Non-inferiority: Issues for today and developments for tomorrow

Kevin Carroll

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

Dr R Pazdur, Director, Division of Oncology Drug Products, CDER, FDA

- "For regular approval of a drug, the sponsor must demonstrate that the drug is safe and effective in adequate and well-controlled trials. The effectiveness must be demonstrated on an endpoint that the agency believes to represent clinical benefit, usually survival, disease symptom amelioration or established surrogates for these."

- "The sponsor is not obligated to show that the drug is safer and/or more effective than an approved drug."

Oncologic Drugs Advisory Committee, July 27, 2004: Transcript, pg 17:
http://www.fda.gov/ohrms/dockets/ac/04/transcripts/2004-4060T1.htm

The statutory requirement for regulatory approval

- The drug is effective
- There is a positive benefit:risk

- New drugs do not have to be shown to have superior efficacy vs existing agents, otherwise only one therapy could ever be approved and available at any given time
  - Approval of a new drug would logically necessitate the currently approved standard to be withdrawn
- This would result in chaos and a total loss of therapeutic options to the patient and physician

How might we show effectiveness?

- Directly, via a placebo controlled trial
  - This is, and remains, the gold standard for effectiveness
- Indirectly via an active-controlled trial
  - When a placebo controlled trial is either unethical or impractical

There really is no such thing as an 'NI' trial

- There are only active-control (AC) trials with differing objectives
- When placebo control is either unethical or impractical, effectiveness is established via an AC trial by either
  - Showing drug is better than control and, thus, drug is better than placebo
    or
  - Showing indirectly, by reference to historical data, that drug is better than placebo

The first objective of an AC 'NI' trial is not NI

- The true and first regulatory purpose of an active-controlled 'NI' trial is to establish indirectly that a new therapy would have beaten placebo if a placebo controlled trial could have been conducted.

- Examination of relative efficacy of new to control by, say, showing a given amount of the control effect, say 25%, 50% or 75%, has been retained might be considered as a further, descriptive and subsidiary objective.

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

Two approaches to 'NI' assessment in AC trials

- Preservation of control effect
  - Superiority to putative placebo not considered sufficient
  - The new drug must, in addition, preserve some fraction of active control's effect
- The fixed margin
  - The largest loss of effectiveness that could be tolerated clinically for drug relative to control
  - Denoted by δ
  - Must show that the CI for the difference between drug and control does excludes δ

Parameters for an AC trial

- Define:
  - γ_{TC} = Effect of drug vs control
    - Estimated as B_{TC} (with variance V_{TC}) in AC trial
  - γ_{CP} = Effect of control vs placebo
    - Estimated as B_{CP} (with variance V_{CP}) from historical trial(s))
  - γ_{TP} = Effect of drug vs placebo
    - Indirectly estimated from as B_{TC} (with variance V_{TC}) from B_{TC} (V_{TC}) and B_{CP} (V_{CP})

Preservation of effect

- For preservation of effect, want to show:
  - $\gamma_{TC} \geq (1\text{-}f)^{*}\gamma_{CP}$
  - Where $0 \leq f \leq 1$ is the preservation factor

- Declare 'NI' if:
  - The 95% CI for $\gamma_{TC} - (1\text{-}f)^{*}\gamma_{CP}$ excludes zero
  - $\{B_{TC} - (1\text{-}f)^{*} B_{CP}\} - 1.96^{*}\sqrt{(V_{TC} + (1\text{-}f)^{2*} V_{CP})} > 0$

Fixed margin

- For fixed margin, not uncommon to:
  - Base δ on lower end of 95% CI for γ_{CP}
  - $\delta = (1\text{-}f) * [B_{CP} - 1.96*\sqrt{(V_{CP})}]$
- Declare 'NI' if:
  - The 95% CI for effect of drug vs control excludes the margin, i.e. if
  - $B_{TC} - 1.96*\sqrt{(V_{TC})} > -\delta$
  - $\{B_{TC} - (1\text{-}f)* B_{CP}\} - 1.96*\{\sqrt{(V_{TC})} + (1\text{-}f)*\sqrt{(V_{PC})}\}] > 0$

