Complex Time to Event Data: Design and Statistical Inference for the INVESTED Trial

KyungMann Kim, Ph.D.
University of Wisconsin-Madison
kyungmann.kim@wisc.edu

Outline

- INVESTED Trial
  - Overview
  - Organizations
  - DCC responsibilities
- Complex time to event data
  - Design
  - Statistical Inferences
    - Non-randomized cohorts
    - Mediation analysis

INVESTED Trial

- INfluenza Vaccine to Effectively Stop cardioThoracic Events and Decompensated heart failure (INVESTED) trial
- ClinicalTrials.gov Identifier: NCT02787044
  - https://www.clinicaltrials.gov/ct2/show/NCT02787044?term=NCT02787044&rank=1

THE UNIVERSITY of WISCONSIN MADISON

INVESTED

NIH National Heart, Lung, and Blood Institute

https://www.investedtrial.org

INVESTED: Overview

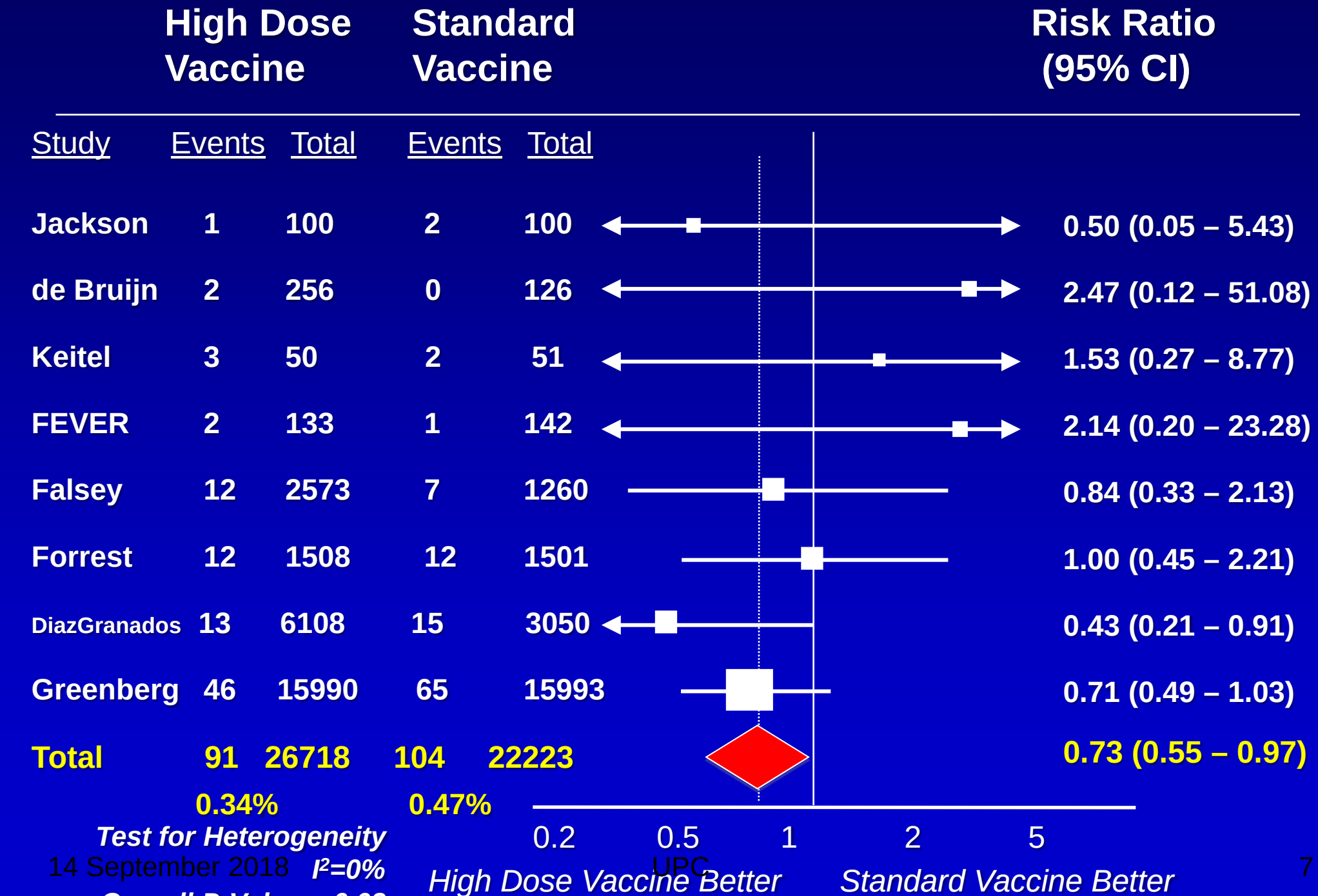
- Large, simple, adequately powered, double-blind and pragmatic trial
- Comparative effectiveness research
- Assess whether high-dose trivalent influenza vaccine (HD TIV/IIV3-HD) compared with standard dose quadrivalent influenza vaccine (SD QIV/IIV4-SD) will reduce cardiopulmonary events including death and hospitalization
- A high-risk cardiovascular population
  - MI within a year
  - HF within two years

