



U.S. Department of
Health and Human
Services



National Institutes
of Health



National Heart, Lung,
and Blood Institute

Peter G. Kaufmann, Ph.D.
National Heart, Lung, and Blood Institute

Issues in Clinical Trials:

Generalizability, Sample size, and Interpretation

Third Annual NIH/OBSSR Summer Institute
on Randomized Clinical Trials

Airlie Conference Center, Warrenton, Virginia

July 25, 2003

Overview

- Generalizability
- Sample size and stability of results
- Equipoise and trial ethics
- Mechanistic and ancillary studies
- Alternative designs
- Scientific integrity

Generalizability

- The extent to which participants in a trial represent the population with the condition under study.
 - Representativeness is limited by:
 - » Eligibility criteria
 - » Consent
 - » Other unknown factors
- Stipulate: RCT results can always be generalized to similar patients in settings in which the trial was conducted.

Generalizability

- Treatment – related questions:
 - How easily can a Tx be applied within contemporary clinical practice?
 - Is qualified staff available (or can be trained) to deliver the Tx?
 - How well suited is an intervention for application:
 - » To a large population
 - » To a specific population
 - » To a specific patient

Generalizability

- Treatment – related questions:
- What proportion of individuals with the condition will respond if the Tx is delivered in a modality that is different than the one tested?
- What will it cost to treat individuals who could benefit from the Tx?

RE-AIM

“...public health and community based...Multilevel interventions that incorporate policy, environmental, and individual components should be evaluated with measurements suited to their settings, goals, and purpose.”

Glasgow, Vogt, and Boles, Am J Public Hlth, 1999, 89,1322-1327.

RE-AIM

Reach: The percentage of persons who are affected by a policy or program.

Efficacy: Positive outcomes, negative outcomes, biological and behavioral outcomes.

Adoption: Proportion and representativeness of worksites, communities, health departments that adopt a given policy or program.

Implementation: Extent to which a program can be delivered as intended.

Maintenance: Extent to which innovations become a stable, enduring part of an individual's, organization's, or community's behavioral repertoire.

RE-AIM

“...although often efficacious for those participating, traditional face-to-face intervention modalities will have limited impact if they cannot be delivered consistently to large segments of the target population.”

Glasgow et al., Patient Educ & Counseling, 2001, 44, 119-127.

RE-AIM

“Program planners should make decisions regarding implementing and funding health services based on multiple dimensions, rather than only considering efficacy in randomized clinical trials.”

Glasgow et al., Patient Educ & Counseling, 2001, 44, 119-127.

Community, Public Health, Policy

Suitability depends on:

Prevalence of the targeted condition.

Severity of the targeted condition.

e.g., Fluoridation of water

Vaccination

Modality of the Intervention

e.g., Lifestyle messages

CONSORT

- A list of requirements for uniform reporting of clinical trials with the overall aim of improving the reporting of RCTs, to facilitate their critical appraisal, and to facilitate their inclusion in systematic reviews.

Published in 1996, revised 2001

CONSORT

- Primarily aimed at first reports of two-group parallel design RCTs.
(equivalent to Phase III trials)

CONSORT

- CONSORT standards represent only minimum reporting requirements for RCTs
- Informative reporting of RCTs has been a notoriously difficult problem
- CONSORT standards have been accepted primarily by medical journals

Summing up

Exercise judgment:

Does the rule or recommendation apply to the case under consideration –

- design and objectives of a specific RCT
(e.g., prevention vs. treatment)
- review of results from a completed RCT
- review of a new grant application

Phases of drug development

- Phase I: determine dose and toxicity
- Phase II: establish biological activity & adverse event rates
- Phase III: evaluate effectiveness in comparison to other treatments
- Phase IV: long term monitoring, with or without comparison groups
- Other: dissemination/translation research, community interventions

Phases of behavioral intervention development

- Phase I: test intervention for acceptability
- Phase II: establish behavior change rates, adherence, estimate treatment intensity needed for effect
- Phase III: evaluate effectiveness in comparison to other treatments
- Phase IV: long term monitoring, with or without comparison groups
- Other: dissemination/translation research, community interventions

Generalizability:

Is “reach” always a virtue?

- Example: does 80 mg aspirin daily reduce re-infarction rate?
 - “Run in:” Mailing back aspirin labels for 30 days.
 - This is an efficacy trial.