Logical inconsistencies with preservation of effect (1)

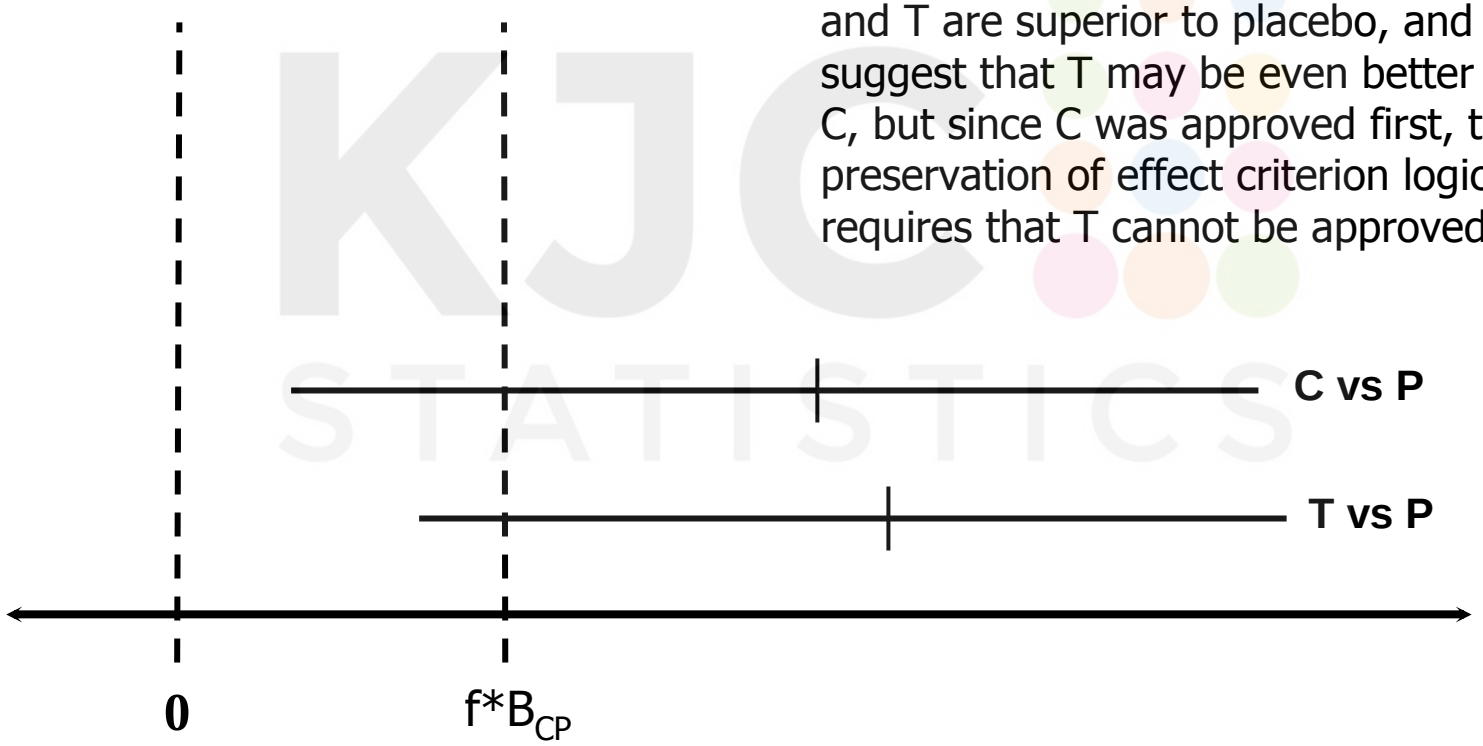
- The requirement for approval is to show effectiveness, $\gamma_{TP} > 0$
- For preservation of effect, need to show $\gamma_{TP} > f^*\gamma_{CP}$
- **Why should this requirement depend on the study design?**

Logical inconsistencies with preservation of effect (2)

- The regulatory bias inherent in applying the preservation criterion leads to logical inconsistencies, as illustrated by the following example:
  - Suppose C and T were both evaluated in placebo-controlled trials, but C was approved first
  - Assume $\gamma_{TP} > \gamma_{CP}$
  - FDA's requirement that T preserve f% of C's benefit over placebo will in many plausible instances lead to rejection of T (even though T may be a better drug than C !)