Impact of Influenza in US

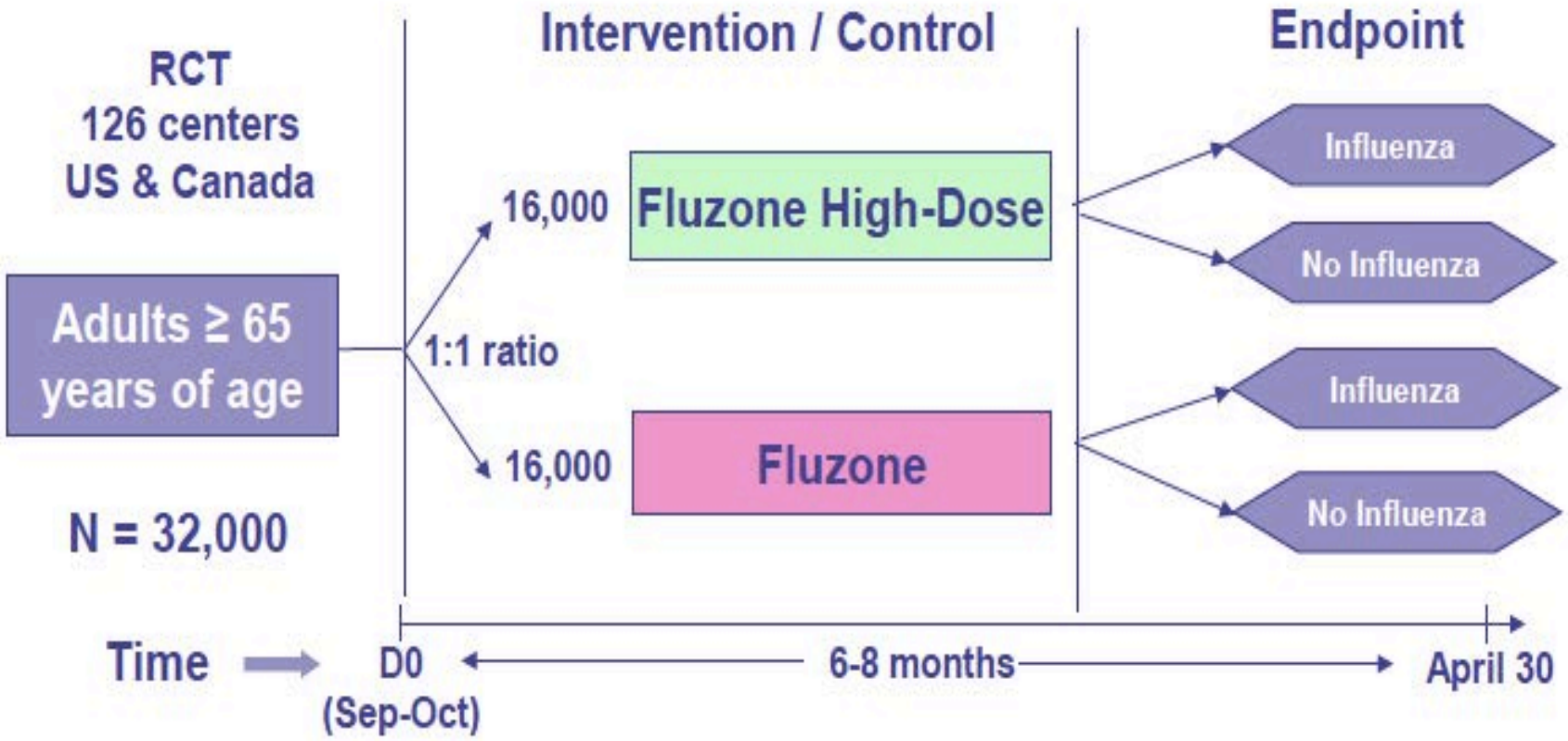
- Approximately 36,000 influenza-associated deaths during each influenza season
- Over 200,000 influenza-related excess hospitalizations
- Several analyses have documented an association between acute respiratory infections and cardiovascular (CV) events

Thompson et al JAMA. 2003;289:179-86
Thompson et al JAMA. 2004;292:1333-40
Madjid et al. EHJ. 2007;28:1205-10

More Intensive Influenza Vaccine Reduces CV Events: Meta Analysis



High vs Standard Dose Influenza Vaccine RCT in Healthy Elderly Individuals



Study conducted over two influenza seasons

Primary endpoint based on influenza caused by any influenza strain associated with a protocol-defined ILI

