"Strategies for Monitoring and Evaluation of Resource-limited National Antiretroviral Therapy Programs: The Two-phase Design"

Matthew Blake, Maria Fernandes, Deja Washington

Sebastien Haneuse, PhD,
Claudia Rivera, PhD
Department of Biostatistics

HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH
Department of Biostatistics

Background

- Malawi, Africa high number of HIV/Aids cases (one million, Dr. Harries et al)
- Implementation of National Antiretroviral Treatment (ART) Programs
- Focusing on 2005-2007 timely systems for Monitoring and Evaluation (M&E)

Background

- Antiretroviral treatment (ART) programs rely on monitoring and evaluation (M&E)
- Collected data:
  - Planning
  - Managing
  - Addressing potential problems
  - epidemiologic research.

Challenges

- Complete data (ideal)
- Comprehensive data collection is expensive
- World Health Organization (WHO) devised relying on “quarterly clinic-cohorts” aggregated data
- Aggregated data may result in ecological bias

	Children	Adults	Total
Male	3	6	9
Female	7	4	11
Total	10	10	20

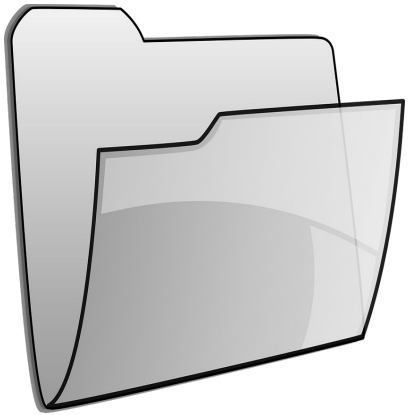
Challenges

- Complete data (ideal)
- Comprehensive data collection is expensive
- World Health Organization (WHO) devised relying on “quarterly clinic-cohorts” aggregated data
- Aggregated data may result in ecological bias

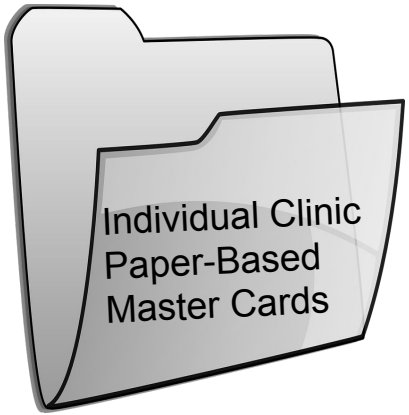
	Children	Adults	Total
Male			9
Female			11
Total	10	10	20

