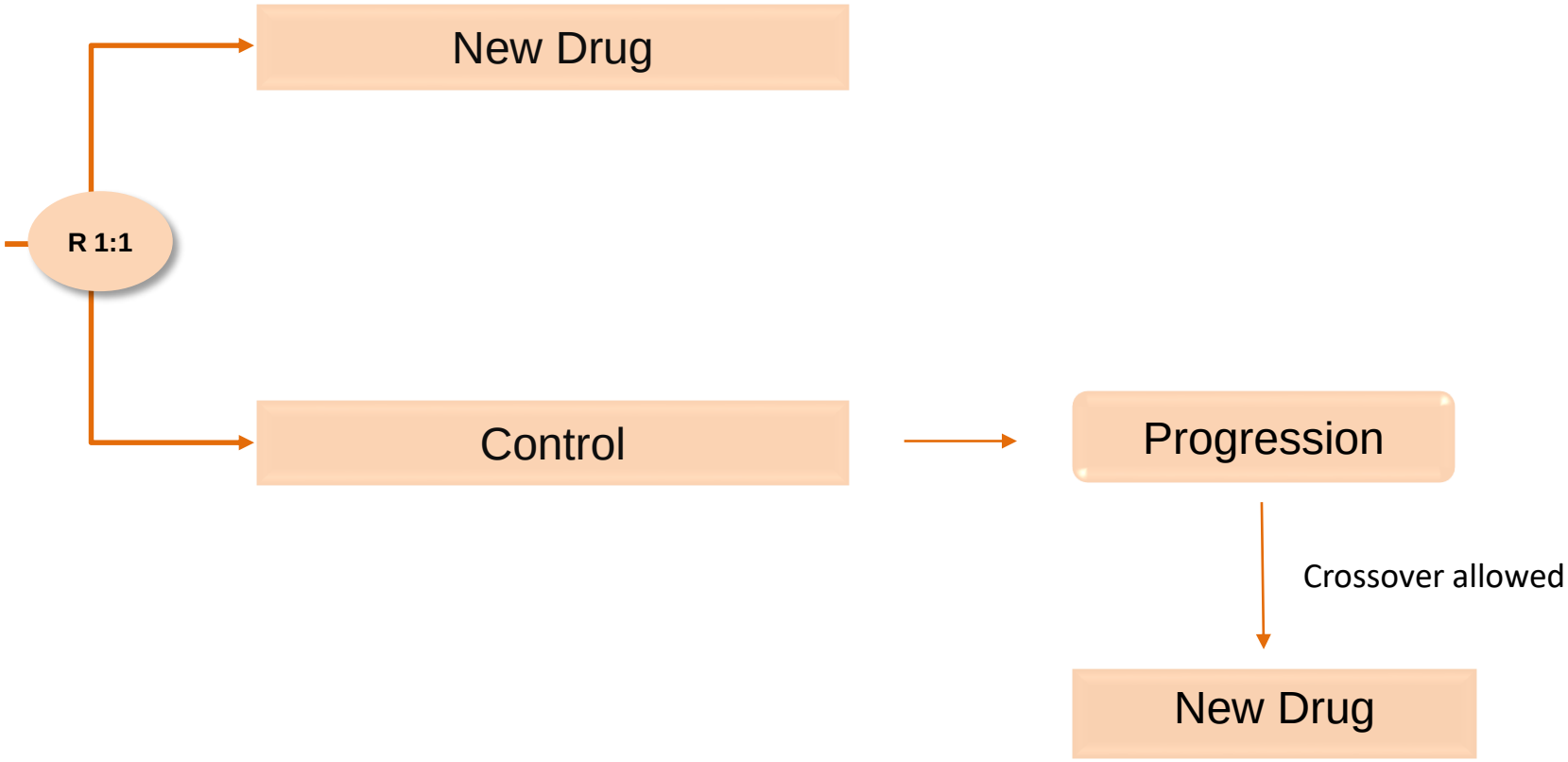
Impact of IRC Review on PFS Analysis in Randomized Oncology Studies

Tao Zhang, Ph.D
Hengrui Medicine

Contents

- Introduction to IRC
- IRC assessed PFS vs. investigator assessed PFS
- Impact of IRC assessment on PFS conclusion

Introduction to IRC



Primary endpoints: IRC assessed PFS

Secondary endpoints: investigator assessed PFS

Introduction to IRC

Independent Review Committee (IRC)

- Independent central review of imaging data, such as computed tomography (CT), magnetic resonance imaging (MRI), and conventional radiography.

IRC aims to reduce variability and bias caused by

- Variable Image acquisition protocols across sites
- Subjective assessment of images (therefore IRC is particularly valuable in open-label studies)
- Different interpretation of data

IRC review process

- A double read by two independent radiologists
- A third radiologist acts as adjudicator to resolve differences of interpretation between the two readers.
- There may also be a requirement for a separate review of clinical and laboratory data by an independent oncologist.

A central review allows an auditable, rigorous, and uniform process of evaluation. This provides greater consistency across sites.

