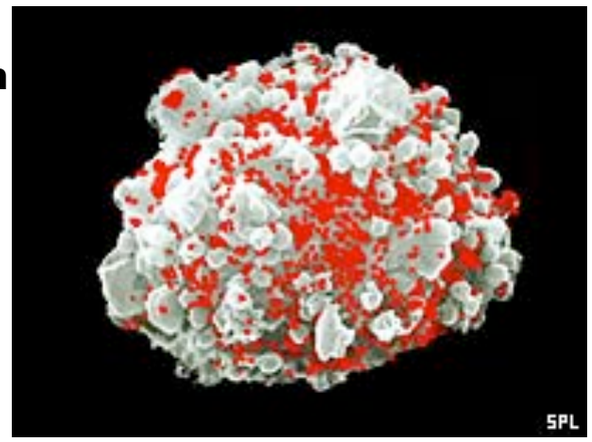
Basics of Interim Analysis

Dave Glidden
Division of Biostatistics, UCSF
11 April 2008

Last Updated: Friday, 21 September 2007, 21:52 GMT 22:52 UK

Merck abandons HIV vaccine trials

International drug company Merck has halted trials on an HIV vaccine that was regarded as one of the most promising in the fight against Aids.



The vaccine was loaded with copies of three HIV genes

Merck stopped testing the vaccine after it was judged to be ineffective.

In trials, the vaccine failed to prevent HIV infections among volunteers who were at risk of catching the virus, including gay men and sex workers.

Merck had previously expressed high hopes for the drug, which it spent 10 years developing.

'Headed for failure'

Merck's international trial, called Step, began in 2004 and involved 3,000 HIV-negative volunteers from diverse backgrounds, between the ages of 18 and 45.

Merck said that 24 of 741 volunteers who got the vaccine became infected with HIV, the virus that causes Aids.

SEE ALSO
Early success with new HIV drug
13 Apr 07 | Health
African HIV vaccine trial launch
08 Feb 07 | Health
New test for HIV drug resistance
08 Jan 07 | Health
'DNA target' to block HIV found
26 Feb 06 | Health
HIV vaccines show promise
26 Feb 02 | Health

RELATED INTERNET LINKS
Merck & Co
HIV Vaccine Trials Network
The BBC is not responsible for the content of external internet sites

TOP HEALTH STORIES
'Six-way' kidney transplant first
Heart risk may be set in the womb
Growth spurts may cause obesity
News feeds

MOST POPULAR STORIES NOW
MOST E-MAILED | MOST READ
1 'Breakthrough' at Stonehenge dig
2 IMF slashes world growth

The New York Times
nytimes.com

December 14, 2006

H.I.V. RISK HALVED BY CIRCUMCISION, U.S. AGENCY FINDS

By DONALD G. MCNEIL JR.

Circumcision appears to reduce a man's risk of contracting AIDS from heterosexual sex by half, United States government health officials said yesterday, and the directors of the two largest funds for fighting the disease said they would consider paying for circumcisions in high-risk countries.

The announcement was made by officials of the National Institutes of Health as they halted two clinical trials, in Kenya and Uganda, on the ground that not offering circumcision to all the men taking part would be unethical. The success of the trials confirmed a study done last year in South Africa.

AIDS experts immediately hailed the finding. "This is very exciting news," said Daniel Halperin, an H.I.V. specialist at the Harvard Center for Population and Development, who has argued that circumcision slows the spread of AIDS in the parts of Africa where it is common.

In an interview from Zimbabwe, he added, "I have no doubt that as word of this gets around, millions of African men will want to get circumcised, and that will save many lives."

Uncircumcised men are thought to be more susceptible because the underside of the foreskin is rich in Langerhans cells, sentinel cells of the immune system, which attach easily to the human immunodeficiency virus, which causes AIDS. The foreskin also often suffers small tears during intercourse.

But experts also cautioned that circumcision is no cure-all. It only lessens the chances that a man will catch the virus; it is expensive compared to condoms, abstinence or other methods; and the surgery has serious risks if performed by folk healers using dirty blades, as often happens in rural Africa.

What goes into those decisions?

