Evaluating Intermittent Androgen-Deprivation Therapy Phase III Clinical Trials: The Devil Is in the Details

Maha Hussain, Catherine Tangen, Celestia Higano, Nicholas Vogelzang, and Ian Thompson

See accompanying editorial on page 211

A B S T R A C T

Purpose

Intermittent androgen deprivation (IAD) has been widely tested in prostate cancer. However, phase III trials testing continuous androgen deprivation (CAD) versus IAD have reached inconclusive and seemingly contradictory results. Different design and conduct issues must be critically evaluated to better interpret the results.

Patients and Methods

Seven published phase III trials were examined for prespecified design and outcomes. Treatment specifications; primary end point; superiority versus noninferiority design assumptions, including magnitude of assumed versus observed noninferiority margin (NIM); duration of follow-up; and quality-of-life (QOL) outcomes were considered in terms of the results and conclusions reported.

Results

Five trials had a superiority and three had a noninferiority primary hypothesis. Only three trials had a uniform population and overall survival (OS) end point. All trials observed better outcomes in terms of OS and progression-free survival (PFS) than assumed at time of study design, translating into prespecified NIMs or hazard ratios that reflected larger absolute differences in OS or PFS between arms. Lower-than-expected event rates also reduced statistical power for the trials. Other factors, including length of follow-up, cause of death, QOL, and primary end point, and their impact on trial interpretation are discussed.

Conclusion

No trial to date has demonstrated survival superiority of IAD compared with CAD. Trials concluding IAD is noninferior to CAD were based on wide NIMs that included clinically important survival differences, not likely to be considered comparable by physicians or patients. Interim analyses relying on short follow-up and including a majority of non-prostate cancer deaths will favor a noninferiority conclusion and should be interpreted cautiously. Adequate follow-up is required to ensure capture of prostate cancer deaths in both superiority and noninferiority trials.

J Clin Oncol 34:280-285. © 2015 by American Society of Clinical Oncology

INTRODUCTION

In 1941, Huggins demonstrated the critical role of androgens in stimulating prostate cancer (PC) growth, providing the foundation for androgen-deprivation therapy (ADT) for metastatic PC. Irrespective of modality, a majority of patients will experience progression to castration resistance, with a median survival of 4 years. Although several strategies have been tested to optimize ADT, to date, continuous androgen deprivation (CAD) is standard.

The potential to delay disease progression, improve quality of life (QOL) by minimizing adverse effects, and decrease treatment costs (medical castration) fueled international interest in testing intermittent androgen deprivation (IAD). Preclinical

data in an androgen-dependent tumor model (not PC model) demonstrating that IAD prolonged time to castration resistance by three-fold compared with CAD¹ provided the scientific rationale for testing IAD.

Several systematic reviews have concluded that IAD is as effective as CAD and therefore should be offered to patients. Nonetheless, a number of trials testing CAD versus IAD reached inconclusive and, on the surface, rather contradictory results. To properly interpret these studies in such a way to determine clinical implications, several factors must be considered. Specifically, trial design aspects, including disease stage, treatment specifications, primary end point, and design assumptions for superiority versus noninferiority for IAD versus CAD, are critical to interpretation of the results.

Maha Hussain, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI; Catherine Tangen, SWOG Statistical Center; Celestia Higano, University of Washington, Seattle, WA; Nicholas Vogelzang, US Oncology, Las Vegas, NV; and Ian Thompson, University of Texas Health Science Center, San Antonio, TX.

Published online ahead of print at www.jco.org on November 9, 2015.

M.H. and C.T. contributed equally to this work.

Authors' disclosures of potential conflicts of interest are found in the article online at www.jco.org. Author contributions are found at the end of this article.

Corresponding author: Maha Hussain, MD, FACP, FASCO, 1500 E. Medical Center Dr, 7314 Cancer Center, Ann Arbor, MI 48109-0946; e-mail: mahahuss@umich.edu.

© 2015 by American Society of Clinical Oncology

0732-183X/16/3403w-280w/\$20.00

DOI: 10.1200/JCO.2015.62.8065

In this review, we will discuss the impact of these factors on the scientific integrity and interpretation of these trials. By highlighting these considerations, we will summarize the phase III trials in the context of these factors and make recommendations for interpretation of the data.

METHODS

We included all phase III trials comparing CAD with IAD, with a primary end point of overall (OS) or progression-free survival (PFS), that were published in a peer-reviewed journal (abstracts excluded). We performed a Medline search and examined systematic reviews and meta-analysis publications²⁻⁶ focused on CAD versus IAD to ensure that all relevant trials were included. We extracted information about design, conduct, and results from the primary publication of each trial. We evaluated trials for eligibility, primary and other end points, statistical design, follow-up duration, and cause of death. We compared prespecifications for each trial and contrasted these with observed outcomes, particularly in terms of accrual and event rates. By assessing these components of the trials, we sought to provide accurate interpretation of outcomes.