Logical inconsistencies with preservation of effect (3)

The results here indicate that both C and T are superior to placebo, and they suggest that T may be even better than C, but since C was approved first, the preservation of effect criterion logically requires that T cannot be approved.



Logical issues with the fixed margin

- Two sponsors compare their new drugs T1 and T2 against some active control C
- First sponsor specifies NI limit =1.30
- Second sponsor specifies NI limit=1.15
- Suppose upper 95% CL for T1:C is 1.25 and for T2:C is 1.20.
- Conclusion: T1 is non-inferior while T2 is not, despite being able to rule out a lesser for disadvantage for T2.

Do We Need a Non-Inferiority Margin?

- If the goal is demonstration of effectiveness:
  - Need to show $\gamma_{TP} > 0$, or $\gamma_{CP} + \gamma_{TC} > 0$
  - This is estimated by $B_{CP} + B_{TC}$, with Variance $V_{CP} + V_{TC}$
  - The **Synthesis Method** provides and efficient test for non-inferiority: $(B_{CP}+B_{TC})/\text{sqrt}(V_{CP}+V_{TC}) > 1.96$

- ***Where Does the Need for a Margin Come From?***

One Standard of Evidence

- The standard of evidence for effectiveness of a new treatment T is:
  - $\rightarrow$ Statistically significant evidence that $\gamma_{TP} > 0$
- Why should an arbitrarily higher standard of evidence ($\gamma_{TP} > y > 0$) be applied when an AC trial has been performed?
- The preservation margin is necessarily arbitrary, in the sense that there will be values below the margin for which there is no meaningful clinical difference in efficacy from a value above the margin.
  - Preserving less than f% does not imply that T is an ineffective treatment.
  - In contrast, $\gamma_{TP} = 0$ has a definite objective clinical meaning.
- Requiring a higher standard of evidence for AC trials institutes a regulatory bias in favor of the first drug to be approved.

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

PhRMA CDIG PISC Team
(Non-Inferiority / Active-Controlled Trials)

- Steven Snapinn (Amgen): Chair

- Kevin Carroll (Astra Zeneca)
- Christy Chuang-Stein (Pfizer)
- Yu-Yun Ho (Novartis)
- Qi Jiang (Amgen)
- Gang Li (J&J)
- Patrick Peterson (Lilly)
- Yong-Cheng Wang (J&J)
- Matilde Sanchez (Arenapharm)
- Rick Sax (Astra Zeneca)

Proposed Approach

- Basic Principles
  - Required degree of efficacy to support approval should be independent of study design
  - If "weaker" drug is approvable then the "Better" drug should be approvable
  - Methodologic weaknesses of NI trials should Be addressed separately
- Proposal
  - Rigorous demonstration of "Any Efficacy" ($\gamma_{TP} > 0$)
  - Magnitude of treatment effect evaluated through point estimate via the synthesis method
  - Term "Non-Inferiority" seems inappropriate – prefer "Active Control Efficacy Trial"

Common criticisms

- We're allowing for new drugs to be ineffective when we do NI trials; we want positive superiority trials
  - A superiority trial with $p<0.05$ does not rule out a new drug is ineffective
  - We accept a 1 in 40 chance that the drug is ineffective
- Constancy is the Achilles heel of NI trials
  - But this issue applies to all AC trials regardless of the objective – if the control arm 'under performs', then $p<0.05$ might not be convincing – if it over performs then we increase the risk of a false negative
- Assay sensitivity
  - Regardless of the objective, a poorly conduct AC trial raises a legitimate concern and PP and ITT analysis would always need to be close
- Methods in development to examine assay sensitivity and constancy (discounting, Snappin 2008)

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

Three stage proposal for design and analysis

Stage 1 Design

Calculate the number of deaths to establish indirectly new treatment is better than (putative) placebo with 90% power, 2.5% significance.

