ISCTM
THE INTERNATIONAL SOCIETY FOR CNS CLINICAL TRIALS AND METHODOLOGY

DETERMINING HOW EFFECTIVE A TREATMENT WILL BE FOR AN INDIVIDUAL PATIENT

Helena Chmura Kraemer, Ph.D.
Stanford University (Emerita)
University of Pittsburgh

Background

- MacArthur Network and what followed (Skimming the top).
- Theme: The value of clinical research is measured by the degree to which it improves the impact of clinical decision- making on individual patients.
- A study that well addresses that problem is likely to be more powerful, more cost-effective, and to have a greater impact on patient care.
- Insanity: doing the same thing over and over again and expecting different results. (Albert Einstein).

Effect Sizes - I

- Meehl: All null hypotheses (of randomness) are false. True?
  - Large enough samples sizes and enough RCTs will eventually result in one or two or more $p<.05$s.
- The p-value indicates the adequacy of the design (sample size, measurement) to detect the deviation from randomness, not the clinical importance of the treatment choice.

Effect Sizes - II

- An appropriately chosen effect size: a population parameter (estimated in a sample) that indicates the clinical importance of the finding.
  - In each RCT report effect size and its confidence interval (CONSORT, APA)
  - When multiple RCTs done, meta analysis using that effect size will give a more precise estimator of the population parameter. `

Which Effect Size to compare T1 vs T2 in a RCT?

- Cohen's d: To compare normal distributions with equal varainces, yes.
- Odds Ratio: Never!
- AUC=Prob(T1>T2)+.5Prob(T1=T2)
  - AUC estimated by U/(mn), where m and n are the sample sizes.
  - Where d appropriate AUC=Φ(d/√2).
- SRD=Prob(T1>T2)-Prob(T1<T2)=2AUC-1.
  - For success/failure outcome, SRD=Success Rate Difference.
- NNT=Number Needed to Treat=1/SRD.

Definition of Moderators of Treatment - I

- The overall effect size comparing T1 and T2 in a RCT indicates the AVERAGE effect size over the patients in the population, NOT the effect on any individual patient.
- M is a moderator of treatment choice (T: T1 versus T2) on the outcome (O) ***if*** the effect size of T on O changes depending on what M is.
  - Personalized Medicine.

Definition of Moderators of Treatment - II

- MacArthur Model criteria for Moderators:
  - Temporal Precedence: M precedes T precedes O in time.
  - Correlation: M and T uncorrelated.
  - Effect Size: The effect size of T on O changes depending on what M is.
- In a RCT, any baseline (pre-randomization) variable satisfies the first two criteria.