Introduction to IRC

抗肿瘤药临床试验影像终点程序标准技术指导原则（征求意见稿）- 国家药品监督管理局，2020年4月

拟支持注册的关键研究，存在以下情况时，建议使用独立影像评估：
⑴ 单臂设计的试验（当前须采用BICR评估）；
⑵ 无法设盲、可能存在评估偏倚的试验；
⑶ 有效性统计假设/预期获益可能不十分显著的随机对照试验；
⑷ 影像源数据质量易出现偏差，需要设置IRC对影像源数据质量进行控制的试验；
⑸ 使用特殊的评估标准的试验，如需要特殊影像量化方法：少数罕见病如神经母细胞瘤间位腺代苄胍扫描(meta-iodobenzylguanidine, MIGB)或PET，或借助特殊软件处理影像数据，该标准或软件操作在实施时的一致性较难控制。

若能够充分证实影像评估数据在常规临床诊疗环境中有很好的一致性和可重复性，或在双盲随机对照的大型III期设计中，试验组疗效显著优于对照组，这种情况下独立影像评估也不是必要的。

Guidance for Industry: Clinical Trial Imaging Endpoint Process Standards - FDA, April 2018

In open-label clinical trials, availability of clinical information might influence a site-based image interpretation because the expected relation of clinical features to outcome is known, and therefore, a site-based image interpretation could raise concern about potential bias. A centralized image interpretation process, fully blinded, may greatly enhance the credibility of image assessments and better ensure consistency of image assessments

Contents

- Introduction to IRC
- IRC assessed PFS vs. investigator assessed PFS
- Impact of IRC assessment on PFS conclusion

IRC assessed PFS and investigator assessed PFS

- Analysis of primary endpoint based on adjudicated events, however. investigator assessed data are also analyzed as secondary analysis

- It is important to assess and understand concordance between IRC and investigator assessment.

抗肿瘤药临床试验影像终点程序标准技术指导原则（征求意见稿）- 国家药品监督管理局，2020年4月

IRC评估数据的稳定性是保证评估结果可信的重要指标之一，包括同一个临床试验在评估实施初期和结束时，例如按每1/3病例或事件数分析的不同阶段数据的稳定性；IRC评估者之间评估结果的稳定性，同一个评估者前后评估结果的稳定性，以及IRC的总体评估结果与研究者的总体评估结果之间差异的稳定性，包括试验组和对照组之间差异的方向和幅度是否稳定。评估的独立性不过度强调IRC与研究者在最终评估结果的一致性。

抗肿瘤药物临床试验统计学设计指导原则（征求意见稿）- 国家药品监督管理局，2020年7月

对于拟进行注册申报的试验，主要PFS分析通常应基于盲态独立中心审查委员会（BICR）评估的肿瘤测量和缓解评估，但应尽量减少研究者和BICR之间评估的不一致（如对各中心研究者进行适当的培训和教育）。

研究者和BICR对疾病进展评估的差异性是PFS分析中的重要问题。

A 2 X 2 Table

表 研究者评估PFS与独立读片结果不一致比对情况

	总计（N=xxx） 独立影像学评估		
研究者评估	事件	删失	不一致率（ % ）
事件	xx（xx）	xx（xx）	xx（xx）
删失	xx（xx）	xx（xx）	

Amit Method

早期不一致率
(a3+b)/(a+b)

晚期不一致率
(a2+c)/(a2+a3+b+c)

a1 = 均判定为PFS事件,判定事件的时间也一致.
a2 = 均判定为PFS事件,但研究者评估的事件发生较晚些.
a3 = 均判定为PFS事件,但研究者评估的事件发生较早些.
b: 研究者评估判定为PFS事件,而独立影像学评估判定为不是PFS事件.
c: 独立影像学评估判定为PFS事件,而研究者评估判定为不是PFS事件.
d: 独立影像学评估和研究者评估均判定不是PFS事件.
a = a1+a2+a3

EDR (Early Discrepancy Rate):
- Indicates investigator declares progression early relative to the IRC
- Negative value indicates a bias in investigator favoring the treatment arm

LDR (Late Discrepancy Rate):
- Indicates investigator declares progression later than the IRC
- Positive value indicates a bias in investigator favoring the treatment arm

O. Amit et al. European Journal of Cancer. 2011.

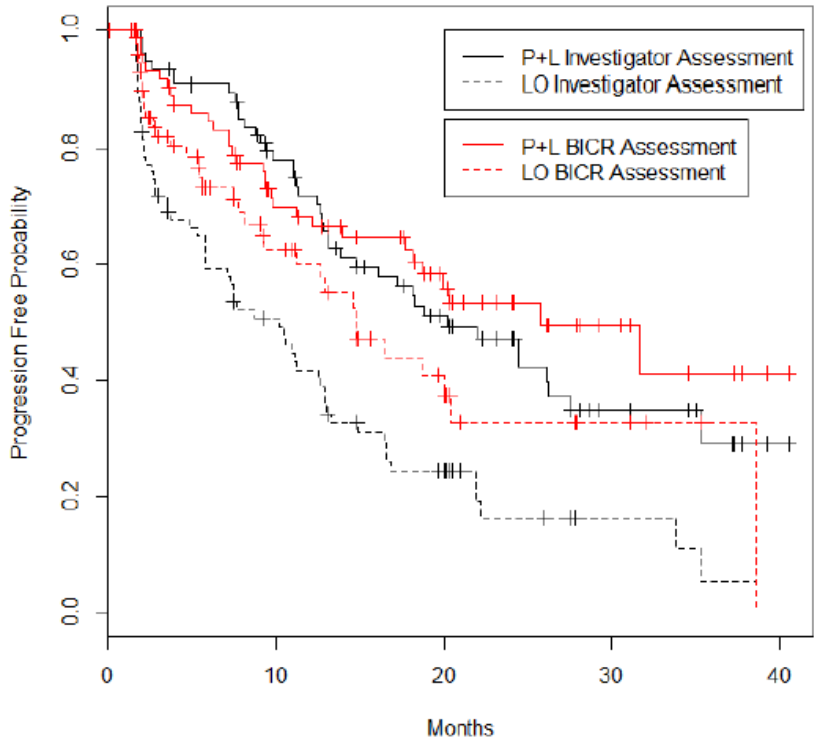
PALOMA-1： Palbociclib + Letrozole in BC

Investigator Assessed PFS

	Part 1 + Part2	
	P+L N=84	L N=81
Number of events	41 (48.8%)	59 (72.8%)
Censored	43 (51.2%)	22 (27.2%)
Median PFS (months)	20.2	10.2
95% CI	(13.8, 27.5)	(5.7,12.6)
Hazard Ratio	0.488	
95% CI	(0.319 – 0.748)	
Nominal p-value	<0.01	

BICR Assessed PFS

	Part 1 + Part2	
	P+L N=84	L N=81
Number of events	31 (36.9%)	33 (40.7%)
Censored	53 (63.1%)	48 (59.3%)
Median PFS (months)	25.7	14.8
95% CI	(17.7, NR)	(9.3, 20.4)
Hazard Ratio	0.621	
95% CI	0.378	
Nominal p-value	0.0595	



Due to the possibility of bias in an open-label study, FDA requested the sponsor to conduct a 100% BICR review.