DiazGranados CA, et al. NEJM 2014;371:635-45

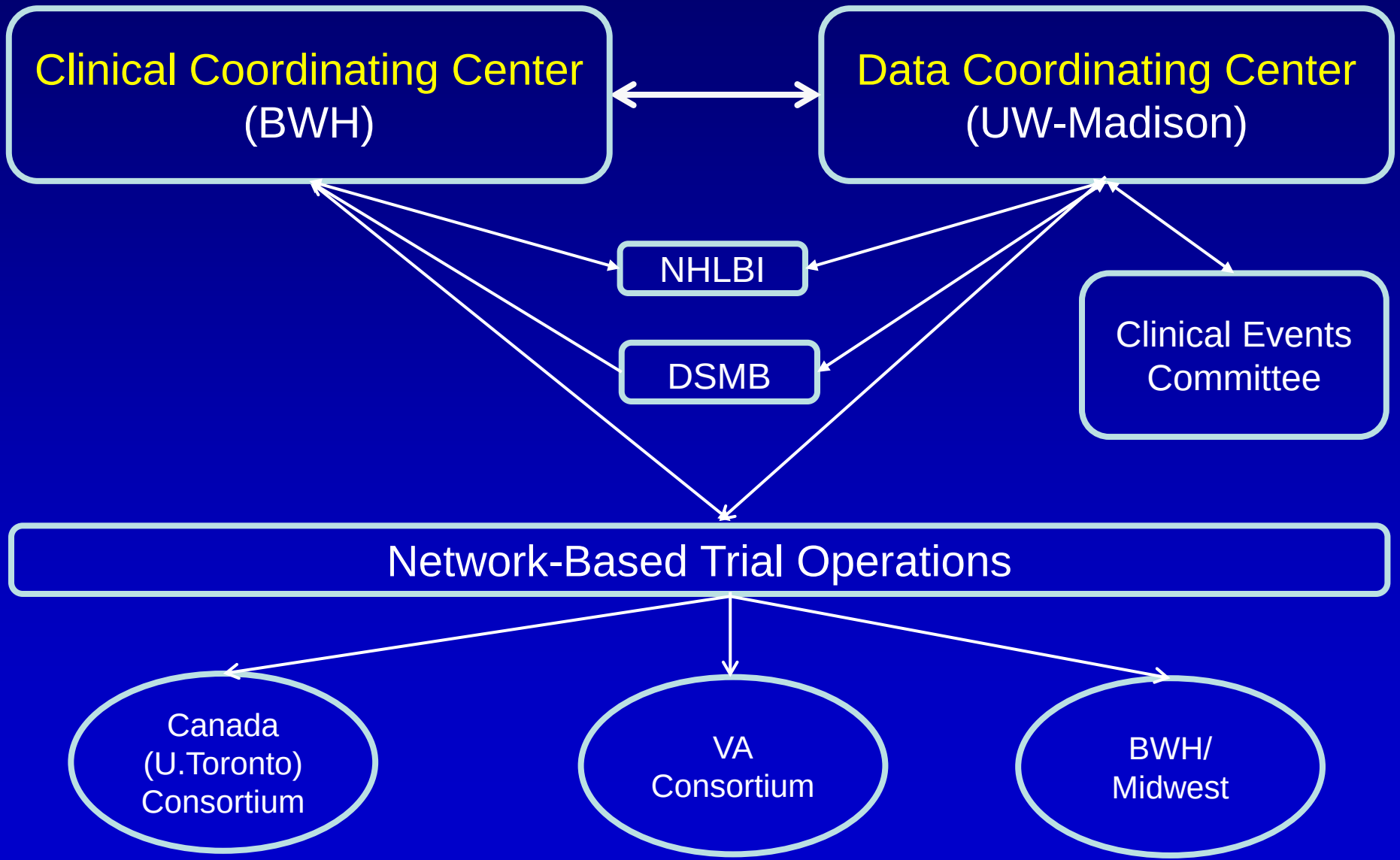
Features of the FLUZONE trial

- Design Hemagglutinin (HA) as influenza antigen
  - 1:1=IIV3-HD (60 μg HA/strain) : IIV3-SD (15 μg HA/strain)
  - Primary efficacy endpoint: Influenza-like illness 14 days after vaccination until the end of the influenza season
  - 30,000 to detect a relative efficacy of 30% with 1-β=0.8 at α=0.05 with an incidence of 2% for IIV3-SD
- Results: 09/06/11-05/31/13, 31,989
  - Year 1 (09/06/11-10/09/11): 14,500 new
    - H1N1, H3N2 (A/Victoria/210/2009) & B/Brisbane/60/2008
  - Year 2 (10/09/12-10/21/12): 17,489=7,645+9,844 new
    - 7,645 from year 1 re-randomized in year 2
    - H1N1, H3N2 (A/Victoria/210/2009) & B/Texas/6/2001
  - Relative efficacy: 24.2% reduction in the incidence of influenza-like illness (relative risk of 0.758)

INVESTED: Rationale

- Influenza is associated with and may trigger cardiovascular events, and may lead to disease exacerbation, especially in immune compromised conditions such as heart failure (HF)
- Influenza vaccine provides cardiovascular benefit in a meta-analysis of vaccine trials
- High risk patients, including those with HF or recent acute coronary syndrome/myocardial infarction (MI), may derive greater benefit from vaccination
- Patient with heart failure exhibit reduced immune responses to influenza vaccination which can be overcome with a higher dose of influenza vaccine
- In several analyses, high dose vaccine is associated with reduction in CV events
- High dose vaccine is currently approved for healthy older adults only; CDC's Advisory Committee on Immunization Practices does not preferentially recommend one vaccine formulation over another