Generalizability: What is the question?

- Q1: Does 80 mg aspirin daily reduce re-infarction rates?
- Q2: If daily aspirin is recommended to a population, will infarction rates decline?
- Q3: If a major public information campaign is waged based on advertising, public service announcements, physician education etc.,
a) will daily aspirin use increase?
b) will infarction rates decline?

Generalizability: Effect of positive outcome on the recommendation

Q1: Does 80 mg aspirin daily reduce re-infarction rates?
If you had an MI and take aspirin, you may decrease risk of reinfarction.

Q2: If daily aspirin is recommended to a population, will infarction rates decline?
Public agencies (AHA) should endorse taking 80 mg aspirin daily.

Q3: If a major public information campaign is waged based on advertising, public service announcements, physician education etc., a) will daily aspirin use increase? b) will infarction rates decline?

Significant funds should be allocated for intensive promotion of daily aspirin use.

Sample size in behavioral trials

Case: Hypertension

- Most thoroughly studied area in behavioral medicine.
- Eisenberg meta-analysis, 1993:
 - Review of published clinical trials involving cognitive behavioral therapies.
 - Screened 800+ published studies.
 - Found 26 that met eligibility criteria.
 - Average sample size = 50.

Case: Hypertension

- Conclusion:
 - Cognitive therapies are superior to no therapy but not superior to credible sham techniques or to self-monitoring.
 - No single technique seems to be more effective than others.

Meta-Analysis of Psychoeducational Programs for CHD Patients

Number of trials included:	37
Median / Mean sample size:	147 / 243
Number randomized:	28 (76%)

Mean reduction in cardiovascular events: 34%

Conclusion: Programs which influence proximal targets affect distal targets (morbidity and mortality).

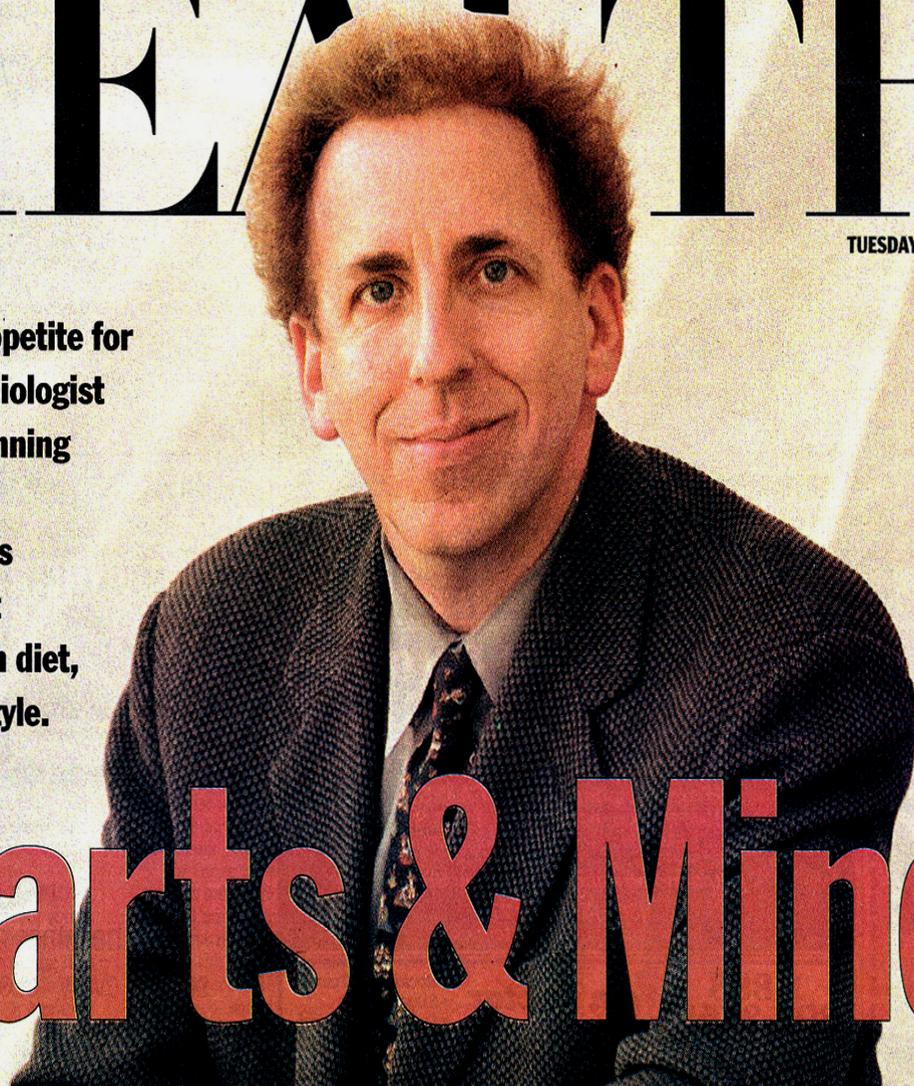
MEDICAL FRONTIERS | *Battling Heart Disease*