PALOMA-1： Palbociclib + Letrozole in BC

Table 8: Early and Late Discordance Rates

	Part 1 + Part 2		Part 1 + Part 2 – Bone Only	
	P+L N=84	L N=81	P+L N=78	L N=67
EDR	46.3%	50.8%	37.1%	35.6%
Difference	-4.5%		1.6%	
Rand Test Quantile	32%		55%	
LDR	55.8%	33.3%	64.9%	48.4%
Difference	22.5%		16.5%	
Rand Test Quantile	98%		91%	

FDA: For the LDR, there does appear to be investigator bias towards the treatment group.

PALOMA-1： Palbociclib + Letrozole in BC

FDA stat review:
At this time, it appears that Palbociclib + letrozole treatment has a longer PFS time than letrozole treatment. Based upon the primary analysis of investigator-assessed PFS, the Palbociclib +letrozole has an estimated median PFS time of 20.2 months and the letrozole arm has a median PFS time of 10.2 months. Using the BICR assessment of PFS, the Palbociclib + letrozole has an estimated median PFS time of 25.7 months and the letrozole arm has a median PFS time of 14.8 months.

Nevertheless, due to poor study conduct, numerous protocol violations, data driven changes to the protocol, possible investigator bias towards the treatment arm, and a biomarker selected population in Part 2 of the study, the magnitude the difference in median PFS time remains uncertain at this time.

IMmotion 151： Atezolizumab + Bevacizumab in RCC

	IMmotion 150	IMmotion 151
Primary endpoint	PFS in ITT (IRC) ✗ PFS in PDL1+ (IRC) ✗	PFS in PDL1+ (INV) OS in ITT
Secondary endpoint	PFS in ITT (INV) ✗ PFS in PDL1+ (INV) ✓	

Phase II study supports potential for Roche's TECENTRIQ (atezolizumab) plus Avastin (bevacizumab) for people with locally advanced or metastatic renal cell carcinoma

Basel, 18 February 2017

- **Proof of concept study in first line mRCC (a type of kidney cancer) shows that TECENTRIQ and Avastin can be combined with a manageable safety profile**
- **Study results also showed encouraging efficacy compared to sunitinib in those people whose disease expressed the PD-L1 (programmed death-ligand 1) protein**
- **Roche is evaluating TECENTRIQ plus Avastin in a Phase III study (IMmotion151) in people with previously untreated, locally advanced or metastatic RCC**

IMmotion 151： Atezolizumab + Bevacizumab in RCC

Primary endpoint: investigator assessed PFS
Secondary endpoint: IRC assessed PFS

Table S1. Summary of progression-free survival

	PD-L1+*		PD-L1−†		ITT	
	Atezo + Bev (n=178)	Sunitinib (n=184)	Atezo + Bev (n=276)	Sunitinib (n=277)	Atezo + Bev (n=454)	Sunitinib (n=461)
PFS, investigator assessed						
Median PFS, mo (95% CI)	11·2 (8·9–15·0)	7·7 (6·8–9·7)	11·2 (8·6–13·7)	9·5 (8·2–10·9)	11·2 (9·6–13·3)	8·4 (7·5–9·7)
Stratified HR (95% CI)	0·74 (0·57–0·96)		0·89 (0·72–1·10)		0·83 (0·70–0·97)	
PFS, IRC assessed						
Median PFS, mo (95% CI)	8·9 (6·9–12·5)	7·2 (6·1–11·1)	11·0 (8·3–13·3)	8·4 (7·4–10·1)	9·6 (8·3–11·5)	8·3 (7·0–9·7)
Stratified HR (95% CI)	0·93 (0·72–1·21)		0·84 (0·67–1·04)		0·88 (0·74–1·04)	

Atezo=atezolizumab. Bev=bevacizumab. HR=hazard ratio. IC=tumour-infiltrating immune cell. IHC=immunohistochemistry. IRC=independent radiology committee. ITT=intent-to-treat. PD-L1=programmed death-ligand 1. PFS=progression-free survival. * PD-L1–positive tumours had a PD-L1 IC IHC expression ≥1%. † PD-L1–negative tumours had a PD-L1 IC IHC expression <1%.

Rini et al. Lancet 2019; 393: 2404–15

IMmotion 151： Atezolizumab + Bevacizumab in RCC

Table S3. Early and late discrepancy rates in the ITT and PD-L1+ population[1]

	Early Discrepancy Rates		Late Discrepancy Rates	
	Atezo + Bev (n=454)	Sunitinib (n=461)	Atezo + Bev (n=454)	Sunitinib (n=461)
ITT	0·31	0·31	0·61	0·54
PD-L1+	0·26	0·30	0·65	0·53

Atezo=atezolizumab. Bev=bevacizumab. ITT=intent-to-treat.

Negative EDR or positive LDR indicates a bias

Table S4. Reason for progression for patients with PD per IRC assessment, but not per investigator assessment, in the PD-L1+ population

Reason for IRC PD, no. of PD	Atezo + Bev (n=22)	Sunitinib (n=10)
Target lesion	8	7
Non-target lesion	3	4
New lesion	13	2

Atezo=atezolizumab. Bev=bevacizumab. ITT=intent-to-treat. PD=progressive disease.

More patients in the Atezo+Bev arm have new lesions of lymph nodes.