Dogma of Data Collection

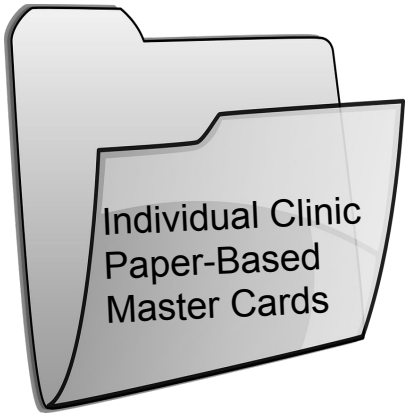
Dogma of Data Collection



Dogma of Data Collection



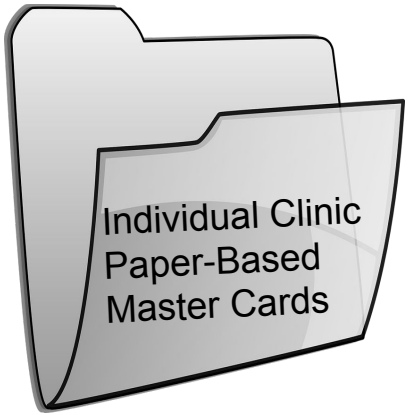
Dogma of Data Collection



Quarterly-Clinic Cohort
Data Collected

By Ministry of Health
*aggregated

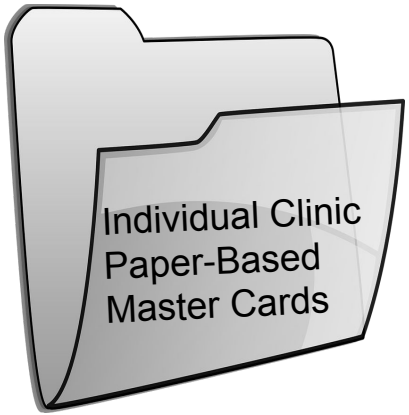
Dogma of Data Collection



Quarterly-Clinic Cohort
Data Collected

By Ministry of Health
*aggregated

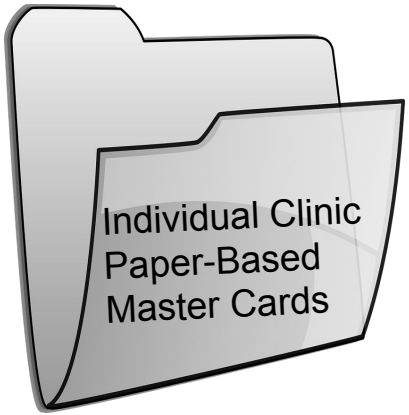
Dogma of Data Collection



Quarterly-Clinic Cohort
Data Collected

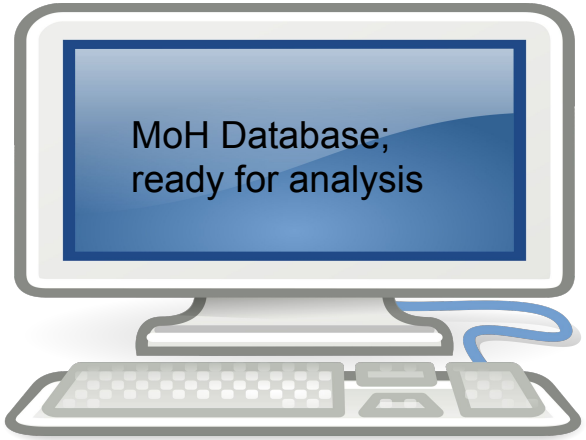
By Ministry of Health
*aggregated

Dogma of Data Collection

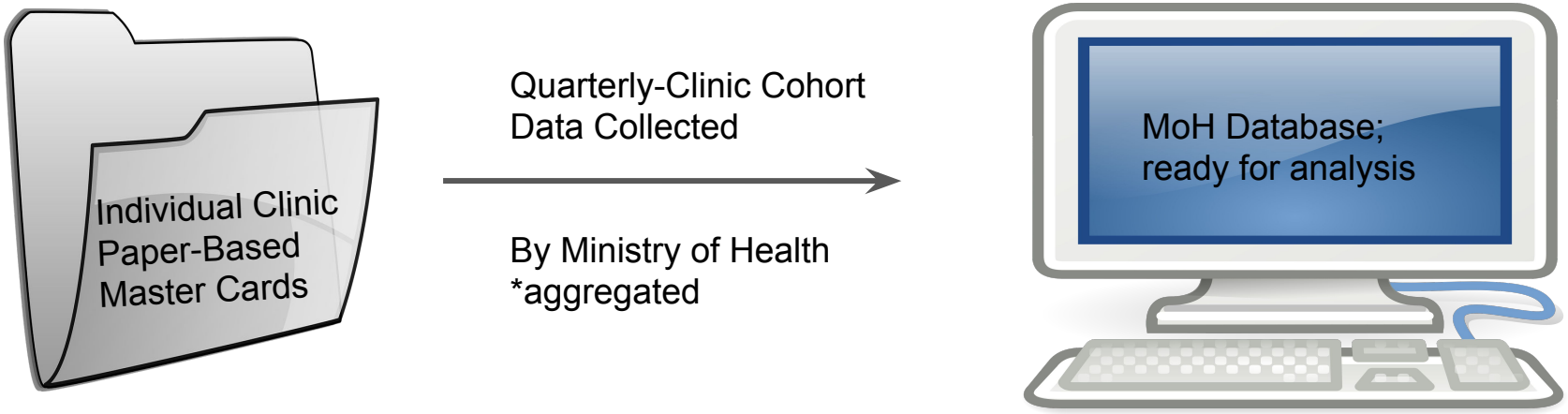


Quarterly-Clinic Cohort
Data Collected

By Ministry of Health
*aggregated



Dogma of Data Collection



For the dates 04/2008-05/2009, Malawian MoH, conducted a one-time cross-sectional survey which covered:

- Demographic characteristics (age, WHO Stage, gender)
- Treatment information (date of ART initiation and regimen)
- Clinic information (location and clinic type)

Data

- De-identification
- Binary outcomes
- Characteristics

Table 1 Patient Outcomes at six months, by patient and clinic characteristics as well as by year of registration

	Status at six months		Rate%
	Non-negative[a]	Negative[b]	
Overall	66,746	16,141	19.5
Age, years			
16-25	7,116	2,130	23.0
26-35	26,460	6,575	19.9
36-45	21,078	4,756	18.4
46-55	8,835	1,888	17.6
56-65	2,785	656	19.1
>65	472	136	22.4
Gender			
Male	25,150	7,172	22.2
Female	41,596	8,969	17.7
WHO stage			
1/2	4,418	314	6.6
3/4	62,328	15,827	20.3
Region			
Central/North	27,662	7,269	20.8
South	39,084	8,872	18.5
Registration year			
2005	12,238	3,514	22.3
2006	23,893	6,181	20.6
2007	30,615	6,446	17.4
Clinic type			
Public	64651	15839	19.7
Private	2095	302	12.6

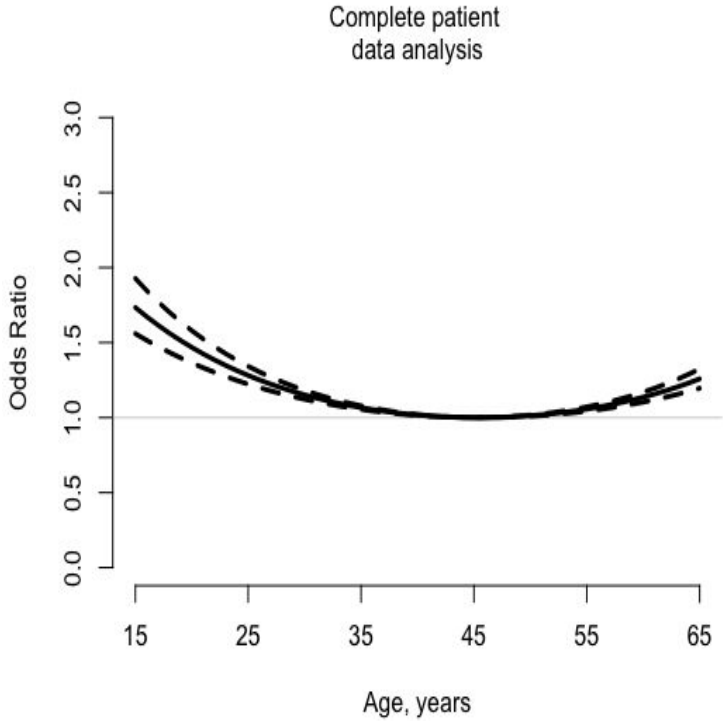
[a]Non-negative = transferred-out or alive and on-treatment.
[b]Negative = stopping treatment, lost to follow-up or death.

- “Artificially” Aggregated dataset - MoH survey data
- Individual vs. Group

	Children	Adults	Total
Male	3	6	9
Female	7	4	11
Total	10	10	20

Table 3 Characteristics of N = 82,887 patients and the corresponding N* = 1,518 quartely-clinic cohorts, from a cross-sectional survey conducted in Malawi between 04/2008-05/2009