INVESTED: Organization



INVESTED CCC

- MPI: Orly Vardeny, U of Minnesota
  Scott Solomon, BWH
- Cooperative agreement: U01 HL130163
- Funding period: 02/15/16-01/31/21
- Responsibilities:
  - Study operations
  - Recruitment of investigators and sites
  - Human subject protection
  - Regulatory affairs
- https://www.nhlbi.nih.gov/events/2011/data-coordinating-centers-best-practices

INVESTED DCC

- PI: KyungMann Kim, UW-Madison
- Cooperative agreement: U01 HL130204
- Funding period: 02/15/16-01/31/21
- Responsibilities:
  - Statistical methods
  - Data management
  - Quality control/assurance
  - Study monitoring
- Data management subcontract with Frontier Science

INVESTED: Executive Committee

Scott D. Solomon, MD
Professor of Medicine
Harvard Medical School
(CCC Co-PI)

Jacob A. Udell, MD, MPH
Assistant Professor of Medicine
University of Toronto
Canadian Co-PI

Keipp Talbot, MD, MPH
Assistant Professor of Medicine
Vanderbilt University

Orly Vardeny, PharmD, MS
Assoc Prof of Pharmacy and Medicine
University of Wisconsin
(CCC Co-PI)

Michael Farkouh, MD, MSc
Professor of Medicine
University of Toronto
Canadian Co-PI

Allison McGeer, MD, MSc
Professor of Laboratory Medicine, Pathobiology, and Public Health Sciences
University of Toronto

KyungMann Kim, PhD
Professor of Biostatistics and Statistics
University of Wisconsin
(DCC-PI)

J. Michael Gaziano, MD, MPH
Professor of Medicine
Harvard Medical School
VA network PI

Adrian Hernandez, MD, MHS
Professor of Medicine
Duke University
PCORnet network lead

NIH Project Team

Lawton Cooper, MD, MPH,
Program Officer
Rebecca Campo, PhD
Nicole Redmond, MD, PhD
Song Yang, PhD

Steering Committee Members:

Janet Wittes, PhD
Jonathan Temte, MD
Brian Claggett, PhD
Clyde Yancy, MD
Shaun Goodman, MD
Christopher Cannon, MD
Deepak Bhatt, MD
Pat Winokur, MD

Clinical Endpoint Committee:

Akshay Desai, MD, Chair
Peter Finn, MD
Jonathan Strongin, MD

DCC: IT Support

- Randomization system
- Treatment inventory utility
- Interface with eSOCDAT at CCC
  - Site management
  - Clinical events classification (soft adjudication)

DCC: Data Management

- Electronic Data Capture (EDC): OpenClinica
  - 21 CFR Part 11 compliant
  - Web-based data entry and management system
  - Audit trails
- Backend RDBMS: Ingres
  - Study database for statistical analysis and reporting

DCC: Quality Control

- Standard Operating Procedures (SOPs)
- Good Clinical Practice (ICH E6)
- Centralized risk-based monitoring
- Delinquency monitoring
- Data consistency and correctness
- Source data verification
  - Random sampling (5%)
  - Remote
- For-cause site visits if necessary