HEALTH

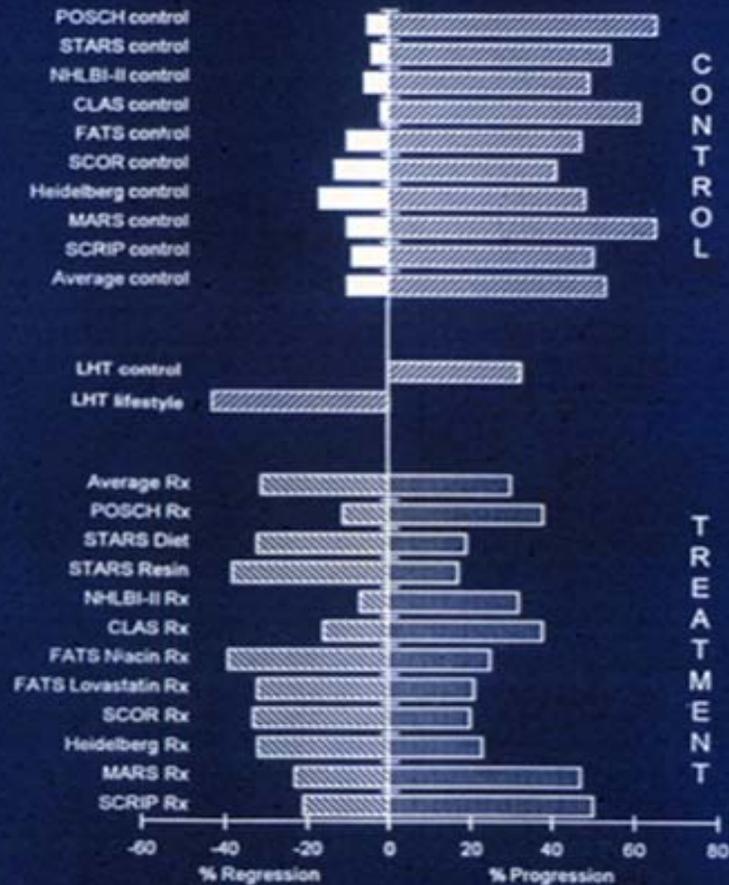
TUESDAY, JULY 24, 2001

Aided by a new appetite for cost-cutting, cardiologist Dean Ornish is winning powerful allies in Washington for his campaign to treat heart disease with diet, exercise and lifestyle.