Rini et al. Lancet 2019; 393: 2404–15

IMmotion 151： Atezolizumab + Bevacizumab in RCC

Roche

Dr Harald Enzmann
European Medicines Agency
7 Westferry Circus
Canary Wharf
London
E14 4HB
United Kingdom

Basel, October 22nd, 2018.

Subject: Withdrawal of Type II variation EMEA/H/C/004143/II/0014 for Tecentriq (atezolizumab)

Dear Dr Enzmann,

I would like to inform you that Roche Registration GmbH has taken the decision to withdraw the type II variation application to extend the use of Tecentriq in combination with bevacizumab, for the first-line treatment of patients with unresectable locally advanced or metastatic renal cell carcinoma (RCC) whose tumours have a PD-L1 expression ≥1%.

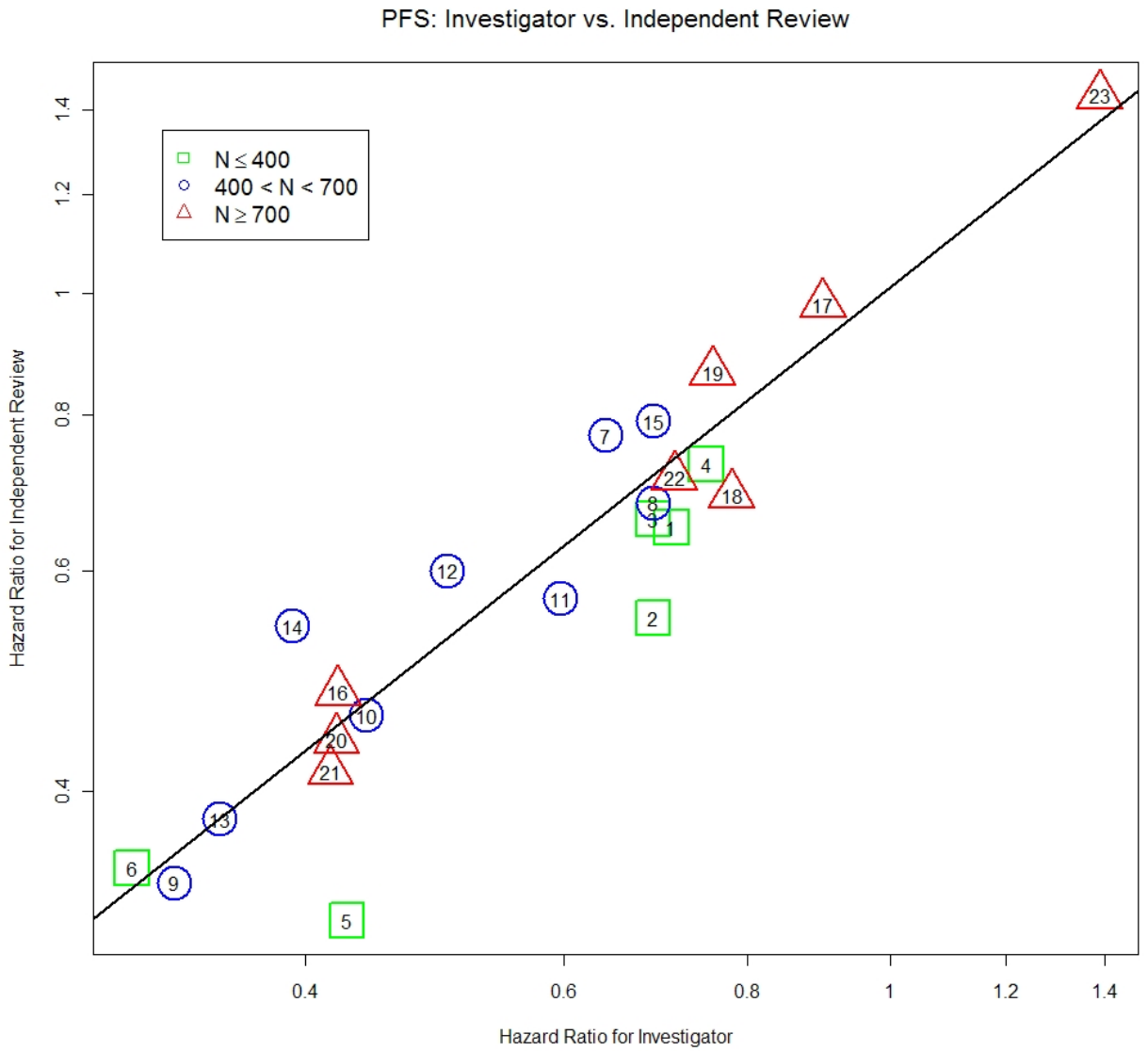
The withdrawal is based on IMmotion151 results that are not sufficient to support an extension of indication at this time. The study will continue as per protocol to the next analysis for overall survival.

A FDA Review

A FDA review (Zhang et al.)

- Evaluated all regulatory trials from 2005-2012 in solid tumors.
- Analysis included those studies with INV and IRC assessments for PFS reported Trials (20)
- Our analysis results revealed a high level of agreement between IRC and IVS assessments of PFS treatment effect. The results were also consistent across various subgroups, especially tumor type and whether the trial was blinded or open label.
- ORR results are similar.

