated previously, however, no hard evidence exists showcasing harms associated with TSE or clinician examination, at least to the knowledge of these authors. Interestingly, the Task Force themselves, as well as supporters of the ‘D’ grade, actually admit to the lack of evidence showcasing not just TCa screening’s benefits (their fulcrum argument), but also its harms.

For example, Lin and Sharangpani (2010) indicate that they failed to identify any evidence directly discussing TCa screening harms and benefits, while the American Cancer Society (2015) states that the lack of evidence prevents them from forming any concrete recommendation about TCa screening, although, as discussed prior, they list on their website that those deemed high risk should consider performing monthly exams. The National Cancer Institute (2015), which distributes information on how to perform TSE, suggests that TCa screening harms are poorly quantified and that there is no available evidence that allows the field to determine TCa screening’s effectiveness in reducing mortality rates associated with the disease. The USPSTF (2011) itself stated that the available evidence was inadequate to determine whether asymptomatic patient TCa screening has greater yield or accuracy for detecting TCa at more curable stages.

In essence, zero new evidence was found that examined either the harms or benefits of screening, yet a ‘D’ grade was offered, suggesting that the evidence was fair enough to warrant a conclusive decision to not support TCa screening (further detail on this is offered in the next section). This logic is obviously flawed. Although the primary outcomes used to build a case against TCa screening are the lack of evidence showcasing the effect of screening upon lowering TCa mortality and its potential to increase anxiety stemming from false-positives, there is no evidence measuring either of these variables to even speculate on any conclusion, let alone offering one.

At this point, the lack of evidence would suggest that there is no sound foundation upon which a recommendation can be based. Despite this, the reaffirmation update cited the three studies in establishing their recommendation. To assume that lack of evidence of effectiveness of TCa screening equates to the evidence of absence of effectiveness, according to these authors, is dangerous. CPGs and recommendations are only as good as the evidence on which they are based and insufficient evidence often results in weak recommendations (Francis 2013).

2.3. ‘Fair’ and ‘sufficient’ evidence

The USPSTF (2011) stated that fair evidence exists demonstrating the ineffectiveness of TCa screening and that the harms outweigh the benefits of performing the behavior. This statement, as mentioned previously, was made despite their own admission on the inadequacy of available evidence to conclusively determine if TCa screening is beneficial or harmful.

Calonge (2005) indicates that the USPSTF’s definition of ‘fair evidence’ is:

“Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes.” (p. 2069–2070)

An issue arises, however, as to the use of the word sufficient. There is a sense of ambiguity with the use of that word pertaining to the threshold needed to make a recommendation. Some questions that arise are: Is the USPSTF’s operational definition of sufficiency similar to other professional bodies in the field? At what point was the evidence insufficient and at what point was the evidence sufficient to make the decision? In other words, what are the criteria for insufficiency? The USPSTF refrains from providing any clarity on the threshold for deeming something sufficient, therefore casting doubt on the objectivity of the term. The Task Force, in summary, suggests the existence of sufficient evidence, which leads them to conclude that fair evidence exists on TCa screening’s ineffectiveness, despite their admission that there is an absence of evidence that can lead them to determine if it is harmful or beneficial.

Highlighting the above flawed rationale further, in their 2009 update on methods regarding insufficient evidence, the Task Force recognized limitations with using such language. The statement ‘evidence is insufficient’ was characterized as ‘useless’ or ‘worse than useless’ by clinicians and professional societies who expressed the need for better guidance in recommendations (Petitti et al., 2009). Yet, the definition of this term remains vague at best.

These authors posit that if pro-TSE advocates enact this clause of sufficient evidence, they would emerge as having the superior argument. The historical demonstration of TSE’s benefit (see Mormon, 2000; Ward et al., 2005; Mc Gilligan et al., 2009; Wanzier et al., 2014, Thornton, 2016, among others) compared to the Task Force and others, who, again, have yet to produce a solid piece of evidence sufficiently demonstrating TSE’s harms, is essentially a lopsided debate. The available evidence is exclusively rooted in demonstrating TSE’s benefit as no evidence exists on its harms.
2.4. Outcomes used to determine TCa Screening’s Worth: mortality, anxiety, and costs

The USPSTF (2011) states that males typically detect TCa unintentionally or through self/partner examination. Other research suggests that the best outcomes for TCa treatment stem from cases treated in earlier stages of the disease (see Saab et al., 2014; Gilligan, 2015; Ozturk et al., 2015). It would therefore be logical to assume that TSE or clinician examination promotes optimal wellness within this population. The rationale of the USPSTF (2011), however, is rooted in the notion that males should not undergo any type of physical examination to detect TCa due to the high probability of cure and prevention of death. The USPSTF (2011) essentially suggests that males should not be screened for TCa. If males are possibly diagnosed with the disease, it is very treatable and the individual will (most likely) live. This line of reasoning is troubling as it is essentially a gamble on an individual's expected rate of recovery and response to treatment. As an aside, no studies exist testing the effectiveness of TCa screening upon reducing TCa mortality, yet the Task Force continues to use this as a primary argument against the promotion of TCa screening.