DCC: Quality Assurance

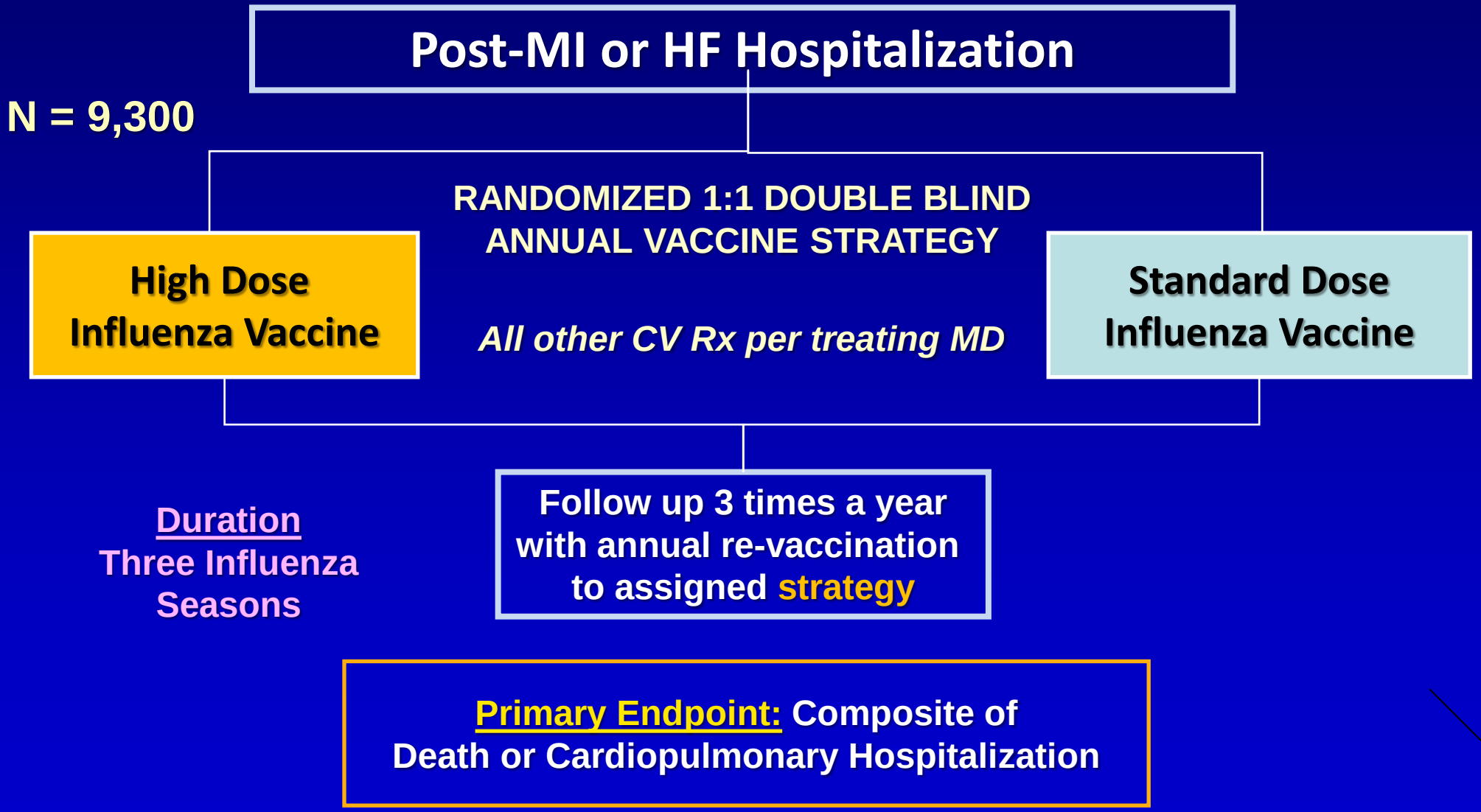
Quality Assurance Manager Oversees all QA activities

Standard Operating Procedures	Monitoring Plans	Management Controls	GCDMP
• Corporate • Project-specific to ensure project-specific goals are achieved • Annually reviewed and updated • Staff compliance monitored as part of employee annual reviews	• Study-specific quality monitoring plans specify deliverables and quality standards • QA/QC plans go beyond monitoring plans, specifying quality review processes for individual data items • Data collection instruments include built-in data validation and quality control	• Organization steering and compliance committee gives general oversight and guidelines to all projects • Individual management groups are established based on project needs • Independent software quality assurance department proactively audits software compliance	• Frontier Science's SOPs are based on GCDMP requirements • Routine annual review of all internal processes in the context of GCDMP ensures new and updated practices are compliant

DCC: Study Monitoring

- Central remote monitoring
  - Enrollment by site and by network
  - Trial conduct and performance
- Source document verification of 5% random samples
  - Informed consent
  - Eligibility
- Safety reporting for suspected unexpected serious adverse reactions (SUSARs) to Health Canada
- Data entry and query resolution
- Lag based on study schedule

INfluenza Vaccine to Effectively Stop CardioThoracic Events and Decompensated Heart Failure in Patients with CVD (INVESTED)



INVESTED Vaccines

- Inactivated influenza vaccine (IIV)
- Fluzone® donated by Sanofi
- Standard dose quadrivalent influenza vaccine (IIV4-SD)
  - Each at 15 μg hemagglutinin (HA)
  - Targets 4 strains:
    - A/H1N1, A/H3N2, B/Yamagata plus B/Victoria
  - Approved for 6 months of age and older
- High dose trivalent influenza vaccine (IIV3-HD)
  - Each at 60 μg HA
  - Targets 3 strains:
    - A/H1N1, A/H3N2, B/Yamagata
  - Approved for 65 years of age and older
- IND exemption from FDA

(Original) Design in Grant Proposal

- Enrollment during three influenza seasons (from September to January)
- Primary endpoint: Time to all-cause death or cardiopulmonary (CP) hospitalization
- Two-tailed log rank test at α=0.05
- Effect size: 18% reduction, i.e. hazard ratio (HR)=0.82
- Control event rates: 9% in 1st season; 8% in 2nd; 7% in 3rd
- Follow up ≥ 6 months with 20% drop-out per year
- 9,300 pts (4,650 in 1st season; 3,100 in 2nd; 1,550 in 3rd)
- 1,088 primary endpoint events
- Power 1-β > 0.90
- Two interim analyses using O'Brien-Fleming

gsSurv

- Sample size and power analysis for clinical trials with time to event endpoint
  - Lachin and Foulkes (1986)
  - Non-uniform entry, losses to follow-up, noncompliance
  - Non-constant event rates
- Group sequential trials with time to event endpoint
  - Kim and Tsiatis (1990)
- gsSurv by Keaven Anderson at Merck
  - Combines the flexibility of Lachin and Foulkes (1986) with the group sequential design of Kim and Tsiatis (1990)
  - Directly applied for design of INVESTED

gsSurv call and results

gsSurv(k=3, test.type=2, sfu="OF", lambdaC=- c(log(.91),log(.92),log(.93)), S=c(1,1), R=c(.5,.5,.5,.5,.5), gamma=c(3,0,2,0,1), hr=0.82, T=3, minfup=0.5, alpha=0.05, beta=0.1, sided=2, eta=0.2)

Time to event group sequential design with HR= 0.82

Equal randomization: ratio=1

Symmetric two-sided group sequential design with 90% power and 2.5% Type I Error.