Sample size formulae provided by Simon (1999) or Rothmann (2003)

Proposal for design and for analysis
(contd)

Stage 2 Analysis (a)

Demonstrate absolute efficacy by indirect analysis vs placebo using historical data

Stage 3 Analysis (b)

Having passed Stage 2, then describe the relative efficacy of new to control on a continuum in terms of to what extent the new drug has retained the efficacy of the control – 'Effect Retention Likelihood'

Effect Retention Likelihood

- Both analysis stages can be simultaneously visualised in an effect retention likelihood plot
- Trial results can be expressed in a natural hierarchy, starting with the likelihood that the new drug is better than a putative placebo, followed by the likelihood that any given fraction of the control effect has been retained through to the likelihood that the new drug is better than the control

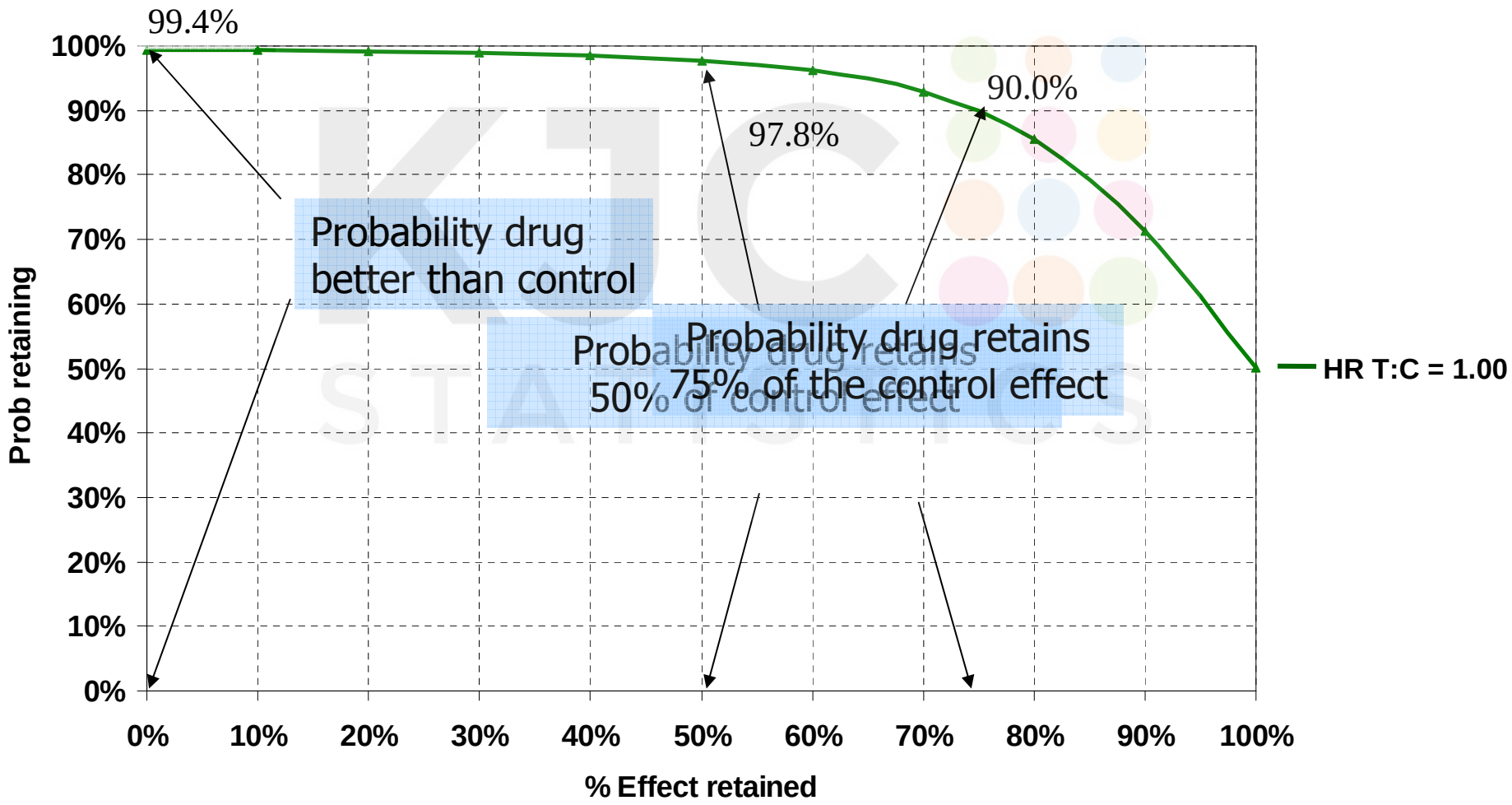
How might this proposal be classified?