We reviewed whether the included population was homogenous or mixed (metastatic and nonmetastatic) and whether the population was preselected for response to ADT. We report study specifications with regard to induction, retreatment trigger for patients in IAD arm, and drug holiday periods that could affect treatment comparisons between CAD and IAD.

We evaluated the prespecified median OS or PFS from the statistical analysis plan and compared it with the actual observed median OS or PFS and the impact that it would have on the prespecified number of events required to report final results. For trials with mixed patient populations, we checked whether the actual mix of enrolled populations matched what was projected, because this too would affect the event rate. In the case of survival, we evaluated the proportion of deaths that were related to PC or attributed to noncancer competing causes. Follow-up time was noted for each trial and patient population to ensure that patients underwent follow-up long enough to allow observation of the natural disease process and deaths resulting from PC for a given patient population. We would expect deaths resulting from causes other than PC (ie, noninformative deaths) to be balanced across the arms of a study; this would have the tendency to bring the hazard ratio (HR) comparison between the arms closer to null (HR, 1.0), thereby making it easier to declare that the two arms are noninferior. In the case of a

superiority trial, this would make it harder to declare superiority of one arm over the other. We also evaluated whether the conduct of the trial matched the prespecified design and adequately addressed the primary treatment comparison objective. We noted whether the trial had a superiority or noninferiority design and whether the sample size and number of required events matched those specified in the design. We checked that type I and II error rates were in the range typical of a phase III trial and evaluated the effect-size HR (HR for superiority design) or noninferiority margin (NIM; relative measure for noninferiority design) that was targeted. In addition to determining the NIM, we examined how the NIM translated into an absolute difference in OS or PFS between treatment arms, and we considered what the clinical significance of that difference would be.

RESULTS

We identified seven phase III trials completed and published to date.⁷⁻¹³ Only three trials had a uniform patient population and survival end point (TAP22, PR.7, and S9346 trials).^{7,10,11} In the following sections, we will describe some of the clinical and statistical design aspects of these trials. In Table 1, we identify the key trial design considerations, and in Table 2, we summarize the observed results.

Statistical Design: Superiority Versus Noninferiority Designs

Trials are typically designed with a superiority end point. Because noninferiority trials typically specify an NIM HR closer to null (1.0) than an HR that would be specified for a superiority trial (eg, HR, 1.20 v 1.30, respectively), the sample size for a noninferiority trial is typically larger than that for a superiority trial. Five trials were designed with a superiority end point, and three were designed with a noninferiority end point.

The absence of a statistically significant benefit with IAD (PFS or OS) does not imply that IAD has noninferior outcomes compared with CAD. Substantially more events (deaths or PFS events) and a long follow-up period are needed to rule out small but potentially clinically

Table 1. Trial Design Considerations for Phase III Trials Testing IAD Versus CAD

Trial (year published)	Patient Population	Preselection for Responsive Disease	Trial Design	No. of Patients or Events		Primary End Point	IAD Versus CAD		Projected Median Survival for CAD Arm
				Random Assignment Goal	Events Required for Analysis		HR, Alternative Hypothesis	NIM	
SEUG 9401 (2009) ⁸	Locally advanced or metastatic disease (approximately 70%/30%)	Yes	Superiority	570	220 (PFS)	PFS	0.70		6.0 years (PFS)
Finn Prostate VII (2012) ^{12,14}	Locally advanced or metastatic disease (approximately 50%/50%)	Yes	Superiority	600	NR	PFS	0.74		20.5 months (PFS)
SEUG 9901 (2013) ¹³	Locally advanced or metastatic disease	Yes	Noninferiority	900	658 deaths	OS	1.21		4.25 years
TULP (2011) ^{9,15}	Metastatic disease	Yes	No primary objective to compare arms		NR	PFS	NR		NR
NCIC-CTG PR.7 (2012) ¹¹	PSA relapse after RT	No	Noninferiority	1,340	800 deaths	OS	1.25		7.0 years
TAP22 (2012) ⁷	Metastatic disease	Yes	Superiority	180	88	OS	0.51		30 months
S9346 (INT-0162) (2013) ¹⁰	Metastatic disease	Yes	Noninferiority	1,512	NR	OS	1.20		35 months

Abbreviations: CAD, continuous androgen deprivation; HR, hazard ratio; IAD, intermittent androgen deprivation; NCIC-CTG, National Cancer Institute of Canada Clinical Trials Group; NIM, noninferiority margin; NR, not reported; OS, overall survival; PFS, progression-free survival; RT, radiation therapy; SEUG, South European Urological Group; TULP, Therapy Upgrading Life in Prostate Cancer.