Hearts & Minds



Regression of Coronary Artery Disease by Lipid Lowering: Angiographic Evidence



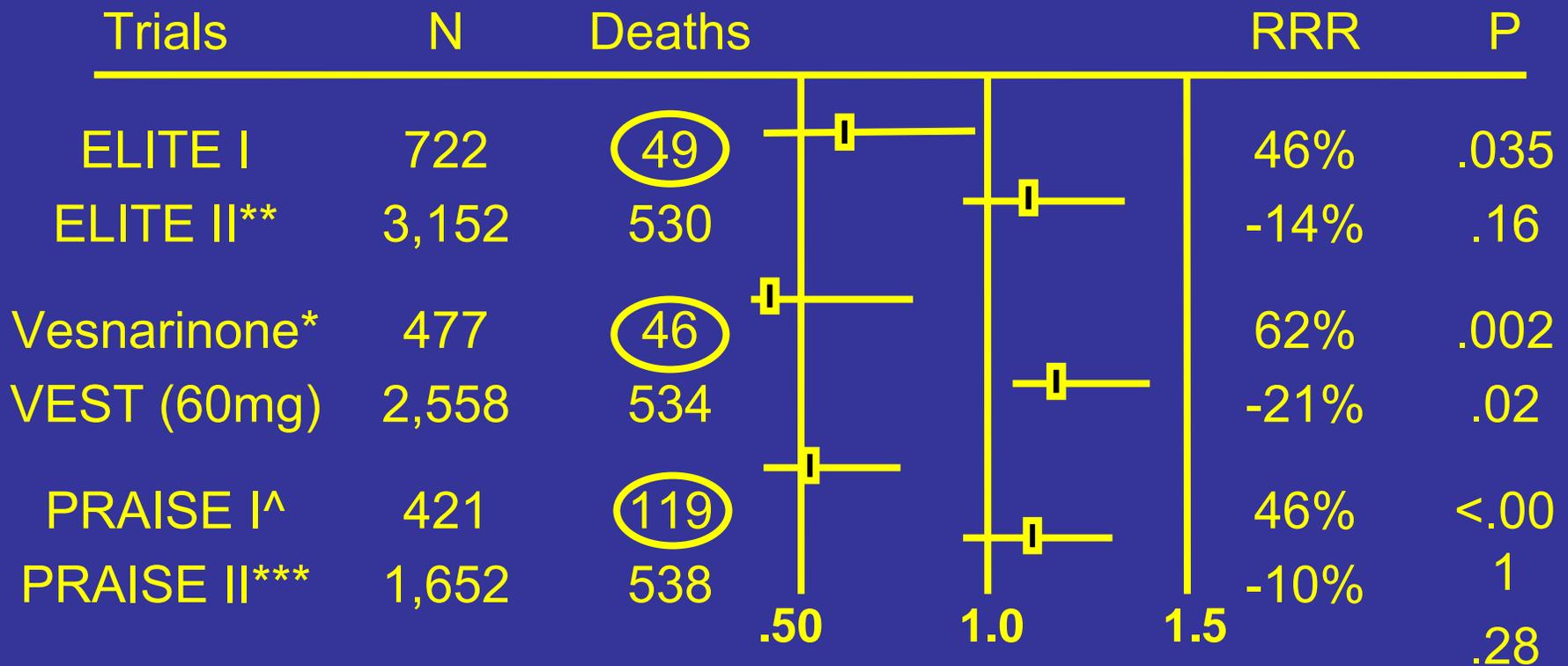
Adapted from Deedwania PC. *Contemp. Issues Cardiol.* 1995;79:973

Sample Size

Deaths or Events	Patients Randomized (Risk = 10%)	Chance of Type II Error *	Comments on Sample Size
0-50	<500	>0.9	Utterly inadequate
50-150	1000	0.7-0.9	Probably inadequate
150-350	3000	0.3-0.7	Possibly inadequate
350-650	6000	0.1-0.3	Probably adequate
>650	10000	<0.1	Adequate

* Probability of failing to achieve $p < .01$ if risk reduction =25%

Small Trials in Cardiology are Unreliable: CHF



^ (non-ischemic only)

*Inotrope; **ACE inhib; ***Ca++ block.

Equipoise

“Perception or belief that uncertainty exists concerning the effectiveness of available treatments, including the treatment to be investigated.”

Equipoise

Perceptions of equipoise are crucial to trial participation: a qualitative study of men in the ProtecT study:

“Belief in clinical equipoise was key to participants' consent to randomization.”

» **Mills et al., Controlled Clin Trials. 2003, 24:272-82.**

Multiple Risk Factor Intervention Trial (MRFIT) 1976-1983

□ 1: Am J Public Health 1979 Oct;69(10):996-1000

The psychological effects of differential treatment of a high risk sample in a randomized clinical trial.

Benfari RC, McIntyre K, Eaker E, Blumberg S, Paul O.

A study was carried out using 616 participants in a randomized clinical trial at the Harvard MRFIT (Multiple Risk Factor Intervention Trial) Clinical Center, to test if there were differences in the psychological dimensions of anxiety, depression, and functional heart symptoms in groups given different levels of treatment in a CHD (Coronary Heart Disease) Intervention Program. A theoretical framework was given to justify a number of hypotheses as to the induction of adverse psychological effects. At the end of two years in the MRFIT Program there were no significant

[Related Articles, Links](#)

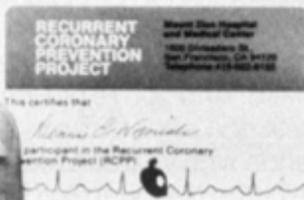
□ 1: JAMA 1976 Feb 23;235

The multiple risk factor intervention trial: primary prevention of coronary heart disease

PMID: 946311 [PubMed - indexed for MEDLINE]

TIME, JULY 17, 1978

THIS CARD JUST MIGHT SAVE YOUR LIFE.