The Task Force and the supporters of TCa screening's apparent harmfulness (beyond their assertion that TCa screening does not reduce mortality), including Case et al. (2011), continue to suggest that anxiety may manifest in the face of a false-positive screening. They fail, yet again, to produce any such evidence supporting their claim. These authors do not doubt that certain levels of anxiety would result from a false-positive TSE as similar trends have been historically observed with false-positive test results. For example, Fylan (1998) discusses the issue regarding abnormal smears in cervical cancer screening. Nelson (2013) discusses it from the standpoint of breast cancer screening, and Brown, Dimeu, and Cameron (2014) present a discussion on HIV self-tests and possible adverse psychosocial effects from false-positive results. However, again, a thread woven throughout the majority of such studies and statements suggesting false-positive anxiety is a reason to not perform self-exams of any kind is that little, if any at all, evidence exists to make such a claim.

A recent review (Saab et al., 2016) found that highlighting the high cure rate of TCa can actually help overcome cancer-related anxiety. This also may serve as part of the counter-argument to the USPSTF (2011) and others where teaching males about TCa and TSE to increase knowledge and awareness of the disease, preventive measures, and treatments would actually decrease anxiety (see also Rovito et al., 2015). Another commentary (Rovito, 2016) provides more insight on the illogical use of anxiety as a central point in the argument against TCa screening.

One outcome the Task Force fails to consider in its rating of TCa screening is the significant impact on the quality of life and risk of adverse mental health outcomes related to late-stage diagnosis. Long-term health implications of more invasive and toxic treatments during late-stages cases (e.g. chemotherapy, radiation, etc.), previously noted, are discussed by Gilligan (2015). The authors indicate that later-stage cases treated with chemotherapy experience a significantly worse quality of life and manifest higher frequencies of adverse mental health outcomes as compared to their counterparts diagnosed early in the staging process.

Finally, in regards to costs associated with TCa screening, aside from the typical costs associated with a doctor's appointment, performing this service is virtually free. These authors advocate that if males were adequately informed on what to screen for and how to properly palpate their testicles, the issue of unnecessary doctor's appointments to confirm a false-positive would be assuaged significantly. However, if we were to recognize the idea that males will flood clinician offices with concerns of lumps on their testicles and almost all of them being false-positives, Aberger et al. (2014) offer a comprehensive discussion on cost–benefit analysis on the topic, essentially concluding that the costs associated with the previous assertion are far outweighed by just one late-stage TCa diagnosis treatment. Although we trumpet this finding as a very important data point in defense of promoting TSE, the Task Force does not use this is their rubric to judge TSE's usefulness.

These authors caution the field on the limitations of using TCa mortality as the primary outcome to determine TCa screening’s worth as it may lose sight of survivorship and lifespan health post-diagnosis. In other words, the morbidity of the disease and treatments are not being considered in this primary outcome. This notion supports a growing body of research beginning to challenge survivorships' supremacy as the ultimate goal for determining TCa screening's benefit (see Rovito et al., 2015; Ozturk et al., 2015).

If males did not undergo the various forms of physical examination (TSE, partner, and/or physician), it logically implies that most, if not all cases (theoretically speaking, at least), would probably not be diagnosed until the disease is metastatic and produces other overt symptoms. Therefore, not recommending that TCa screening should be offered as a preventive service, whether by clinician or self-examination, is irresponsible and fails to take into account the harms of not screening, which far outweigh the potential harms associated with screening.

In summary, the USPSTF (2011) argues that since the disease is so curable that it justifies a reason not to screen. Again, morbidity of chemotherapy and major abdominal, chest, and brain surgery is missed with this argument. The Task Force also argues that because TCa does not have a high incidence, screening will lead to over treatment and investigation. The cost effectiveness argument has been investigated by Aberger et al. (2014), who indicate that evaluation is possibly cost-effective compared to treatment of one missed metastatic disease. Finally, the notion that small, early-detected testicular masses may become metastatic already at presentation (therefore offering no benefit to the patient) is a moot argument as there is no evidence to support this, and most other, claims of harm resulting from TCa screening.