Spending computations assume trial stops if a bound is crossed.

Analysis	N	Z	Nominal p	Spend
1	363	3.47	0.0003	0.0003
2	725	2.45	0.0071	0.0069
3	1088	2.00	0.0225	0.0178
Total			0.0250	

++ alpha spending: O'Brien-Fleming boundary

Complex Time to Event Data: Design/Statistical Inference Options

1. randomize once; first event counted; follow until end of study; analysis stratified by season (original ITT plan)
2. randomize once; first event counted (across seasons); follow until patient refuses vaccine; analysis stratified by season
3. randomize once; first event each year counted; follow until patient refuses vaccine; analysis stratified by season
4. randomize each year; first event each year gets counted; follow until patient refuses vaccine; analysis stratified by season

Revised Primary Endpoint

- Time to first occurrence of all-cause death (30%) or cardiopulmonary hospitalization (70%) within each season (from 14 days after vaccination until July 31)
- Except for death, CP hospitalizations will be counted for multiple vaccinations
- Examples (non-exclusive):

• **non-fatal myocardial infarction**	• **unplanned revascularizations**
• **non-fatal stroke**	• **arrhythmia**
• **non-fatal cardiac arrest**	• **pulmonary embolism**
• **unstable angina**	• **respiratory tract infections**
• **incident or acute heart failure**	• **pulmonary disease exacerbations**

Secondary Endpoints

- (Original) Primary endpoint over the entire study (ITT)
- Recurrent CP hospitalizations subject to competing risk of death
- Primary endpoints only during “influenza season” (until end April-mid May)
- Individual components of the primary endpoint
- Other secondary endpoints representing composites of key CV and pulmonary events

Sample Size/Power Analysis

- Effect size: 18% reduction or hazard ratio (HR) 0.82
- Control event rates: 9% in 1st season; 8% in 2nd; 7% in 3rd
- 30% : 70%=death : CP hospitalization
- 30% not returning for subsequent years' vaccinations
- Primary endpoint events: 279, 448 and 549 in 1st, 2nd, 3rd
- A total of 1,276 events over three seasons
- Power=0.94 to detect HR=0.82 at a two-sided α=0.05 log-rank test
- Two interim analysis using O'Brien-Fleming

Analysis of Efficacy Endpoints

- Subject's clock for each influenza season resets 2 weeks after influenza vaccination
  - Primary endpoint counted until July 31 of each season
  - Each subject can contribute primary endpoint events in more than one influenza season (considered independent?)
- Primary efficacy analysis (Specific Aim 1)
  - Log-rank test stratified by season, unadjusted estimate of HR
  - Cox proportional hazards regression, adjusted estimate of HR
- Secondary efficacy similar to primary efficacy
  - Recurrent events analysis subject to competing risk of death
- Additional efficacy analysis (asked by influenza experts)
  - In season analysis (events counted until end April-mid May)

Randomized vs Non-randomized

	2016-2017	2017-2018	2018-2019	2019-2020
2016-2017	494	298		
2017-2018		2,502		
2018-2019				
2019-2020				

Statistical Analysis Plan

- No re-randomization
  - As a strategy trial
  - To avoid dilution of effect due to possible carry-over effects
- After the initial randomization, in subsequent seasons
  - Bias due to differential survivorship
  - Bias due to differential drop-out
  - Two treatment groups no longer comparable
  - Randomization analysis maybe problematic
- Solutions: Causal inference?
  - Principal stratification
  - Matching based on propensity score
  - Inverse probability of treatment weighting

Statistical Analysis Plan

- Causal inference
  - Complex composite endpoint
  - Recurrent events subject to competing risk of death
- Potential methodology research topics
- Lu Mao, Co-I
- Potential dissertation topics

Analysis of Immune Responses

- Analysis of Immune Responses in HA inhibition (HAI)
  - T-test for geometric mean titers (GMTs)
  - Chi-square tests for seroconversion (SC) and seroprotection (SP)
  - Log-rank test of primary endpoint by status of SC and SP
  - Cox regression model with GMT as a model term, while adjusting for treatment, SC and SP and the interaction between treatment and match for circulating B (Victoria)-lineage to estimate HR for each doubling of GMT
- Association between immune response and primary endpoint (Specific Aim 2)

Association between Immune Response and Clinical outcomes

- Gilbert et al. (2014)
- Association between fold rise in varicella zoster virus (VZV) antibody titers and protection from herpes zoster, i.e. shingles
  - Zostavax Efficacy and Safety Trial (ZEST)
  - Correlate of Protection (CoP): Fold rise in antibody titer level
- No VZV antibody titers measured from placebo
- Validation of CoP as a surrogate endpoint
  - Prentice framework (1989)
  - Principal stratification or vaccine efficacy (VE) framework