Example: MTA search

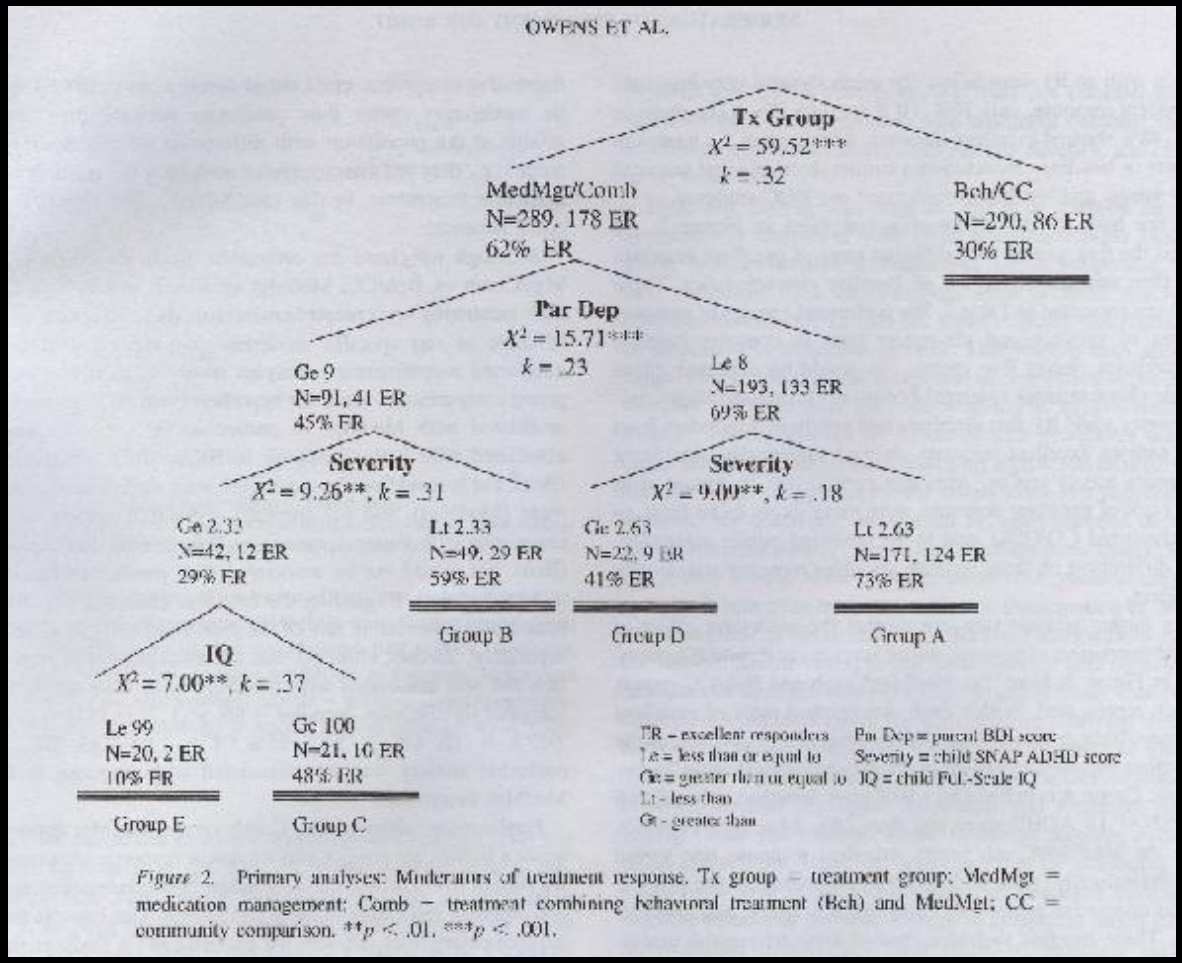


Figure 2. Primary analyses: Moderators of treatment response. Tx group = treatment group; MedMgt = medication management; Comb = treatment combining behavioral treatment (Beh) and MedMgt; CC = community comparison. $^{**}p < .01$. $^{***}p < .001$.

Messages - I

- For any T1 vs T2 comparison, there are many baseline factors, but only a few moderators of treatment on outcome: Need for exploration (hypothesis-generation) to detect moderators.
- The result of exploration is rationale and justification for **hypotheses** to be tested in future RCTs (and the information needed to design powerful such studies). It is not **conclusions.**