Outline

- General idea of interim analysis
- Statistical methods for stopping
  benefit, harm, futility
- Numerous considerations
  not just statistical
- Special issues

General Idea

Clinical Trials

- Requires a degree of equipoise
  can erode as data accumulates
- Imperative on ensuring patient safety
- Investigators: poor positioned for these
- Independent, expert, rigorous assessments
- Interim Analysis

Approach

- Investigators develop plan for monitoring
  prior to starting the trial
- Convene a board of independent experts
  clinical area, statistics, ethics
  known as a DSMB
- Reach consensus with DSMB
- Perhaps, retain an independent statistician
  responsible for reporting to DSMB

Group Sequential

- Sequential: examine emerging results
- Group: defined by period of time/data
- Monitoring plan specifies timing
- May examine efficacy, safety, or data quality

Reasons to Stop a Trial

- Treatment benefit
- No treatment difference (futility)
- Severe toxicity/side effects
- Outside information renders trial unethical
- Poor data quality
- Problems with trial conduct

Stopping for Efficacy

Pressures to Stop Early

- Preserves resources
- Minimizes exposure to inferior rx
- Move on to new questions

Pressures to Keep Going

- Precisely estimate treatment effect
- Controlled safety data
- Suspicion of short-term trends
- Data on secondary outcomes, subgroups

Naive Analysis

- Suppose we plan 5 interim analyses
- At approximately equal periods
  more on this later
- Significance at $p < 0.05$
- True type I error probability = 0.142

Pocock (1977)

- Adapted the previous example
- To set the overall type I error
  calculated using Brownian motion
- Reject at any analysis if $p < 0.0158$
- Including the final analysis

O'Brien/Fleming (1979)

- Grows out of awkward feature of Pocock
  if trial completed, big penalty interim analysis
- Noted better to be stringent early
- Relax criteria at later analyses
- Final analysis at near desired alpha

Haybittle-Peto

- Interesting compromise
- Interim analyses at p=0.0027
  reject for Z-statistic greater than 3
- No matter how many interim analyses!
- Final analysis at near 0.05

Critical Values

	1	2	3	4	Final
Pocock	0.0153	0.0153	0.0153	0.0153	0.0153
O'Brien Fleming	$5\ 10^{-6}$	$1.3\ 10^{-3}$	$8.4\ 10^{-3}$	0.023	0.041
Haybittle Peto	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	0.047

Critical Values

	1	2	3	4	Final
Pocock	0.0153	0.0153	0.0153	0.0153	0.0153
O'Brien Fleming	$5\ 10^{-6}$	$1.3\ 10^{-3}$	$8.4\ 10^{-3}$	0.023	0.041
Haybittle Peto	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	$2.7\ 10^{-3}$	0.047

Alpha Spending

Lan and DeMets

- Based on concept of information fraction
  what % of data do you have?
- Creates bounds which vary with this
- Flexible number, timing of analyses
- Bounds dictated by "spending function"
  key choice
- Can't have timing dictated by results

Practical Considerations

- Desirable to be conservative early
  need strong early evidence
- Good to do final analysis near 0.05 level
- Pocock has smallest average sample size
  if effect is larger than expected
- Pocock has largest power cost
  power lost by allowing for interim analysis

More Considerations

- Interim analysis may be unplanned
  DSMB checks accrual sees trend

- O'Brien Fleming: very conservative
  at early timepoints

- Haybittle has simple criterion:
  ideal for unplanned analysis
  not excessively conservative

Best Practices

- Develop a detailed charter
  written buy-in is essential
- Explore scenarios with team and DSMB
- Allow for unplanned analyses
  Haybittle or alpha spending
- Develop plan for stopping in the unexpected direction

Stopping for Futility

STEP Interim Results

	HIV +	HIV-
Placebo	21	741
HIV Vaccine	24	717

logrank Z = 0.43, p = 0.66

Statistical Question

- Do we already have all the data we need?
- What is the probability of a reversal?
- Power to reject in favor in vaccine if study runs to completion
- Calculate so-called *conditional power*

Conditional Power

- Tool for monitoring futility
- Depends on data accrued
- How far into the study you are
- Original power of study
- Projection about trend in future data
  e.g., true log hazard ratio β