Table 2. Results of Phase III Trials Testing IAD Versus CAD

Trial	No. Randomly Assigned	Median Follow-Up (years)	Primary Outcome	Observed Median Primary Outcome (PFS or OS; years)	IAD Versus CAD		No. of PFS Events			No. of Deaths (OS)			Deaths Related to PC (%)
					HR	95% CI	Total	IAD	CAD	Total	IAD	CAD	
SEUG 9401 ⁸	626	4.25	PFS	NR	1.23 (PFS)	0.95 to 1.59	234	127	107	339	170	169	43
Finn Prostate VII ^{12,14}	554	5.4	PFS	IAD, 2.9; CAD, 2.5	0.93 (PFS)	0.78 to 1.11	372	177	195	392	186	206	63
SEUG 9901 ¹³	918	5.5	OS	NR	0.90 (OS)	0.76 to 1.07	299	168	131	525	258	267	31
TULP ^{9,15}	193	2.6	PFS	IAD, 1.5; CAD, 2.0	NR		NR			NR			NR
NCIC-CTG PR.7 ¹¹	1,386	6.9	OS	IAD, 8.8; CAD, 9.1	0.98	0.83 to 1.16	445	202	243	524	268	256	CAD, 34; IAD, 41
TAP22 ⁷	173	3.7	OS	IAD, 3.5; CAD, 4.3	NR		137	70	67	94	49	45	NR
S9346 ¹⁰	1,535	9.8	OS	IAD, 5.1; CAD, 5.8	0.91	0.80 to 1.03	NR			928	483	445	CAD, 73; IAD, 80

Abbreviations: CAD, continuous androgen deprivation; HR, hazard ratio; IAD, intermittent androgen deprivation; NCIC-CTG, National Cancer Institute of Canada Clinical Trials Group; NR, not reported; OS, overall survival; PC, prostate cancer; PFS, progression-free survival; SEUG, South European Urological Group; TULP, Therapy Upgrading Life in Prostate –Cancer.

meaningful differences between the two arms. For example, South European Urological Group (SEUG) 9401⁸ randomly assigned 626 patients with M0 and M1 disease. The trial was designed so that the IAD arm would be considered superior if the risk of a PFS event was reduced by 30% (ie, HR, 0.70). The resulting PFS HR (IAD v CAD) was 1.23 (95% CI, 0.95 to 1.59; $P = .11$); the survival HR was 1.01 (95% CI, 0.81 to 1.25; $P = .84$). Not only did neither end point reach statistical significance for superiority of IAD over CAD, but the PFS HR estimate was in the opposite direction. The upper confidence limit for each end point suggests that we cannot rule out a 59% increase in the risk of progression or death or a 25% increase in the risk of death with IAD compared with CAD with 95% confidence. These are large relative margins that would be considered clinically inferior.

Trials of patients with nonmetastatic disease require far more participants and longer follow-up because of lower progression and cause-specific death rates. Some of the trials listed in Table 2 observed modest numbers of events and therefore had minimal power either to detect superior treatment differences or to evaluate noninferiority.

Follow-Up and Cause of Death

Median follow-up ranged from 2.6 to 9.8 years (Table 2). Trials with OS as the primary end point generally had longer follow-up than those with a PFS end point. Deaths resulting from PC were not reported in two of the trials,^{7,9} but in the largest trial, which included only men with metastatic disease (S9346) with a median follow-up of 9.8 years, a majority died as a result of PC (CAD, 73%; IAD, 80%).¹⁰ By comparison, in the one study that included nonmetastatic disease only (PR-7),¹¹ only 38% of deaths resulted from PC, and median follow-up was 6.9 years in this better-prognosis population, which is short for this disease setting. In nearly every trial, the prespecified median outcome measure for the CAD group was shorter than what was actually observed. When the observed primary outcome in the study is better than expected, either follow-up must be longer or additional patients must be included to account for the lower event rate to maintain study power.

Irrespective of design (superiority or noninferiority trial), deaths resulting from other causes have the tendency to bring the HR comparison between the two arms closer to null (HR, 1.0). In the case of a superiority trial, this would make it harder to declare superiority of one arm over the other. For noninferiority trials, noncancer events have the

effect of making it easier to declare that the two arms are noninferior. Depending on the patient population, PC-related deaths often occur later in the follow-up process (eg, patients with M0 disease), whereas earlier deaths as the trial proceeds are less likely to be PC related. Therefore, caution must be taken when interpreting planned interim analyses, because early noncancer deaths may have a significant impact on the results.