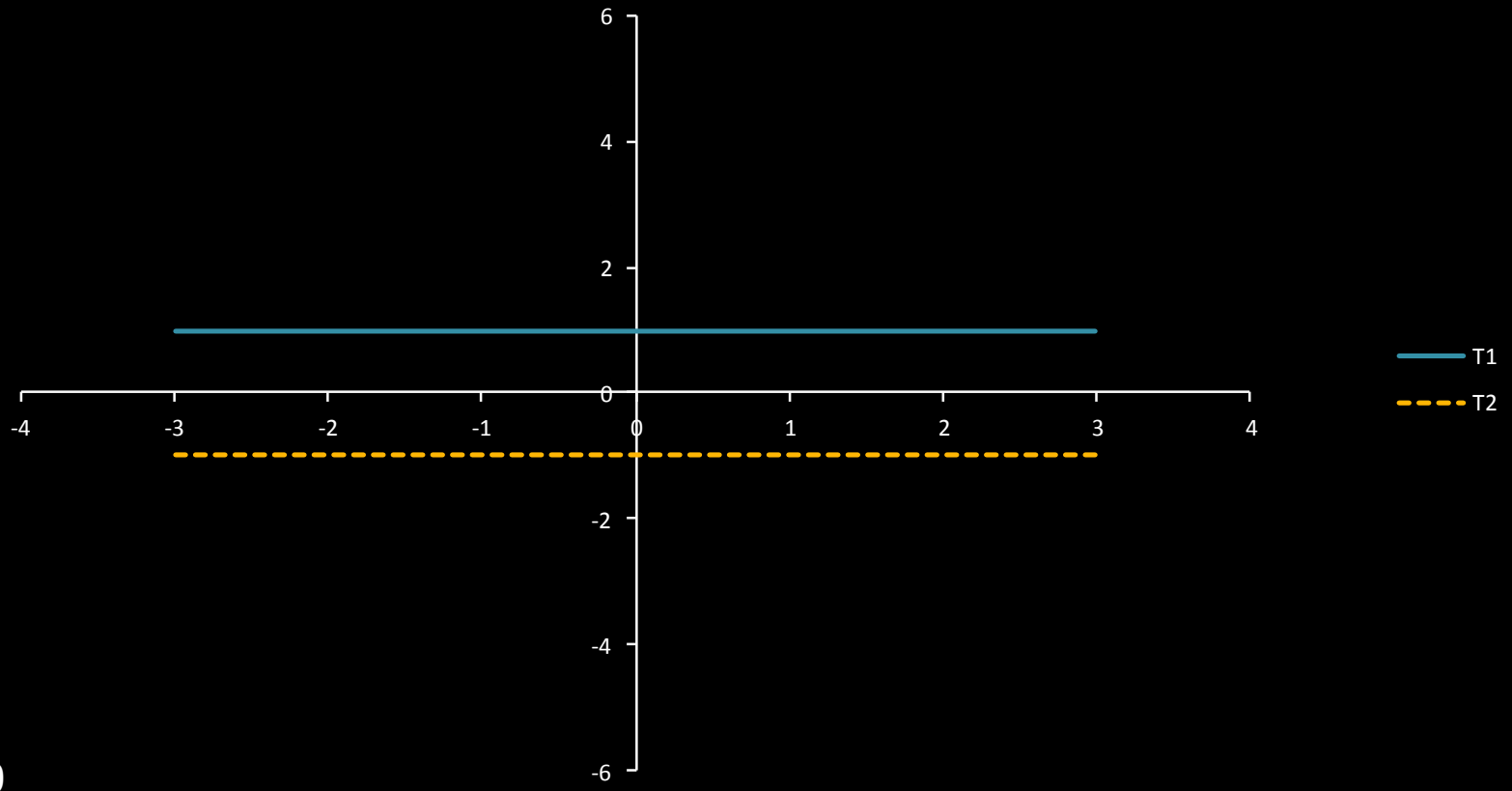
Messages - II

- What you expect will moderate, may not. Surprise!
- A single moderator may not much improve clinical decision-making; decision rules based on **multiple moderators** will likely do much better, and will bring medical decision-making closer to the goals of Personalized Medicine.

How to search for a moderator—Linear Model

- Linear Model: $O=b_0+b_1T+b_2M+b_3TM+e$
- Centering: T coded +1/2 and -1/2; M standardized to a mean of 0 and variance 1.