- Consistent with thinking in CHMP guidance
- Bayesian in nature
  - Historical data are prior, the active–control trial is the data, and the effect retention likelihood is the posterior
- Coincides with Simon, 1999, under uninformative priors for effect of new drug and effect of placebo

A simple example

- HR[placeo:control]=1.5, p=0.005, 95% CI (1.13, 1.99)
- An AC trial with new drug vs control is planned
- To show efficacy with 90% power, 2.5% 1-sided $\alpha \Rightarrow$ 800 events, delta 1.26

Drug vs. control effect retention likelihood for trial conducted with 800 events



Setting a delta in the protocol

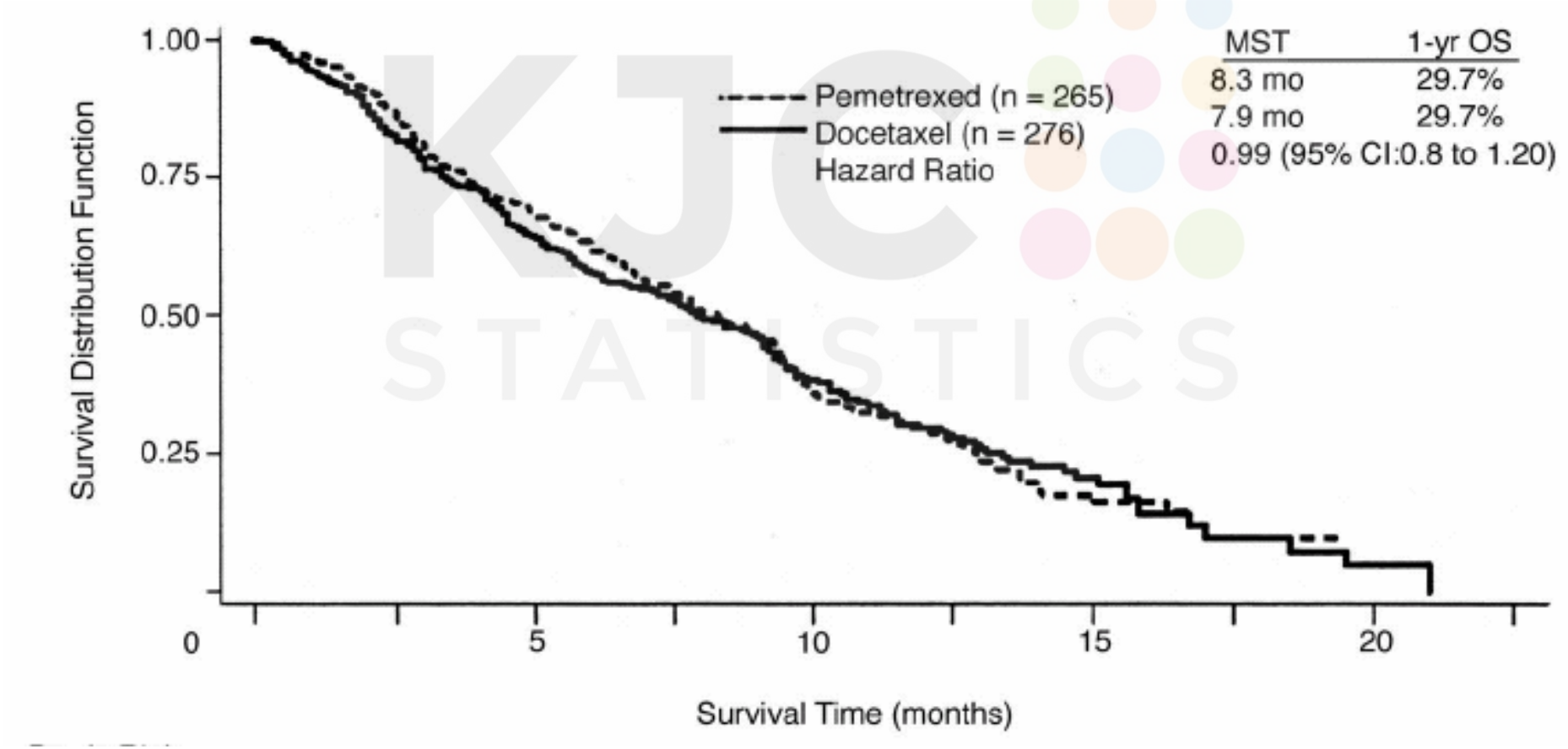
- "Analysis will take place with 800 events. With this amount of information,noting the effect of the control has previously been estimated to be 0.50, p=0.005 and accepting constancy, this trial will be able to show that New is significantly better than placebo with 90% power. In practice, this equates to observing a HR of new to control of 1.09 or less, upper 97.5% CL <1.26."