In PR.7, patients had nonmetastatic disease with rising prostate-specific antigen (PSA) after radiotherapy. Median survival was projected to be 7 years in the CAD arm. The final analysis was to be conducted when 800 deaths had occurred in the CAD arm. An interim analysis was conducted when 524 deaths (66% of expected) had occurred, of which fewer than half were related to PC. The interim analysis conclusion, at a median follow-up of 6.9 years, was that the IAD arm was not inferior (HR, 1.02; 95% CI, 0.86 to 1.21). The observed median survival of the CAD arm (only 69 patients were at risk at 10 years) was 9.1 years—two years longer than projected. As expected, final results, based on this interim analysis, included deaths occurring relatively early in follow-up, a majority of which were not PC related. It would be ideal if the follow-up for survival in PR.7 continued to better assess the impact of CAD on long-term PC deaths. This would provide useful information in evaluating CAD versus IAD in this patient population over the full survival duration.

NIMs: Relative and Absolute

Three trials (S9346, SEUG 9901, and PR.7) were survival-based noninferiority trials,^{10,11,13} a design that tests whether a new treatment is not unacceptably worse than an existing treatment. One key design challenge is selection of a margin that defines how much worse the experimental arm can perform and still be considered clinically acceptable as noninferior to the existing treatment (ie, NIM). The NIM is specified at study design and must be clinically acceptable as having potentially comparable survival while taking into account potential tradeoffs of improved QOL with the comparator treatment. S9346 specified 1.20 as the upper limit of the NIM after extensive discussions among the SWOG genitourinary community. This NIM thus allows a 6-month difference in median survival between arms based on an estimated 35-month median survival with CAD at time of study design.

The SEUG 9901¹³ study, which included a mixed population, selected a NIM of 1.21, translating into a median survival difference of 8.9 months between arms based on a prespecified median OS of 4.25 years in

the CAD arm. The prespecified NIM in PR.7¹¹ was 1.25, with a pre-specified median OS in the CAD arm of 7.0 years, translating into a median survival in patients receiving IAD that could be 1.4 years shorter than that in patients receiving CAD and still be considered noninferior. However, the observed median OS in the CAD arm was 9.2 years, so the NIM of 1.25 translated into a difference in OS between arms of 1.8 years, which would be considered noninferior. It is doubtful that either clinicians or patients would consider a survival difference of 1.8 years as comparable. If PR.7 used a NIM of 1.2, the interim results would not have led to the early reporting of the study and the conclusion of noninferiority of IAD. In contrast, if the S9346 had chosen a 1.25 NIM, the IAD OS results would have been reported as noninferior. The observed median OS for CAD in S9346 was 5.8 years, so the more conservative NIM of 1.20 translated into an absolute difference of roughly 1 year (instead of 6 months as designed); this difference is too large to be considered clinically acceptable. As pointed out by Burotto et al,¹⁶ rarely are NIMs justified in the study design of a trial. A summary of NIMs, prespecified medians, and observed medians and their impact on absolute differences in survival is listed in Table 3.

End Point: PFS Versus All-Cause Mortality (OS)

OS was the primary end point in three of the trials,^{7,11,13} PFS was the primary end point in three trials,^{8,9,12} and OS and QOL were coprimary end points in one trial.¹⁰ The definition of progression varied among trials, but all had one thing in common: Disease progression during a drug holiday in the IAD arm was not counted as progression, and the patient was allowed to be re-treated. Progression was only counted during an active treatment period. This definition artificially lengthens the PFS interval for patients in the IAD arm. For that reason, PFS is not an optimal unbiased end point in this setting, because ascertainment is biased in favor of the IAD arm. PC-specific mortality would be a better end point but would also be challenging because of cause-of-death ascertainment in an elderly population, with other competing causes of death resulting from comorbidities. OS is the optimal end point, although it requires longer follow-up and/or a larger sample size; it is also affected by subsequent treatments. Of the three end points (OS, PFS, and PC-specific mortality), OS would be the optimal and least biased end point for evaluating the real differences between IAD versus CAD.

Patient Population: Uniform Versus Mixed

Four trials^{7,9-11} had uniform patient populations (metastatic, three; PSA relapse, one), whereas three trials targeted mixed patient

populations including locally advanced and metastatic disease.^{8,12,13} Patients with locally advanced disease (M0) have a considerably lower risk of progression and death resulting from PC compared with those with metastatic disease (M1); indeed, there may be opposite relative effects of IAD in these groups, seriously confounding study outcomes. If a greater proportion of patients with M0 disease were enrolled onto a study than was specified in the trial design, the assessment of the primary hypothesis would be compromised because of fewer events. Therefore, a one-size-fits-all approach in one study is not biologically or clinically justified. SEUG 9901 had a mixed population of locally advanced and metastatic disease and expected 35% of randomly assigned patients to have M1 disease; however, in contrast, when the study completed accrual, only approximately 11% of the patients had M1 disease.¹³