Association between Immune Response and Clinical endpoint

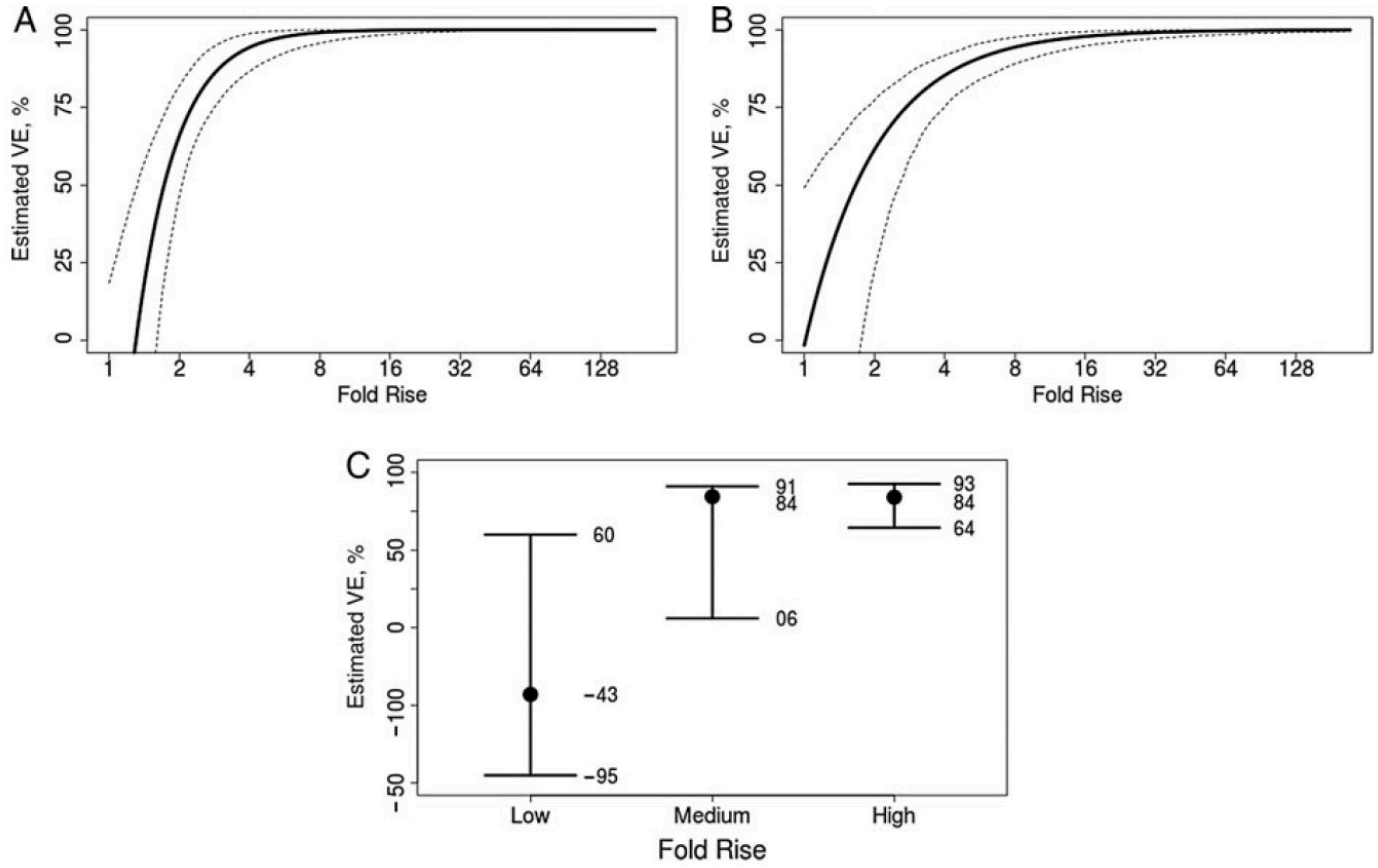


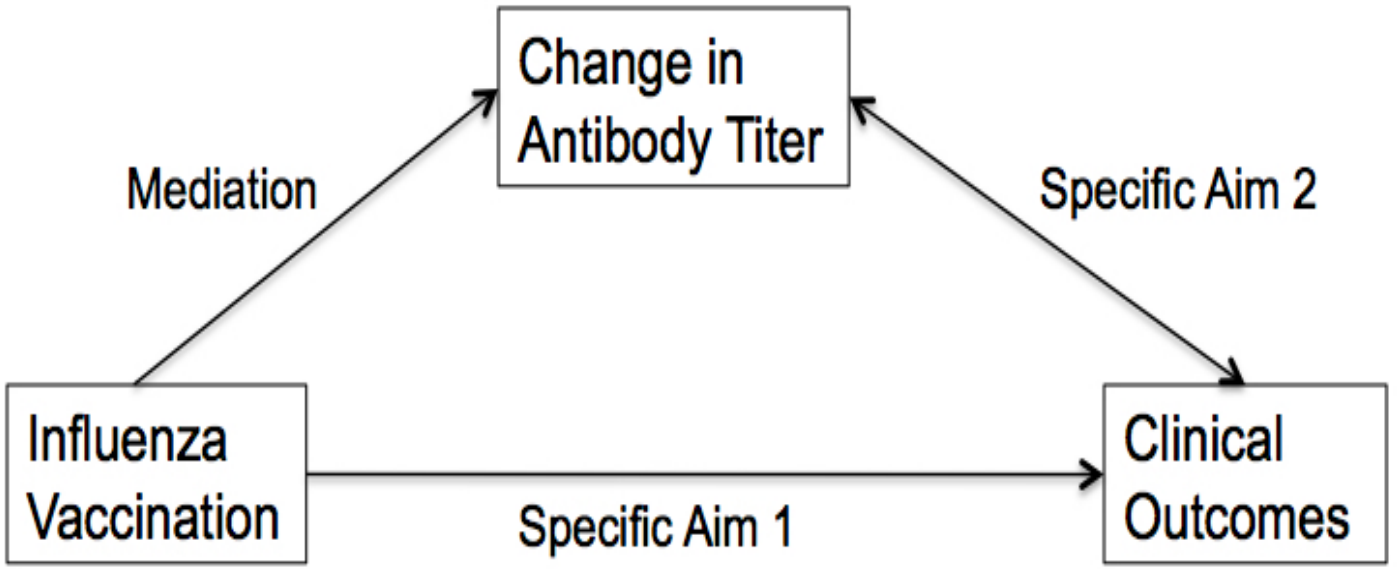
Figure 3. *A* and *B*, Estimated vaccine efficacy curves across levels of vaccine-induced fold rise in titers from baseline to week 6, using the probit estimated likelihood method [27] and the Weibull estimated likelihood method [35], respectively, with 95% bootstrap confidence intervals. The lower *x*-axis indicates the multiplicative fold rise in titers. *C*, Estimated VEs with 95% bootstrap confidence intervals for subgroups defined by the lower, middle, and upper tertiles of vaccine-induced fold rise in titers, using the nonparametric estimated likelihood method [27].

Other Challenges

- Competing Risks
  - Non-terminating individual components of the composite endpoint analyzed using methods for competing risks
  - Analysis of the rate of hospitalization with death as a competing risk
- Mediation analysis of immune response
  - No available method for Cox proportional hazards model
- Missing Data
  - Guided by the National Research Council report (2010)

Mediation Analysis

Fig. 2 Direct and Indirect Effects of Influenza Vaccination with Immune Modulation



Mediation Analysis

- Baron and Kenny (1986)
- Structural equation modeling (SEM)
- Most available methods deal with linear models
- Time to event data requires intrinsically non-linear models for hazard function or some transformation of it
- Wesley Chang's thesis topic
  - Linear transformation models (Cheng et al., 1995)

Efficacy Stopping Rules

- For efficacy comparisons
- At the end of each influenza season (calendar-driven) based on the design (information-driven)