The Rough Calculation

- T: log rank test,
  reject in favor of vaccine if < -1.96
- D: data accrued to date
- β: true effect of the vaccine
- CP(D, β) = Pr(T < -1.96 | D, β)
- The vaccine effect is not really known

Lan and Wittes

- Simple method for calculating conditional power
- Calculate at 3 effects for vaccine:
  current trend, no effect, original alternative
- This gives a range of powers
- Declare futility is power is low in all settings

STEP Conditional Power

Scenario	β	Conditional Power
Current Trend	0.13	0.005
No Effect	0.00	0.012
Original Effect	-0.69	0.167

Easy Call

- Power low in all scenarios
- Concern about harm
  higher incidence in vaccinated
  infections in >1 injection: 19 vac, 11 plac
- Continuing unethical and unproductive

MIRA Study

- DSMB suggest termination due to futility
- Six months of FU left in study
- Would only shorten study by months
- Leave the study with a stigma
- DSMB reversed their decision

Complications

- Cond. power varies by β
  possibly by a large margin
  predictive power does Bayesian averaging
- Controversy about whether to stop
  accrual, treatment, follow-up
- May effect secondary endpoints

Sample Size Re-estimation

Internal Pilot

- Sample size calculations depend on unknown parameters
- Circularity is very frustrating
- Pilot data is often small/suboptimal for instance, from other populations
- Gives rise to interest in internal pilot study

Internal Pilot

- Begin desired study: fix a minimum sample size: n_0
- After some fraction, estimate parameters
- Estimate new sample size n_1.
- Choose n_1 if it exceeds n_0 (by some amt)

Issues

- n_1 shouldn't depend on rx difference
  leads to *type I error inflation*
  if negative study extended until positive
- Can depend on SD, baseline prevalence, etc
- Suggested technique: blinded estimation of these quantities.
- Inflation of type I error is slight
- Power reduced if pilot is small

A Few Observations

- Not aware of many examples in practice
  quick search only turned up 2 examples
- Less compelling for survival data
- Interesting not used to reduce sample size
  reflects conservatism
  undesirability of underpowered study

Analysis Following Sequential Testing

Point Estimation

- Study's stopped: biased treatment effects
  tend to be too large
- Bias reduction is not straightforward
  complex expressions for bias
  bias may be estimated
  `correcting' for bias can increase variance
- Does not appear implement in practice

Adjusted p-values

- O'Brien-Fleming monitoring
- p-value 0.0013 at 2nd analysis
  right on the boundary
- Truth is, it is barely significant at overall alpha
  kinda like a p of 0.05
- P-value: pr of result as or more extreme under H_0
- Z=-3.01 at second analysis, Z=-4.20 third analysis
  which is more extreme?

Alignment Problem

- Not straightforward to order the sample space
- Several ways to do it
- Affects ability to “adjust” confidence intervals and p-values
- Results often similar
- Appears to be rare to adjust CI/p-values

Overall Issues

- Interim analysis permits flexibility
  ability to stop or extend
- Can complicate analysis, interpretation
  bias, hard to get CIs
- Context important in decisions
- Experience too
  hard to benefit from
- Consider some examples

Breast Cancer Prev.

- Breast Cancer Prev Trial, 4/92-3/98
- Double-blind, tamoxifen v. placebo
- Tamoxifen expected to have +/- effects
  benefits to heart and breast cancer
  increase in endrometrial cancer
- Need to weight these effects

Trial Details

- Powered for a 30% reduction
- N=13,000 high risk women
- alpha = 0.01
- develop global index of outcomes
- O'Brien Fleming boundaries

DSMB

- Extensive pre-trial deliberations including voting on contrived scenarios
- Concern for balancing effects
- Imperative to protect healthy volunteers
- Concern about long-term benefits

Unexpected Issue

- Ocular substudy
- Excess of cataracts developed
- 37 placebo, 59 tamoxifen
- Consent revised
- Letter to physicians