Heterogeneity of Treatment Specifications

Table 2 lists whether a trial included a run-in period with CAD to ensure preselection for hormonally responsive disease. If patients with less hormonally responsive disease (ie, likely to develop early castration resistance) are included in a trial and randomly assigned to IAD, their treatment and outcomes will likely closely resemble those of patients receiving CAD, because they would likely stop receiving therapy for a relatively short period. The lower the PSA threshold for re-treatment, the more treatment a patient in the IAD arm will receive and the closer the IAD arm will mimic the CAD arm. Although we did not provide the details from each trial, summary data of percent of time receiving androgen deprivation during the study are important descriptors for evaluating the actual IAD received. This factor in turn depends on the patient population being studied and disease aggressiveness. Currently, there is no consensus on the optimal timing for restarting ADT in the intermittent setting.

The time off therapy for IAD groups varied considerably between studies and often decreased over a patient's duration on study (eg, patients receiving IAD in TULP [Therapy Upgrading Life in Prostate Cancer] study spent mean of 65% of time off therapy in cycle one, but only 14% of time after cycle three⁹). Some patient subsets may gain little benefit from IAD because they require re-treatment and fewer drug holidays compared with patients with better prognoses, so they are essentially receiving CAD despite being randomly assigned to the IAD arm. Unfortunately, none of these trials were powered to definitely identify patient subgroups more or less likely to benefit from IAD. However, results from subset analyses should be interpreted with caution.

Table 3. Corresponding Absolute Difference in Survival for Specified NIM and Median OS Measures for CAD

Trial	Relative Prespecified NIM	Prespecified		Observed	
		Median OS in CAD Arm (years)	Absolute Difference Between Arms Corresponding to NIM (months)*	Median OS in CAD Arm (years)	Absolute Difference Between Arms Corresponding to NIM (months)*
S9346	1.20	3	6	5.8	12
PR.7	1.25	7	16.8	9.1	22
SEUG 9901	1.21	4.25	8.9	5.8†	12

Abbreviations: CAD, continuous androgen deprivation; NIM, noninferiority margin; OS, overall survival; SEUG, South European Urological Group.

*Calculated as: median CAD – (median CAD/NIM), where respective observed or prespecified median CAD measure is used. Maximum difference between arms that would still be considered noninferior based on NIM.

†Extrapolated from Fig 1 (Kaplan-Meier curve of OS for CAD) from primary publication.¹³

QOL Assessment

Noninferior survival must be balanced by meaningful gains in QOL with IAD. Although the primary focus of this article is to expand on considerations for evaluating PFS and OS objectives of phase III trials, a few comments about QOL are warranted. QOL comparisons between CAD and IAD arms are challenging, because objectively, QOL is expected to be a function of testosterone recovery after cessation of luteinizing hormone–releasing hormone agonist. The term recovery itself is subjective, and it is not clear whether recovery to baseline testosterone or to any level over castrate level is needed. Recovery in turn is dependent on baseline testosterone and possibly other factors (eg, age, comorbidities). Typically, a fixed time point is prespecified to compare QOL between arms. Sometimes this can be a change score from baseline. However, each patient has his own on/off treatment pattern in the IAD arm, which possibly reflects the time to testosterone recovery. The prespecified analysis time point needs to allow for enough time to elapse postbaseline for any androgen-deprivation effect to diminish in the IAD arm or to be established in the CAD arm. For example, if QOL is measured later after patient enrollment, it will be more likely that patients in the IAD arm will be receiving active treatment with androgen deprivation when the comparison is made, dampening any QOL improvements with IAD. Selecting a fixed landmark at which to compare arms can provide some information about QOL, but this would not be patient specific, providing aggregate data. Clinical experience has highlighted significant variability among patients in the course of IAD, some with longer off-treatment periods and some with shorter periods, with disease response and rate of androgen recovery likely contributing to these differences. To better understand these differences, obtaining more frequent QOL measures is necessary, a step not pursued by any study we identified.

For the most part, QOL benefits observed in these trials were modest at best; as a result, the anticipated tradeoff between survival and gains in QOL has not been as impactful as initially thought. For the purposes of this article, we are interpreting OS and PFS without any additional weighting for QOL gains.

DISCUSSION

This review highlights many issues in conducting and evaluating phase III trials, specifically those comparing IAD versus CAD, and underscores numerous concerns related to design and interpretation of trials conducted to date. These observations can be applied in the design of future trials.