Pemetrexed[1] example

- Pemetrexed vs. docetaxel in 2nd line NSCLC, N=571
- HR=0.99 (0.82-1.20), 409 deaths
- NI limit in protocol = 1.11
  - 78%* docetaxel effect retention
- Docetaxel previously improved survival cf placebo (BSC) with HR=0.56 95% CI (0.35, 0.88)

[1] Nasser Hanna et al, J Clin Oncol 22:1589-1597, 2004. * retention required to give p=0.025 (1-sided) for non-inferiority

Randomized Phase III Trial of Pemetrexed Versus Docetaxel in Patients With Non–Small-Cell Lung Cancer Previously Treated With Chemotherapy



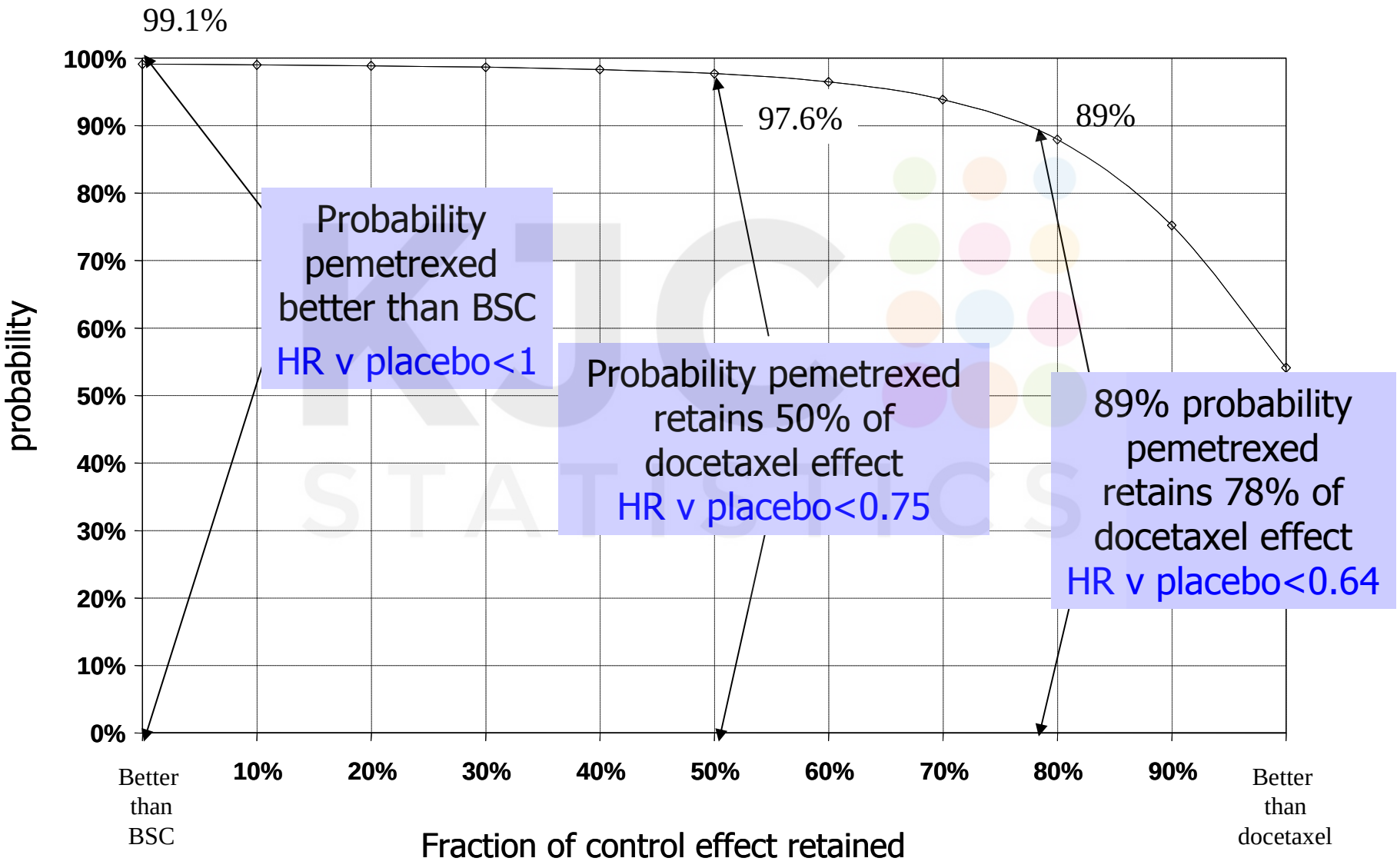
Pemetrexed not non-inferior to docetaxel (p=0.11)[1]

- Cannot say with >97.5% confidence that true HR <1.1 or pemetrexed retained >78% of docetaxel effect
- 'The primary objective in the present Alimta trial was not achieved. Neither superiority nor non-inferiority to docetaxel were adequately demonstrated.' [2]
  - Dr R Pazdur, Director Office of Oncology Drug Products, FDA

1: 1-sided test that true HR<1.1

2: http://www.fda.gov/ohrms/dockets/ac/04/transcripts/2004-4060T1.htm

Effect retention likelihood for pemetrexed



EPAR Scientific Discussion

‘Although non-inferiority was not formally established, the data submitted are robust enough to conclude that a clinically significant inferiority of pemetrexed to docetaxel in terms of efficacy in this population is unlikely’

Contents

- Showing drug effectiveness and requirements for approval
- Approaches to 'NI' assessment
- PhRMA CDIG PISC team view
- A possible mulit-stage approach to AC design and analysis
- Summary

Summary (1)

- One standard of evidence (superiority to placebo) should be maintained for the assessment of effectiveness ($\gamma_{TP} > 0$); this standard should not be determined by trial design.

- In situations where a placebo control is not possible necessitating an AC trial, requiring an arbitrary non-zero preservation of the control effect leads to serious logical inconsistencies; the requirement for approval should be convincing demonstration of effectiveness, $\gamma_{TP} > 0$.

- Fixed NI margins are equally arbitary, statistically inefficient and highly conservative and should be abandoned

Summary (2)

- Analysis should be based on the synthesis method
- The synthesis approach can test for superiority to placebo, accounting for the variability within both the AC trial and the historical data.
- Methodologic issues associated with AC trials should be addressed as far as possible in the design, but will inevitably be a key feature in evaluating and interpreting the data
- A natural extension of the synthesis approach is a retention likelihood analysis which can be used to simultaneously assess the strength of evidence for $\gamma_{TP} > 0$ and gauge the relative efficacy between drug and control right through to $\gamma_{TC} > 0$

Thank you

Any questions?