- Lan-DeMets type I error spending function à la O'Brien-Fleming group sequential method

Analysis at the end of influenza season	Information time	Number of primary endpoint events	Upper efficacy boundary	Nominal p-value
1st	0.219	279	4.65	<0.0001
2nd	0.570	727	2.75	0.0060
3rd	1.000	1,276	1.98	0.0481

Efficacy Stopping Rules: Calendar/Duration paradigm

- For efficacy comparisons
- At the end of each influenza season based on observed
- Lan-DeMets type I error spending function à la O'Brien-Fleming group sequential method
- Observed so far grossly different from expected based on the design
- How to determine the group sequential boundary
- Information vs duration paradigm
  - Lan and DeMets (1989)
  - Lan and Lachin (1990)
  - Kim et al. (1995)

Discussion

- INVESTED trial
  - Large, simple trial
  - Pragmatic trial
  - Comparative effectiveness research
- Challenging statistical inference issues
  - Recurrent events subject to competing risk of death
  - Causal inference due to non-random cohorts after the 1st vaccination
  - Medication analysis for time to event data with immune responses as mediator
  - Interim analysis and group sequential boundary

References

- Anderson K. https://cran.r-project.org/web/packages/gsDesign/index.html
- Baron RM, Kenny DA. J Pers Soc Psychol. 1986 Dec;51(6):1173-82.
- Cheng SC, Wei LJ, Ying Z. Biometrika. 1995;82:835-45.
- Gilbert PB, et al. J Infect Dis. 2014;210 (10):1373-81.
- Kim K, et al. Biometrics. 1995 Sep;51(3):988-1000.
- Kim K, Tsiatis AA. Biometrics. 1990 Mar;46(1):81-92.
- Lachin JM, Foulkes MA. Biometrics. 1986 Sep;42(3):507-19.
- Lan KKG, DeMets DL. Stat Med. 1989;8:1191-8.
- Lan KKG, Lachin JM. Biometrics. 1990 Sep;46(3):759-70.
- Prentice RL. Stat Med. 1989;8:431–40.