Results

	Breast Cancer			
Date	P	T	P-value	Bound.
3/95	70	36	0.028	$1.3\ 10^{-4}$
4/96	89	45	$9\ 10^{-5}$ *	$1.4\ 10^{-4}$
3/97	124	65	$1\ 10^{-5}$ *	$1.5\ 10^{-4}$
3/98	154	85	$6\ 10^{-6}$ *	$1.7\ 10^{-4}$

* crossed boundary

CARET Study

- CARET: 7/88-3/94
- β carotene v. placebo
- Prevention of invasive lung cancer
- Heavy smokers/asbestos workers
- Cancer chemoprophylaxis
- Result of observation epidemiology

Unexpected News

- April 1994
- Finnish study about to be published
- β carotene lung cancer prevention trial
- Approximate 18% increase in lung cancer *among those taking β carotene*
- Triggers a review of CARET Data

Result

- Meeting help in 8/94
- Excess lung cancer in one arm
- Unblinding reveals this to be β carotene
- Did not cross O'Brien Fleming
- Robust discussion follows

Don’t Stop

- It hasn’t crossed the boundary
- Accelerated detection pre-existing cancer?
- Mechanism for harm unclear
- Stopping will take out other trials
- “We need to be sure about this”

Stop Now

- Unlikely to be a chance finding
  given the ATBC results
- DSMB should act to protect participants
- Evidence of harm beyond lung cancer
  some excess heart disease

Decision

- Look again next year
- See if this persists
- Ask statisticians for conditional power
  can this trend be reversed?