3. Conclusion

3.1. In support of TCa screening

Due to the insidious, and often undetectable, nature of cancer (TCa notwithstanding), prompt identification and treatment is warranted. Self-screening is an effective means to identify certain forms of cancer given that an individual is most sensitive to their own anatomy and physiology and changes/anomalies. Given correct information, practice-based skills, and consistent application of these skills in terms of self-examination, TCa screening has the potential to further personal health in males (Rovito et al., 2015).

Often absent in the discussion of TCa screening (i.e. TSE), most notably in the USPSTF (2011) recommendation, are the other benefits of self-exams, particularly if performed correctly. For example, TSE can act to familiarize a man with his body where he becomes more sensitive in noting anomalies such as physical and structural changes (e.g. bumps, lumps, changes in consistencies) and functional issues, such as increased or decreased abilities. Moreover, familiarity also may create a sense of body competence and ‘ownership’ of one’s body possibly leading to a greater likelihood and comfort in reporting and having a dialog with one’s healthcare provider (Rovito et al., 2015). The latter might be equated to an ‘educated consumer’ model.

The public health implications of discouraging TCa screening range far beyond simple cancer detection, as previously noted. At the very least, TCa screening can improve male health literacy in terms of familiarity with his body and possibly improving patient–provider communication. The latter statement is notable as research suggests males are consistently and significantly less likely to engage in healthcare whether for preventative or palliative reasons (Addis and Mahalik, 2003; Leone and Rovito, 2013; Marcell et al., 2015).

Additionally, the age at which TSE is introduced and discussed (likely in a male’s teenage years or early 20s) can help set precedence for positive engagements with the healthcare system and one's provider
(see Dube et al., 2005, for a similar discussion on prostate cancer screening). Discouraging TCa screening via a ‘D’ grade versus an ‘I’ statement limits the practical utility of TSE or clinician examination, or any self-care initiative for that matter. Creating a positive experience through teaching males the value and skills to care for one’s body, at least in these authors’ collective opinions, carry greater value and long-term benefits than the USPSTF ‘D’ grade affords.

3.2. A call for action: changing the current TCa screening recommendation from ‘D’ to ‘I’

The main question of interest to the Task Force of whether a preventative service “works” has evolved in recent years to a more sophisticated approach that considers the level of benefit, the trade-off between harms and benefits, and individual patient and clinical preferences of a particular service (Wooll and Atkins, 2001). However, in the case of TCa screening, the methodological rigor used to produce the evidence to justify the USPSTF’s ‘D’ rating, as well as its reaffirmation, is questionable at best.

Considering the recent work on the cost-effectiveness of TSE and clinician exams, the broad support TCa screening receives in the field, and its versatility to tending to other health concerns beyond cancer, a more concerted effort to change the current rating is warranted. In other words, with insufficient evidence arguing either for or against recommending TCa screening among asymptomatic males, further research, particularly RCTs, are needed before a definitive and conclusive recommendation can be asserted.

The USPSTF has received criticism from clinicians about the frequency of ‘I’ statements in the past and the need for definitive guidance on preventive services. These concerns resulted in the development of a strategy on behalf of the USPSTF, to address the confusion created by the language in ‘I’ statements (Petitti et al., 2009). An ‘I’ statement does not provide clear guidance to practicing clinicians who do not have the luxury of waiting for more evidence to support or negate a certain recommendation and must make decisions at the point of care. In the case of TCa screening, however, a ‘D’ grade essentially limits its full appreciation and application beyond simple detection and treatment of TCa. These authors, therefore, advocate for a robust and continued discussion concerning the holistic value of TSE screening, particularly in terms of how it has a multifaceted potential value in positively impacting male health across the lifespan.

Guirguis-Blake et al. (2007) indicate that the USPSTF aims to “refine and advance its methodology” (p. 122). Therefore, it is strongly suggested that the USPSTF to reexamine the relative value of TSE screening (TSE and clinician exams) in light of the methodological and logical flaws present in the decision-making process of granting the procedure a ‘D’ grade. These authors suggest a reclassification from a ‘D’ grade to an ‘I’ statement as an appropriate next step until more evidence is offered in order to inform future evaluations and recommendations.

Conflict of interest

None.

Acknowledgments

We would like to acknowledge the Men’s Health Initiative, Inc., the Fight Like a Man International Collaborative, the Testicular Cancer Society, and their colleagues/supporters for their tireless work in male health promotion.

References