For every 100 people who have had one or more heart attacks, within five years 45% will have a second heart attack—and half of these people will die.

Frightening? You bet. These are alarming statistics. But something is being done to substantially reduce these odds, and it's being done in San Francisco.

Called the Recurrent Coronary Prevention Project, it's a program

currently underway at Mount Zion Hospital.

Earlier work done by the Mount Zion research group under the direction of Meyer Friedman, M.D., Director of the Harold Brunn Institute at Mount Zion Hospital and Medical Center, has shown that the number of fatalities from a second heart attack can be reduced from five out of ten people to one out of ten.

If you have had a heart attack, here's what you can do. If you are currently a non-smoker, and do not suffer from diabetes, you are eligible to become a part of this program and receive your recurrent coronary prevention card which entitles you to all of the benefits of this project. There's no charge as the project is funded by the National Heart, Lung and Blood Institute.

**Act now
before it's too late.**

For full details, call the
RECURRENT CORONARY
PREVENTION PROJECT
(415) 922-8155

Typography:
Cohn/Comp, San Francisco

This ad prepared as a public service for the Recurrent Coronary Prevention Project by Scroggin, Reed Advertising, 843 Montgomery Street, San Francisco, California.

Mechanistic/Basic investigations

Should clinical trials be postponed until the mechanisms through which psychosocial risk factors act, or are modified, are understood?

The Case of Ventricular Arrhythmias and Sudden Death

- Ventricular Premature Complexes (VPCs) predispose to v. arrhythmias, sudden death.
- Long-acting, Class 1 antiarrhythmic drugs protect against VPCs;
- 3,000-patient post-marketing database.

Cardiac Arrhythmia Suppression Trial (CAST) - 1989

- n = 2309
- Post-myocardial infarction, 6 PVCs/hr
- Class 1 Antiarrhythmics: Encainide, Flecainide, Moricizine
- 10 months follow-up
- Outcome: RR of death 3.6 in favor of placebo for deaths and non-fatal events

“A remedy which is known to work, though nobody knows why, is preferable to a remedy which has the support of theory without confirmation in practice.”

Richard Asher
Lancet, 1961

Ancillary studies

- Defined:
 - Any data collection not related to answering the hypothesis concerning the effect of the intervention on health.
 - Any data collection not in the original design of the RCT. (Exception: Modification to RCT protocol.)
 - Source of funding (original budget or supplement) is irrelevant.

Ancillary studies

- Increase participant burden
- Increase drop-out rate
- Decrease adherence to treatment
- Increase possible adverse events
- Increase strain on resources

Ancillary studies

- New NHLBI policy:
 - Ancillary study proposals must be approved by the RCT Steering Committee:
 - Before implementation
 - Before submitting or writing a grant application
 - Must be approved by NHLBI

Criticisms of RCTs

- Gap between RCTs and clinical practice in psychiatry:
 - Populations in RCTs are highly selected.
 - Interventions in RCTs are simplified, do not reflect complexities of dealing with individual cases.

TenHave et al., Gen. Hosp Psychiatry, 2003

Criticisms of RCTs

- Specific criticisms:
 - High attrition and low adherence to treatment;
 - Contrast between high success rates in treating depression in RCTs vs. success rate in clinical practice;
 - Successful trials enroll highly motivated patients, use resource-intensive interventions;

TenHave et al., Gen. Hosp Psychiatry, 2003

Criticisms of RCTs

- Specific criticisms:
 - Full treatment algorithms involved in clinical practice are not investigated;
 - Contribution of components of algorithms are seldom examined;
 - Fewer than half of potential participants are willing to enroll in RCTs;

TenHave et al., Gen. Hosp Psychiatry, 2003

Alternative Study Designs

- Fixed adaptive design (e.g.: response adaptive randomization)
- Randomized adaptive design (sequential randomization, statistically influenced by previous results)
- Randomized consent (with option to switch at any time after initial Tx)
- Partially randomized patient preference (randomized but allowed to switch immediately)

TenHave et al., Gen. Hosp Psychiatry, 2003

Question

- Will the alternative designs avoid the problems attributed to RCTs, or compound them?