A FDA Review


PFS: Investigator vs. Independent Review
N ≤ 400
400 < N < 700
N ≥ 700
Hazard Ratio for Independent Review
Hazard Ratio for Investigator
0.4
0.6
0.8
1
1.2
1.4

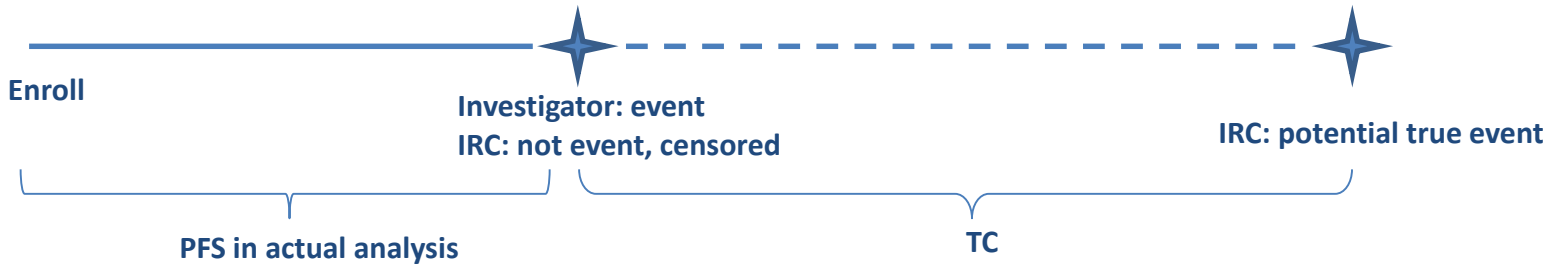
Contents

- Introduction to IRC
- IRC assessed PFS vs. investigator assessed PFS
- Impact of IRC assessment on PFS conclusion
  - ✓ Risk of censored patients: higher or lower
  - ✓ Study design: open or blinded

Impact of IRC on PFS - Simulation

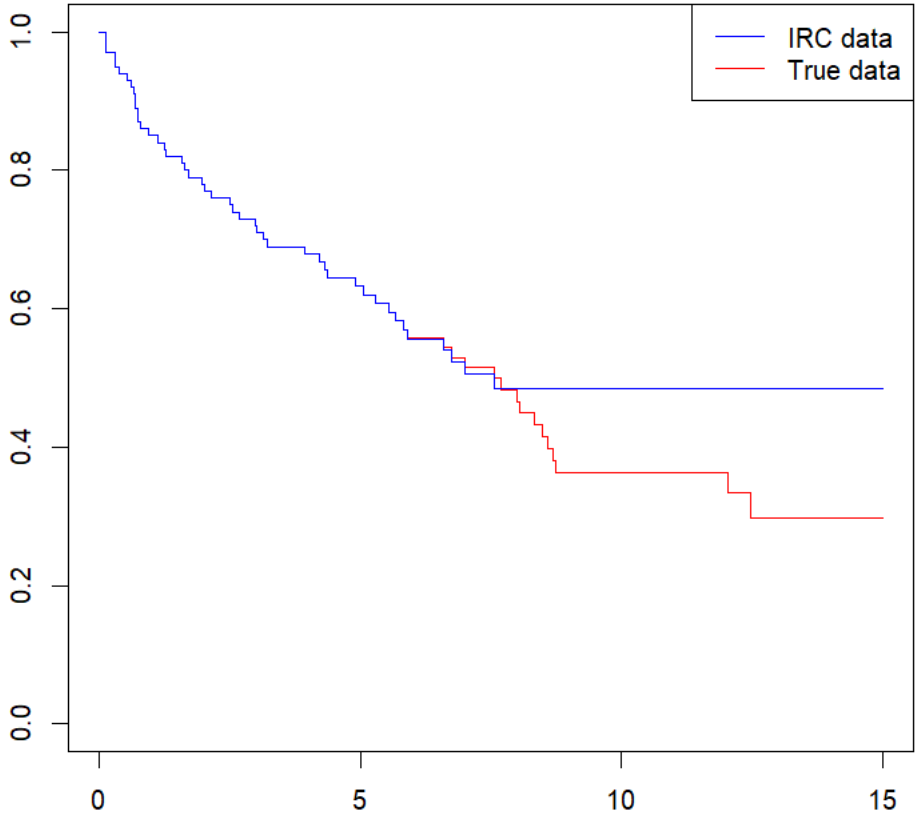
Simulation setting

- Median survival: 7 months
- Sample size: 100 patients
- Recruit duration: 12 months
- Analysis time (calendar time): 15 months
- Number of censoring due to IRC: 10 patients
- Time from IRC censor (investigator progression) to true IRC event: TC

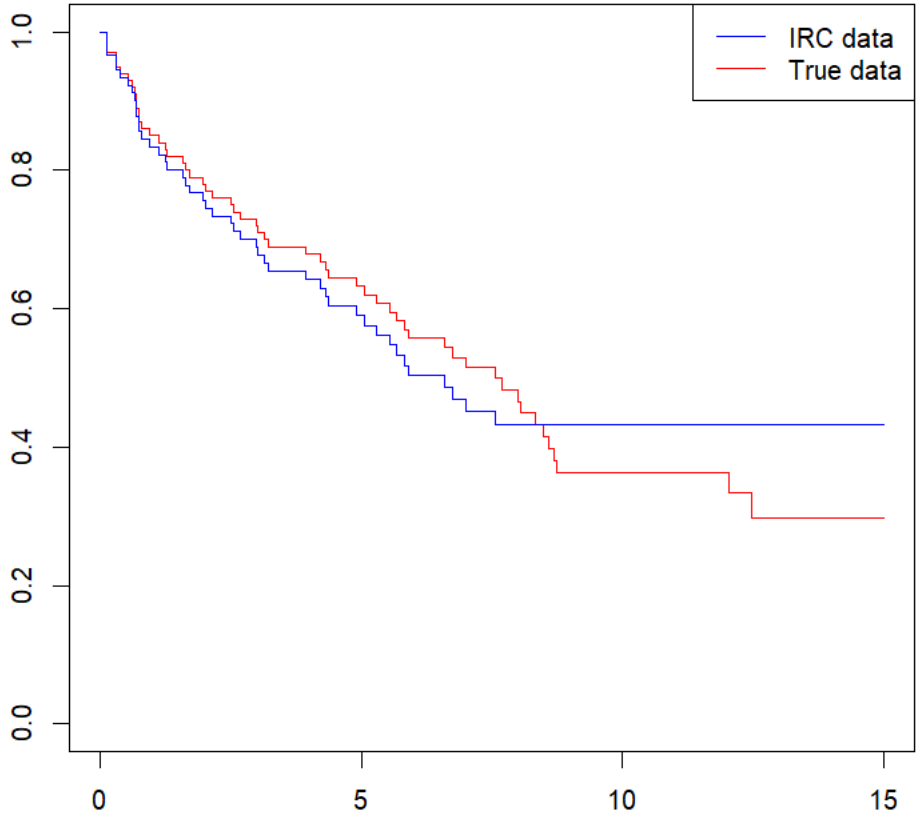


Impact of IRC on PFS - Simulation

Time from IRC censor to true IRC event:
2 months(censored patients has high risk)