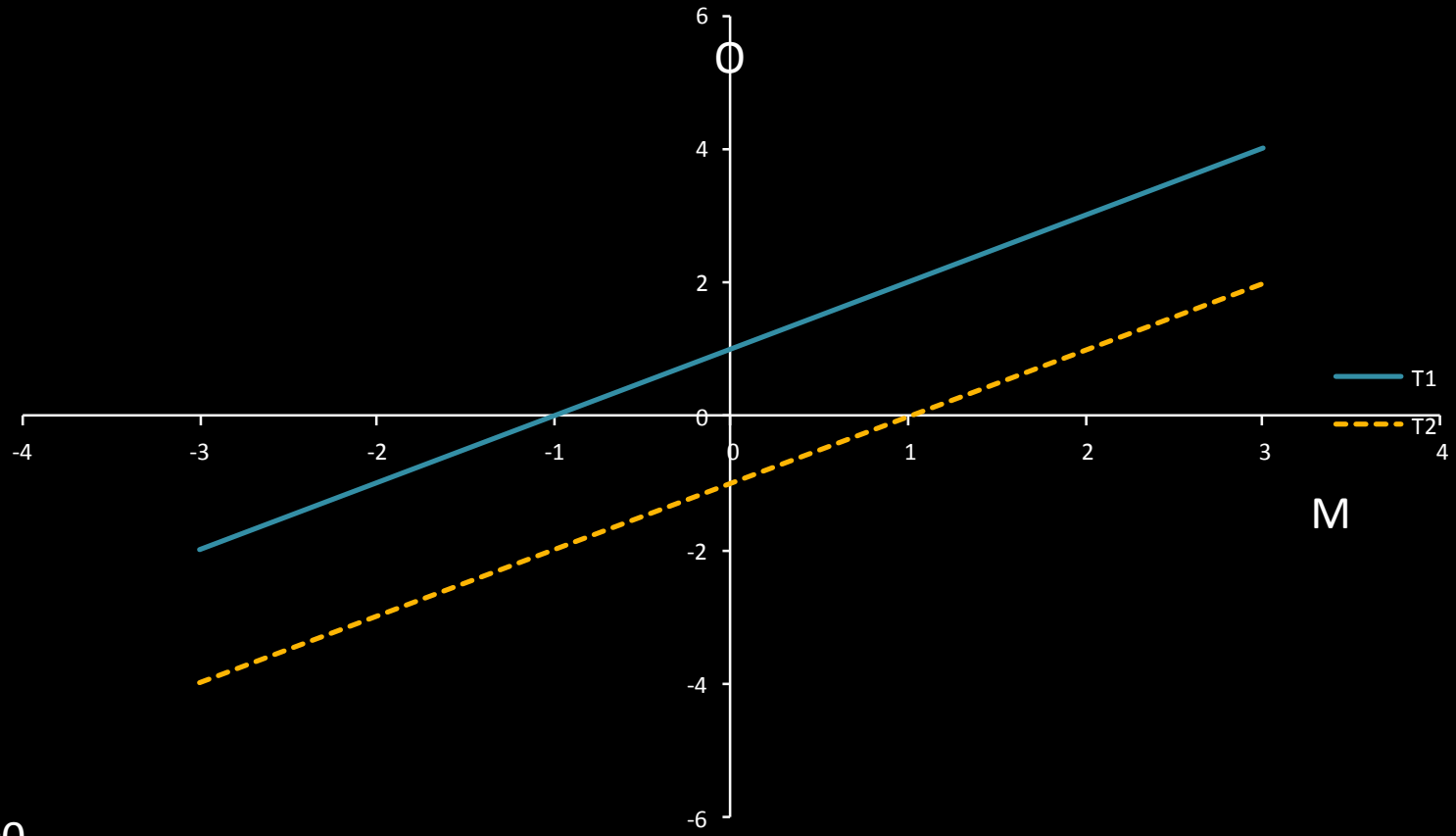
Factor is irrelevant to treatment



$b_2=b_3=0$

$d_1=b_1/\sigma_e$=Overall ES=ES(each M)