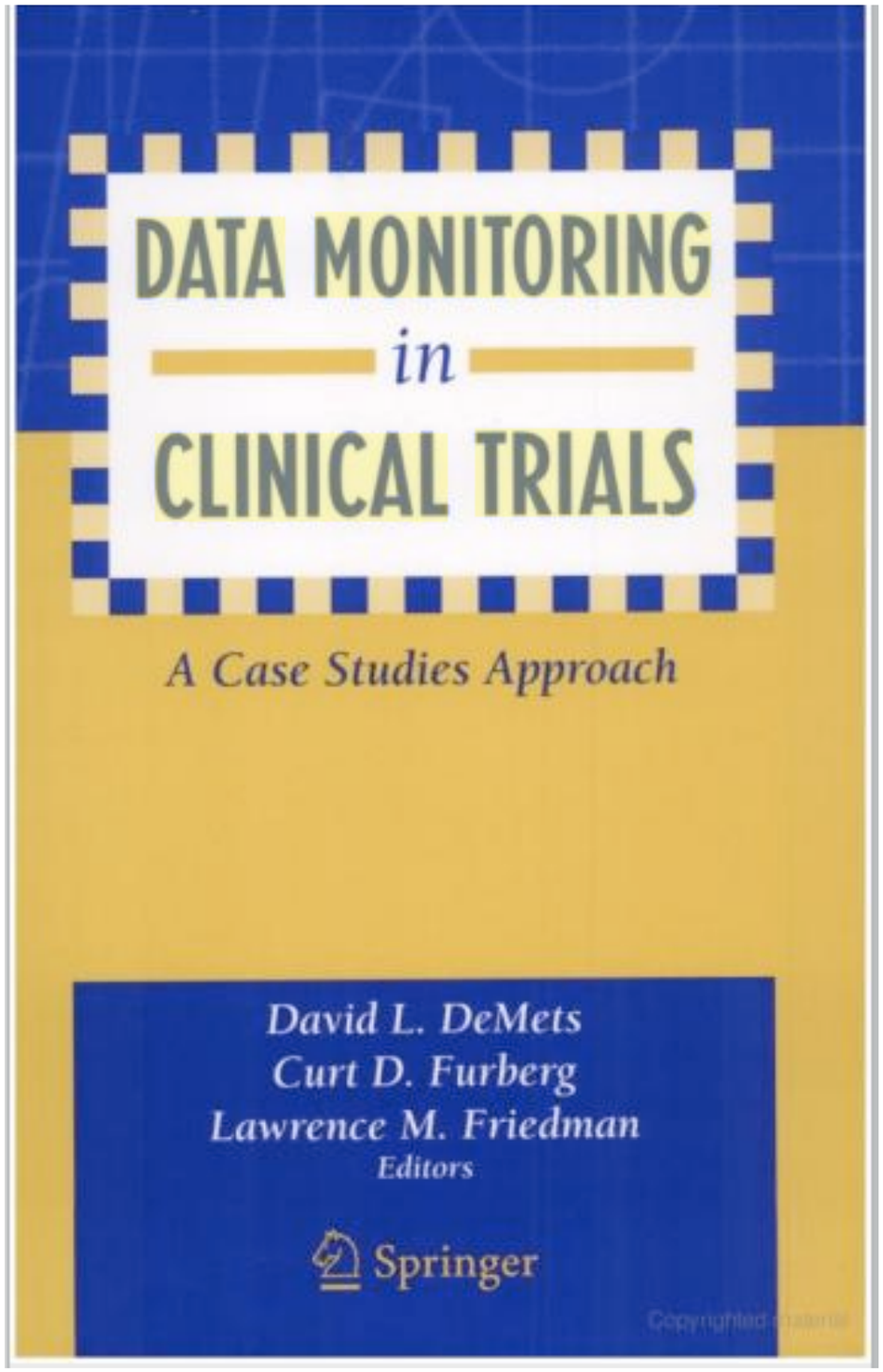
Meeting in 9/95

- Trend has persisted
- Excess of lung cancer
- Excess of cardiovascular disease
- "Extremely unlikely" to show benefit
  based on conditional power calculation
- Unanimous decision to stop

Recommendations

- Statistics doesn't provide unambiguous guidance
- Other considers very important
- Benefit from the experience of others
- Hard because interim analyses not published
- Typically, deliberations secret

Helpful Book

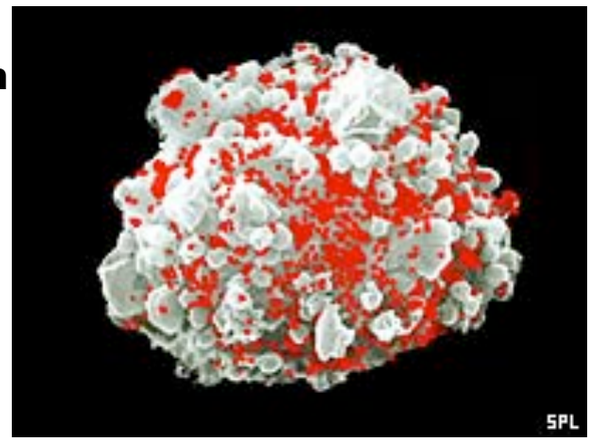


Talks about specific studies, deliberations, lessons learned

Last Updated: Friday, 21 September 2007, 21:52 GMT 22:52 UK

Merck abandons HIV vaccine trials

International drug company Merck has halted trials on an HIV vaccine that was regarded as one of the most promising in the fight against Aids.



The vaccine was loaded with copies of three HIV genes

Merck stopped testing the vaccine after it was judged to be ineffective.

In trials, the vaccine failed to prevent HIV infections among volunteers who were at risk of catching the virus, including gay men and sex workers.

Merck had previously expressed high hopes for the drug, which it spent 10 years developing.

'Headed for failure'

Merck's international trial, called Step, began in 2004 and involved 3,000 HIV-negative volunteers from diverse backgrounds, between the ages of 18 and 45.

Merck said that 24 of 741 volunteers who got the vaccine became infected with HIV, the virus that causes Aids.

SEE ALSO
Early success with new HIV drug
13 Apr 07 | Health
African HIV vaccine trial launch
08 Feb 07 | Health
New test for HIV drug resistance
08 Jan 07 | Health
'DNA target' to block HIV found
26 Feb 06 | Health
HIV vaccines show promise
26 Feb 02 | Health

RELATED INTERNET LINKS
Merck & Co
HIV Vaccine Trials Network
The BBC is not responsible for the content of external internet sites

TOP HEALTH STORIES
'Six-way' kidney transplant first
Heart risk may be set in the womb
Growth spurts may cause obesity
News feeds

MOST POPULAR STORIES NOW
MOST E-MAILED | MOST READ
1 'Breakthrough' at Stonehenge dig
2 IMF slashes world growth