12 months(censored patients has low risk)



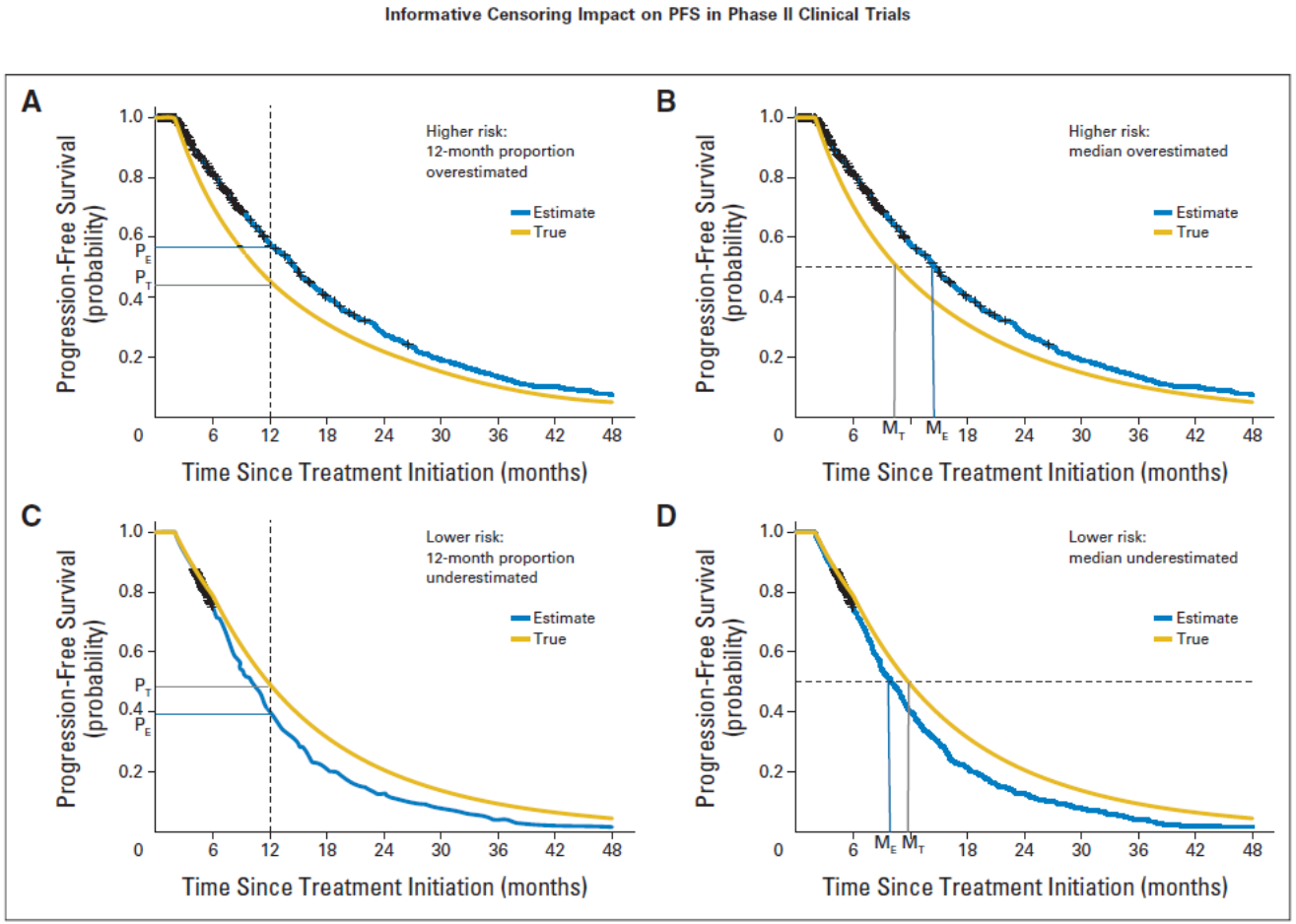


Fig 1. The anticipated direction of the bias in the 12-month and median progression-free survival (PFS) estimates is shown when patients who come off study are at a higher risk for progression (A, B) or a lower risk for progression (C, D). The yellow line represents the true but unknown PFS. The blue line represents the PFS estimate using the Kaplan-Meier method for simulated patients with Waldenström macroglobulinemia in (A) and (B) where 20% start nonprotocol therapy before progressive disease (PD; assuming hazard rate per month of PD $\lambda_1(t) = 0.06$ and $\lambda_2(t) = 4\lambda_1(t)$) and simulated patients with simulated multiple myeloma in (C) and (D) where 40% proceed to autologous stem-cell transplantation before PD (assuming $\lambda_1(t) = 0.108$; $\lambda_1(t) = 2\lambda_2(t)$). PT and MT indicate the true 12-month PFS and the true median PFS. PE and ME indicate the estimates for 12-month and the median PFS.

Campigotto et al., Impact of Informative Censoring on the Kaplan-Meier Estimate of Progression-Free Survival in Phase II Clinical Trials. Journal of Clinical Oncology. 2014.