IAD has been a heavily investigated concept based on preclinical data suggesting better anticancer effect and a clinical rationale that was initially attractive. However, phase III trials have used different criteria; the trials we examined had different study populations (mixed or not), variable induction approaches, different treatment reinitiation criteria, distinct definitions of end points, different outcomes (empirically chosen, not survival), no routine testosterone evaluations, and lacked data on therapy postprogression. Because of this substantial heterogeneity across trials, meta-analytic approaches that attempt to summarize results across the studies should be interpreted with great caution.

It is worth noting that no randomized trial has demonstrated that IAD results in superior or truly comparable survival compared with CAD, although noninferiority has been declared based on questionable NIMs. In fact, survival in two of the M1 trials (S9346 and TAP22) was worse for the IAD arm compared with the CAD arm.^{7,10}

Interestingly, many of the recent reviews concluded that IAD is as effective as CAD in PC^{2-6,17}; however, these reviews often overlooked important issues, such as NIM selection including clinically meaningful differences in survival, follow-up time too short to observe disease-specific end points, suboptimal design of mixed patient populations, and small underpowered trials.

In contrast to the potential favorable effect of IAD on QOL suggested in phase II trials, the totality of the data from randomized trials does not support a large or durable effect on QOL, and in some cases, a paradoxical effect was observed. Only two of the phase III trials found potential benefit in some aspects of QOL. Patients receiving IAD in the Finn Prostate Study VII reported better QOL in terms of activity limitation, physical capacity, and sexual function, but sexuality surprisingly favored CAD.^{12,14} The authors attributed the discrepancy between sexual function and sexuality to a low response rate for this part of the questionnaire. In S9346, there were no significant differences in number of treatment-related grade 3 to 4 adverse events, including cardiovascular events, between IAD and CAD; however, the IAD group reported better erectile function and mental health. These benefits occurred 3 months after random assignment, earlier than the anticipated testosterone recovery, and did not persist beyond. Patients' awareness of treatment assignments and the emotional response to a treatment break with this and other trials may have affected QOL assessments. Measures of QOL related to function did not differ between groups in the PR.7 trial, but differences related to symptoms favored the IAD group.¹¹ In summary, in the phase III setting, benefits of IAD regarding QOL were not what they were perceived to be, and overall, they were modest at best. That, coupled with the open-label nature of all the trials, raises questions regarding the effects of IAD on QOL. Recently, Resnick¹⁸ pointed out that incremental changes in QOL, although statistically significant, may not be meaningful to patients. He recommended use of a method to determine the minimal important difference in QOL results that should guide clinical decision making.

Considering the marginal effect of IAD on QOL, the critical question is: Are the NIMs that were tested acceptable, or should we require a stricter margin (closer to 1.0) to avoid compromising survival outcomes? If S9346 had specified a NIM of 1.15, based on the observed median OS of CAD, an additional 1,000 patients would have needed to be randomly assigned to keep the same type I and II error rates. There often need to be concessions between what is scientifically appropriate and what is feasible.

The decision to recommend IAD or CAD to patients must be put in perspective, because physicians increasingly use new drugs for castration-resistant PC based on OS improvements that are far smaller than those defined by the NIMs specified in the SEUG 9901, PR.7, and S9346 trials. Specifically, docetaxel, sipuleucel-T, cabazitaxel, abiraterone, enzalutamide, and radium-223 were approved based on a median OS advantage ranging from 2.4 to 4.8 months. We therefore urge caution in interpreting the results of noninferiority trials, as previously reviewed¹⁹; there is the distinct possibility that such adoption could have far-reaching consequences for the outcomes of treatment of metastatic PC.

In medicine, therapeutic value is a balance between risks and benefits; the lack of statistically significant inferiority of IAD does not equal clinically insignificant differences; a median OS difference of 8 and 10 months in favor of CAD in patients with metastatic disease in S9346 and TAP22, respectively, cannot be ignored. This, coupled with a modest effect of IAD on QOL, and the accumulating biologic data affirming the importance of androgen receptor (AR) signaling in

disease progression and therapy—specifically, progression to castration resistance is in part adaptive to an androgen-deprived environment, with activation of AR through different mechanisms (eg, gene amplification or mutation, extragonadal androgens) and effect of AR inhibition, validated by OS improvements in metastatic castration-resistant PC^{21,22}—provides the foundation for why CAD continues to be the standard of care for patients with M1 PC. Patients interested in IAD should be carefully counseled regarding the pros and cons of this approach. For patients with PSA-only relapse, no data to date suggest that early ADT before metastatic disease affects survival²³; therefore, if ADT is chosen, a balanced discussion is needed regarding IAD.