Action to Control Cardiovascular Risk in Diabetes

ACCORID



ACCORD Background: Diabetes & CVD

- Increasing glycemia level is associated with increasing CVD risk in observational studies
- Diabetic pts are more likely to have HTN or dyslipidemia
- Question: What is value of intensive control of CVD risk factors in reducing CVD rates?

Three Medical Strategies Tested in ACCORD (Three Trials in One)

- **Intensive glycemic control vs standard control**
- **Intensive blood pressure control vs standard control**
- **Treatment to increase HDL-C and lower TG + LDL-C Tx vs LDL-C Tx alone**

ACCORD Study Design

- Randomized multi-center clinical trial
- Conducted in 7 Clinical Networks w/ 70 clinical sites in the US and Canada
- N = 10,000 patients with type 2 diabetes at high risk for CVD events [Pts w/ and w/o existing CVD]
- Testing 3 questions: glycemia, lipid, BP
- Double 2 X 2 factorial design

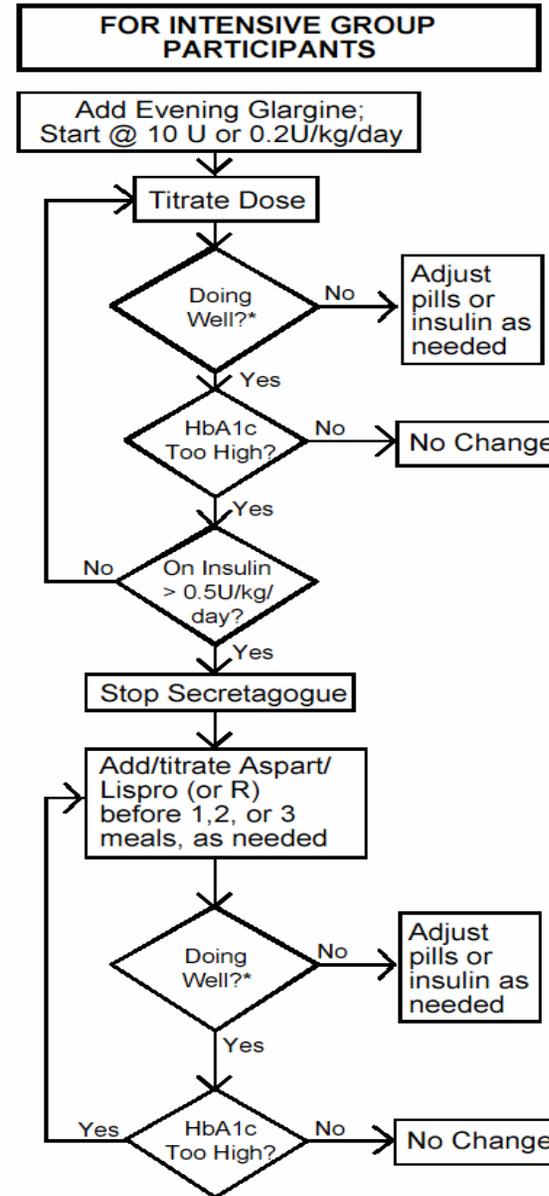
ACCORD Double 2 x 2 Factorial Design

	Lipid		BP		
	Fibrate	Placebo	Intensive	Standard	
Intensive Glycemic Control	1450	1450	1050	1050	5000
Standard Glycemic Control	1450	1450	1050	1050	5000
	2900	2900	2100	2100	10,000
	5800		4200		

Primary Outcome Measure

- **First occurrence of a major cardiovascular disease event:**
 - **Nonfatal MI**
 - **Nonfatal Stroke**
 - **Cardiovascular Death**
- Events adjudicated by committee masked to group assignment

Treatment algorithms used – Example: Suggested Approach to Insulin Use in the intensive glycemia group



*Doing well: no severe hypoglycemic or adverse event

Scientific integrity

- PI is responsible for:
 - Sound design
 - Avoiding bias
 - Instructing staff concerning expectations
 - Serving as role model for staff
 - Conforming to guidelines for publication of results

A Final Thought

- “Success is the ability to go from failure to failure without losing your enthusiasm.”

Winston Churchill, 1874-1965