AB: patient who come off study are at a higher risk.
CD: patient who come off study are at a lower risk.

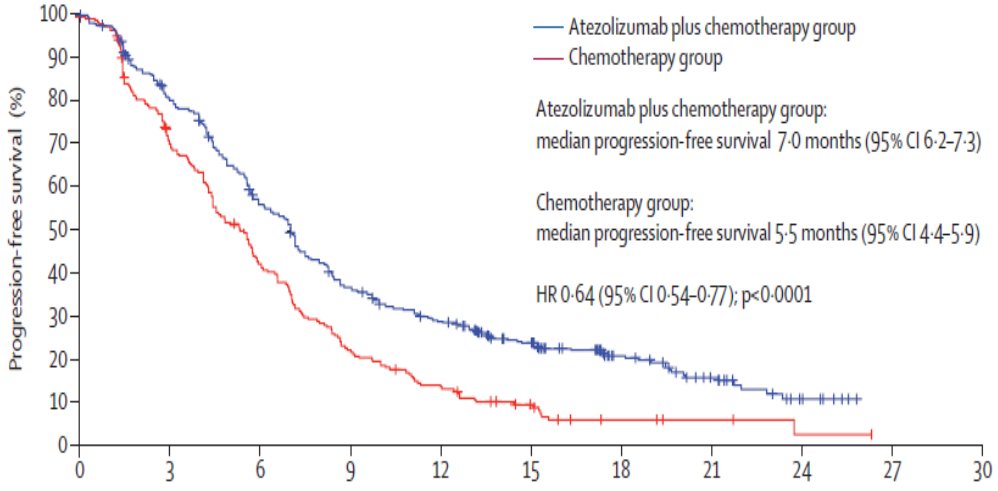
Contents

- Introduction to IRC
- IRC assessed PFS vs. investigator assessed PFS
- Impact of IRC assessment on PFS conclusion
  - ✓ Risk of censored patients: higher or lower
  - ✓ Study design: open or blinded

IMpower 130 (randomized, open-label phase III)

Local assessment

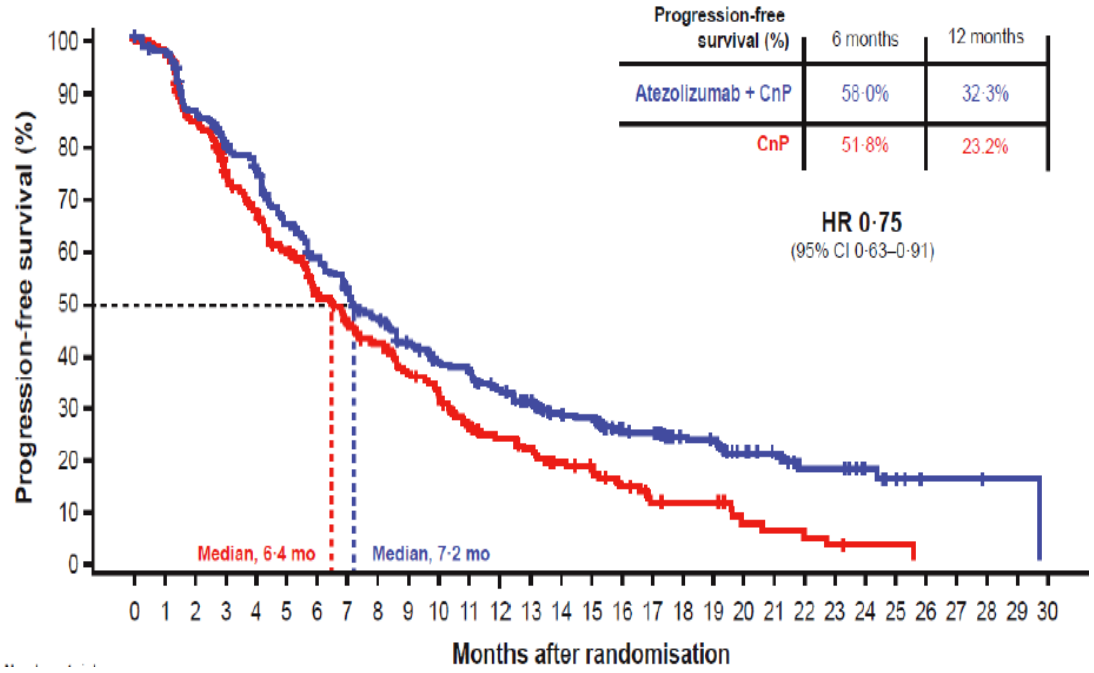
	Progression-free survival at 6 months	Progression-free survival at 12 months
Atezolizumab plus chemotherapy group	56·1% (95% CI 51·5–60·8)	29·1% (95% CI 24·8–33·4)
Chemotherapy group	42·5% (95% CI 35·9–49·1)	14·1% (95% CI 9·4–18·8)



Number at risk (number censored)

Atezolizumab plus chemotherapy group	451 (0)	351 (13)	242 (17)	157 (19)	119 (25)	78 (48)	41 (77)	23 (86)	7 (97)
Chemotherapy group	228 (0)	150 (13)	90 (14)	48 (14)	29 (15)	15 (21)	5 (26)	3 (28)	1 (29)