The principles of study design, analysis, and reporting reviewed herein are important lessons learned in the context of studying IAD versus CAD. Considering that the overarching objective is to improve OS, superiority trial designs will be both more efficient and more desirable. The accumulating biologic and clinical evidence, coupled with the impressive data on ADT plus docetaxel,^{24,25} further highlights the need to focus on impactful treatments, rendering the issue of IAD versus CAD not as relevant; hence, more investment in further testing is not warranted. Future trials should be tailored to disease context,

focus on significantly affecting survival, and capitalize on an expanding portfolio of agents that have shown significant impact in metastatic castration-resistant PC.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: All authors

Collection and assembly of data: Maha Hussain, Catherine Tangen, Celestia Higano

Data analysis and interpretation: Maha Hussain, Catherine Tangen, Celestia Higano, Ian Thompson

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

- Akakura K, Bruchovsky N, Goldenberg SL, et al: Effects of intermittent androgen suppression on androgen-dependent tumors. Apoptosis and serum prostate-specific antigen. *Cancer* 71:2782-2790, 1993
- Abrahamsson PA: Potential benefits of intermittent androgen suppression therapy in the treatment of prostate cancer: A systematic review of the literature. *Eur Urol* 57:49-59, 2010
- Buchan NC, Goldenberg SL: Intermittent androgen suppression for prostate cancer. *Nat Rev Urol* 7:552-560, 2010
- Klotz L, Toren P: Androgen deprivation therapy in advanced prostate cancer: Is intermittent therapy the new standard of care? *Curr Oncol* 19:S13-S21, 2012 (suppl)
- Niraula S, Le LW, Tannock IF: Treatment of prostate cancer with intermittent versus continuous androgen deprivation: A systematic review of randomized trials. *J Clin Oncol* 31:2029-2036, 2013
- Sciarra A, Abrahamsson PA, Brausi M, et al: Intermittent androgen-deprivation therapy in prostate cancer: A critical review focused on phase 3 trials. *Eur Urol* 64:722-730, 2013
- Mottet N, Van Damme J, Loulidi S, et al: Intermittent hormonal therapy in the treatment of metastatic prostate cancer: A randomized trial. *BJU Int* 110:1262-1269, 2012
- Calais da Silva FE, Bono AV, Whelan P, et al: Intermittent androgen deprivation for locally advanced and metastatic prostate cancer: Results from a randomised phase 3 study of the South European Urological Group. *Eur Urol* 55:1269-1277, 2009
- Langenhuijsen JF, Badhauser D, Schaaf B, et al: Continuous vs. intermittent androgen deprivation therapy for metastatic prostate cancer. *Urol Oncol* 31:549-556, 2013
- Hussain M, Tangen CM, Berry DL, et al: Intermittent versus continuous androgen deprivation in prostate cancer. *N Engl J Med* 368:1314-1325, 2013
- Crook JM, O'Callaghan CJ, Duncan G, et al: Intermittent androgen suppression for rising PSA level after radiotherapy. *N Engl J Med* 367:895-903, 2012
- Salonen AJ, Taari K, Ala-Opas M, et al: The Finn Prostate Study VII: Intermittent versus continuous androgen deprivation in patients with advanced prostate cancer. *J Urol* 187:2074-2081, 2012
- Silva FC, Silva FM, Gonçalves F, et al: Locally advanced and metastatic prostate cancer treated with intermittent androgen monotherapy or maximal androgen blockade: Results from a randomised phase 3 study by the South European Urological Group. *Eur Urol* 66:232-239, 2013
- Salonen AJ, Taari K, Ala-Opas M, et al: Advanced prostate cancer treated with intermittent or continuous androgen deprivation in the randomised Finn Prostate Study VII: Quality of life and adverse effects. *Eur Urol* 63:111-120, 2013
- Langenhuijsen JF, Schasfoort EMC, Heathcote P, et al: Intermittent androgen suppression in patients with advanced prostate cancer: An update of the TULP survival data. *Eur Urol* 7:205, 2008 (suppl; abstr 538)
- Burotto M, Prasad V, Fojo T: Non-inferiority trials: Why oncologists must remain wary. *Lancet Oncol* 16:364-366, 2015
- Gruca D, Bacher P, Tunn U: Safety and tolerability of intermittent androgen deprivation therapy: A literature review. *Int J Urol* 19:614-625, 2012
- Resnick MJ: Urological cancer: Walking the tightrope of survival and quality of life with ADT. *Nat Rev Clin Oncol* 10:307-308, 2013
- Piaggio G, Elbourne DR, Altman DG, et al: Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT statement. *JAMA* 295:1152-1160, 2006
- Debes JD, Tindall DJ: Mechanisms of androgen-refractory prostate cancer. *N Engl J Med* 351:1488-1490, 2004
- de Bono JS, Logothetis CJ, Molina A, et al: Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med* 364:1995-2005, 2011
- Scher HI, Fizazi K, Saad F, et al: Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med* 367:1187-1197, 2012
- Garcia-Albeniz X, Chan JM, Paciorette AT: Immediate versus deferred initiation of androgen deprivation therapy in prostate cancer patients with PSA-only relapse. *J Clin Oncol* 32:323s, 2014 (suppl 15; abstr 5003)
- James N, Sydes MR, Mason MD, et al: Docetaxel and/or zoledronic acid for hormone-naïve prostate cancer: First overall survival results from STAMPEDE (NCT00268476). *J Clin Oncol* 33:269s, 2015 (suppl; abstr 5001)
- Sweeney C, Chen Y, Carducci M, et al: Increased survival with chemohormonal therapy in metastatic hormone sensitive prostate cancer. *N Engl J Med* 373:737-746, 2015