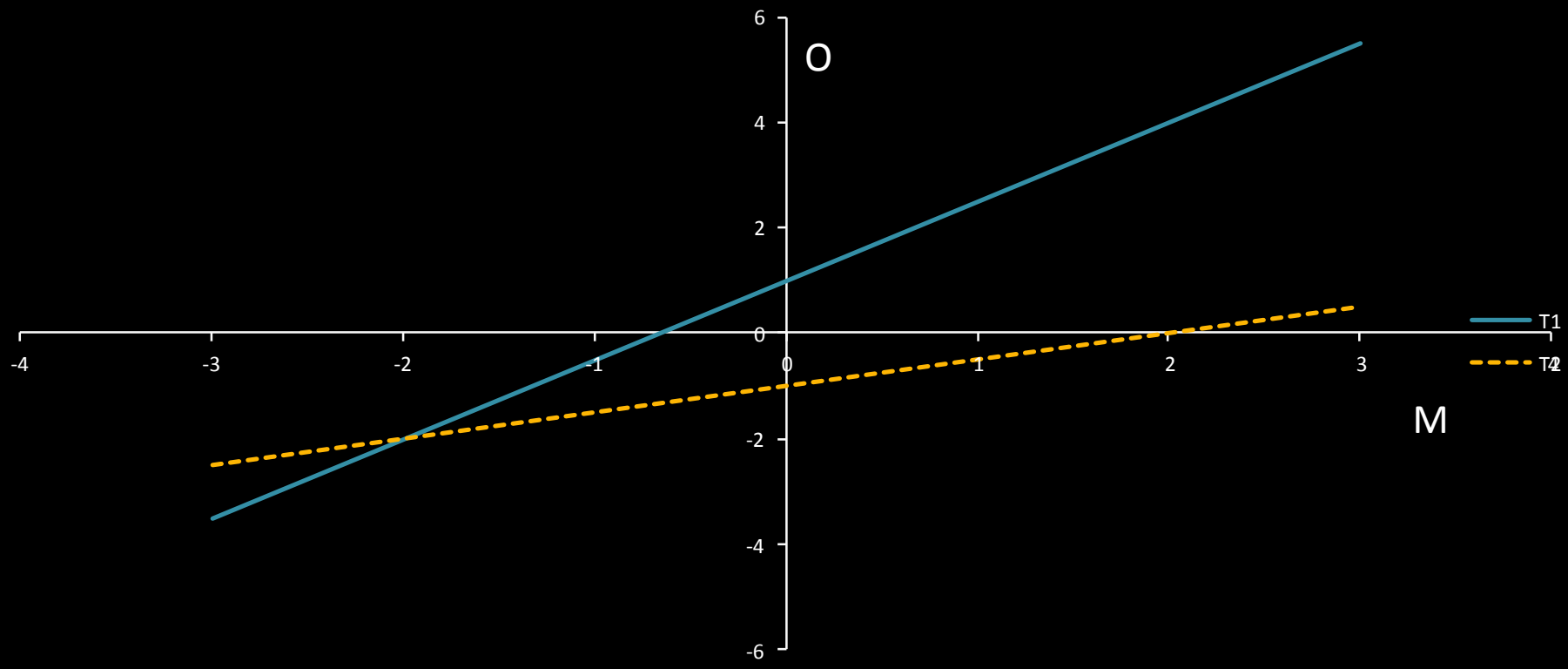
Factor is A non-specific predictor



Here $b_3=0$, $b_2\neq0$

$d_1=b_1/\sigma_e$=ES(each M)

Factor is a Moderator



Here $b_3 \neq 0$.

If T1>T2 for all M, or T1<T2 for all M, then M is a "quantitative moderator".

If T1>T2 for some M and T1<T2 for others, then M is a "qualitative moderator".

$d_1 = b_1/\sigma_e = ES(M=0)$

Moderator Effect Size - I

- We need a Moderator Effect size:
  - To compare and convey the potential impact of different moderators.
- b_3??? Not invariant under linear transformation of O.
- $d_3=b_3/\sigma_e$??? Ignores the relative contributions of M and T separately.

Moderator Effect Size - II

- The treatment effect size T1 versus T2=Overall ES:

  Overall ES=$d_1/\sqrt{(d_2^2+d_3^2/4+1)}$.

  - d_1 is the Cohen's d for subjects with M=0.
  - Attenuation of the magnitude of ES_1 due to d_2 and/or d_3.

- NspES=$d_2/\sqrt{(d_2^2+d_3^2/4+1)}$.
- ModES=$(d_3/2)/\sqrt{(d_2^2+d_3^2/4+1)}$.
- Then Overall ES=$d_1(1-\text{NspES}^2-\text{Mod ES}^2)$.

MODERATOR EFFECT SIZE - III

- Randomly select one patient from T1 and one from T2, compute ΔO, the difference in outcome, ΔM, the difference in their Ms, and AM the average of their Ms.
- Then $\Delta O = b_1 + b_2\ \Delta M + b_3\ AM + \Delta e$.
  - But ΔM and AM are uncorrelated, since the two Ms have equal variances (Pitman).
- Thus:
  - Correlation (ΔO, AM)=ModES.
  - Correlation (ΔO, ΔM)=NspES.

An Optimal moderator - I

- Have possible moderators M_1, M_2, M_3, ……
  - Remove any that don't make sense, that are badly measured.
  - Remove redundancies.
- For each M, test, and estimate the Mod ES and its confidence interval (bootstrap).
  - Compare the Mod ES of the moderators.

An Optimal moderator - II

- Use a multiple regression model with ΔO from all possible randomly paired subjects, with independent measures AM_1, AM_2, ….. as the predictors. This produces weights w_1, w_2, …..
- The optimal moderator is $M^*=\sum w_j M_j$.
  - Get its Mod ES and the confidence interval
- Get independent verification!

Conclusion

- Importance of clinically meaningful effect sizes (and their confidence intervals)
- Importance of moderators in RCTs and in clinical decision-making.
- Exploration to:
  - Identify individual moderators and their effect sizes.
  - Identify an optimal moderator and its effect size.
- Independent Validation
- Treatment decision algorithms for clinical use (apps?).
- Changing the direction of subsequent clinical research by focusing separately on moderated subgroups.

Transition

- Everything depends on having an outcome measure (O) that is reliable, valid, and that is sensitive to crucial individual differences in treatment response among the patients (harms and benefits).
- Ellen….