Central assessment

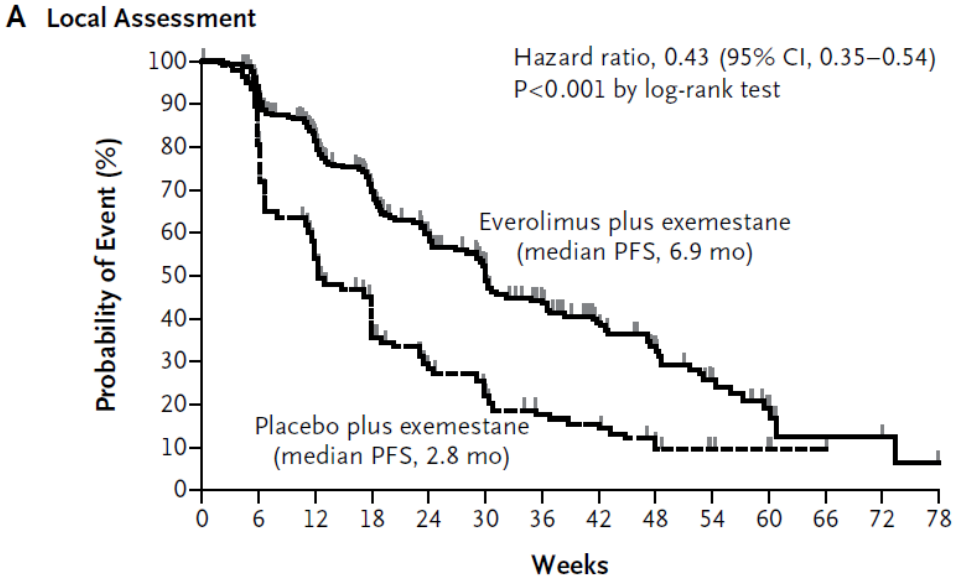


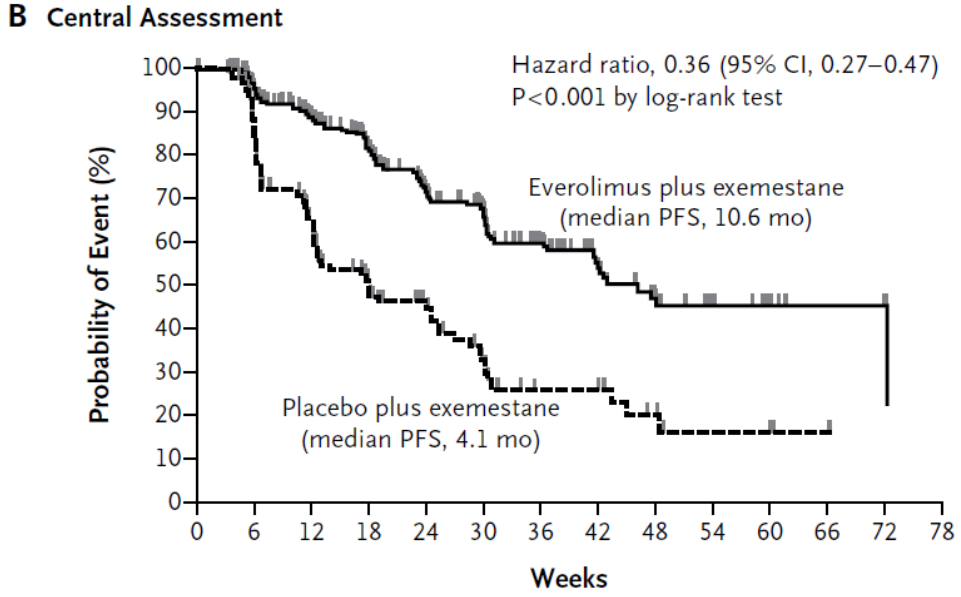
Progression-free survival (%)	6 months	12 months
Atezolizumab + CnP	58·0%	32·3%
CnP	51·8%	23·2%

- IRC assessed PFS are higher than investigator assessed PFS. This trend is more prominent in control arm.
- More patients in control arm are censored due to progression as assessed by investigator. They would progress shortly by IRC assessment.

Howard West, et al. Lancet Oncol 2019; 20: 924–37. NCT02367781

BOLERO-2 (randomized, blinded phase III)





- Blinded design
- There was a substantial difference between treatment arms for patients withdrawn from the study prior to progression because of toxicity or other reasons (24% versus 6% respectively), and some of these patients were censored.

Baselga J, Campone M, Piccart M, et al. Everolimus in postmenopausal hormone-receptor–positive advanced breast cancer. New Engl J Med 2012;366:520–9.

Conclusion & Discussion

- Concordance between IRC and investigator assessment: Amit method and simple method.
- PALOMA-1 accelerated approval; IMmotion 151 submission withdraw.
- According to a FDA review, there is a high level of concordance between IRC and IVS assessments of PFS treatment effect.
- Two important factors that can affect conclusion of PFS conclusion.
  - ✓ Risk of censored patients: higher or lower
  - ✓ Study design: open or blinded
- Some publications recommend to use time-to-treatment-failure (TTF) as an endpoint, where discontinuation of study treatment for any reason is considered an event.

(FDA: TTF is generally not recommended as a regulatory endpoint for new molecular entity drug approval.)
(CDE: 一个合理的支持审批的终点指标应当能清楚地将有效性和药物毒性、患者或医师退出、或患者不耐受区分开，TTF不能将有效性与其他变量进行充分区分。因此，不建议将TTF作为支持药物批准的终点。)

Questions?

Key Reference

O. Amit et al., Blinded independent central review of progression in cancer clinical trials: Results from a meta-analysis. European Journal of Cancer. 2011.

Zhang el al. Evaluation of Blinded Independent Central Review of Tumor Progression in Oncology Clinical Trials: A Meta-analysis. Clinical Trials. 2012.

Dodd et al., Blinded Independent Central Review of Progression-Free Survival in Phase III Clinical Trials: Important Design Element or Unnecessary Expense? Journal of Clinical Oncology. 2008.

Vinay et al., The role of censoring on progression free survival: Oncologist discretion advised. European Journal of Cancer. 2015.

Fleming et al., Issues in Using Progression-Free Survival When Evaluating Oncology Products. Journal of Clinical Oncology. 2009.

Campigotto et al., Impact of Informative Censoring on the Kaplan-Meier Estimate of Progression-Free Survival in Phase II Clinical Trials. Journal of Clinical Oncology. 2014.