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Evaluating Intermittent Androgen-Deprivation Therapy Phase III Clinical Trials: The Devil Is in the Details

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Maha Hussain

Consulting or Advisory Role: Johnson & Johnson, Synthon

Biopharmaceuticals, Essa, AstraZeneca

Research Funding: Genentech (Inst), Astellas Pharma (Inst), Medivation (Inst), Pfizer (Inst), Bayer (Inst)

Patents, Royalties, Other Intellectual Property: Title: Systems and methods for tissue imaging, 3676 our file, serial No. UM-14437/US-1/PRO 60/923, 385 UM-14437/US-2/ORD 12/101, 753US 8, 185, 186 (US patent No.); Systems and methods for tissue imaging (issued patent) EP 08745653.9 (EP application No.); Systems and methods for tissue imaging (pending) CA 2683805 (Canadian application No.); Systems and methods for tissue imaging (pending) US 13/362, 500 (US application No.); Systems and methods for tissue imaging (continuation application of US 8, 185, 186); Title: Method of treating cancer; docket No., serial No. 224990/10-016P2/311733 61/481/671; application filed on May 2, 2011; Title: Dual inhibition of MET and VEGF for the treatment of castration resistant prostate cancer and osteoblastic bone metastases; applicant/proprietor: Exelixis; application No./patent No. 11764665.4-1464; application No./patent No. 11764656.2-1464; application filed on September 26, 2011

Catherine Tangen

Consulting or Advisory Role: Amgen Data Safety Monitoring Committee, Eli Lilly Data Safety Monitoring Committee

Celestia Higano

Employment: Cell Therapeutics Biopharma (I)

Leadership: Cell Therapeutics Biopharma (I)

Stock or Other Ownership: Cell Therapeutics Biopharma (I)

Consulting or Advisory Role: Dendreon, Bayer, Medivation, Ferring, AbbVie, Genentech, Pfizer, BHR Pharma, Orion Corporation, sanofi-aventis, Astellas Pharma, Sotio, Emergent BioSolutions

Research Funding: Algeta/Bayer (Inst), Aragon (Inst), AstraZeneca (Inst), Dendreon (Inst), Genentech (Inst), Medivation (Inst), Millennium

Pharmaceuticals (Inst), sanofi-aventis (Inst), Emergent BioSolutions (Inst)

Travel, Accommodations, Expenses: Bayer, Medivation, Dendreon, Pfizer, AbbVie, Genentech, Amgen, Orion Pharma, sanofi-aventis, Astellas Pharma, Sotio, BHR Pharma, Emergent BioSolutions

Nicholas Vogelzang

Employment: US Oncology

Stock or Other Ownership: Caris Life Sciences

Honoraria: DAVA Oncology, UpToDate, Bavarian Nordic, Endocyte, Dendreon, Pfizer

Consulting or Advisory Role: Janssen Biotech, Amgen, BIND Biosciences, Bayer, Novartis, Genentech/Roche

Speakers' Bureau: Medivation, Dendreon, Bayer, Caris MPI, Millennium/Takeda, GlaxoSmithKline, Novartis

Research Funding: PAREXEL International (Inst), Progenics (Inst), Exelixis (Inst), US Oncology (Inst), Viamet Pharmaceuticals (Inst), Endocyte (Inst), GlaxoSmithKline (Inst)

Travel, Accommodations, Expenses: Genentech/Roche, US Oncology, Dendreon, Novartis, Pfizer, Bayer/Onyx, Exelixis

Ian Thompson

Consulting or Advisory Role: Exosome Diagnostics, Mag Force

Patents, Royalties, Other Intellectual Property: Involved in establishment of new company—NanoTX Therapeutics—to commercialize novel therapy for glioblastoma for our cancer center; on board of directors; it has intellectual property developed by our cancer center; have several patents with colleagues involving novel biomarkers for cancer and two devices for sexual dysfunction and urinary incontinence; no revenues at this time; our university intellectual property office is working with industry to determine if these can be commercialized