Patients			Quartely-clinic cohorts		
	N	%		N*	%
Total	82,887		**Total**	1,518	
Age, years			**Average age, years**		
16-25	9,246	11.2	≤ 30	62	4.1
26-35	33,035	39.9	31-35	344	22.7
36-45	25,834	31.2	36-40	921	60.7
46-55	10,723	12.9	41-45	138	9.0
56-65	3,441	4.2	46-50	30	2.0
>65	608	0.7	>50	23	1.5
Gender			**Percent female**		
Male	32,322	39.0	0%	144	7.5
Female	50,565	61.0	1-40%	123	8.1
			41-50%	201	13.2
			51-60%	339	22.3
			61-99%	606	39.9
			100%	135	8.9
WHO stage			**Percent WHO stage 3/4**		
1/2	4,732	5.7	≤ 90%	255	16.8
3/4	78,155	94.3	>90%	1,263	83.2
Region			**Region**		
Central/North	34,931	42.1	Central/North	720	47.4
South	47,956	57.9	South	798	52.6
Year of registration			**Year of registration**		
2005	15,752	19.0	2005	334	22.0
2006	30,074	36.3	2006	560	36.9
2007	37,061	44.7	2007	624	41.1
Clinic type			**Clinic type**		
Public	80,490	97.1	Public	1,217	80.2
Private	2,397	2.9	Private	301	19.8



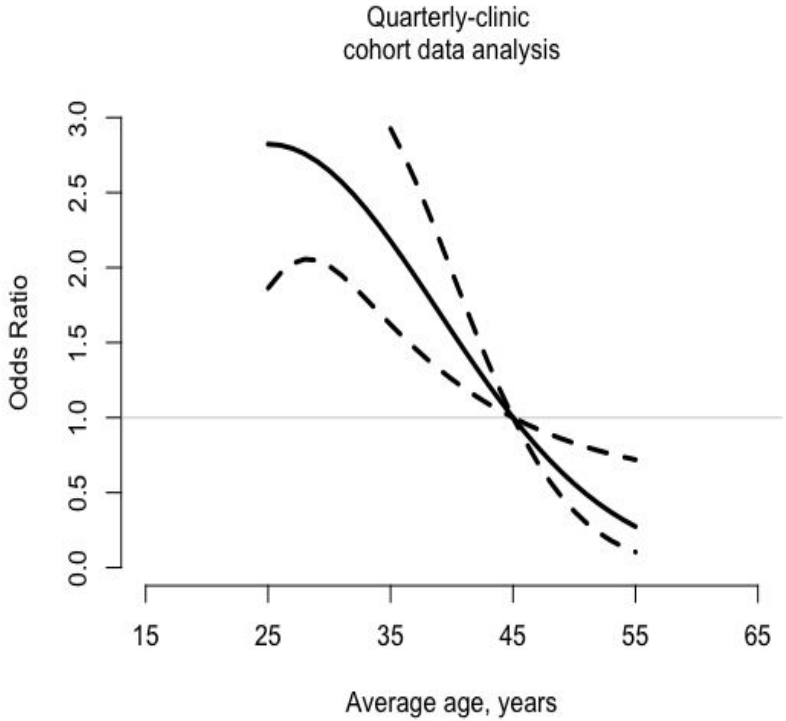


Figure 1 Results on the association between age and negative outcome status based on the complete patient data (N=82,877 patient records) and the quarterly-clinic cohort data (N*=1,518 records). Shown are odds ratio estimates and 95% confidence intervals; the referent age level for the odds ratio associations is 45 years.

Questions

- What strategies do we use to resolve ecological bias?

- Can we use the information we already have to come up with clever designs?

Two-Phase Design: Phase I

- Phase 1
  - Stratification of the entire population on the basis of outcome status and the known aggregated data
- Case-control design does not make use of the routinely collected aggregated quarterly clinic data.
- Two-phase designs used as an alternative

Six possible phase I stratification schemes that use readily-available group-level information collected by the current M&E systems in Malawi

Design #1	Private Clinic					
	No	Yes				
Non-negative status	64,651	2,095				
Negative status	15,839	302				
Design #2	Year of registration/Private Clinic					
	2005/No	2005/Yes	2006/No	2006/Yes	2007/No	2007/Yes
Non-negative status	11,991	247	22,887	1,006	29,773	842
Negative status	3,492	22	6,104	167	6,333	113
Design #3	Percent WHO stage 1 or 2					
	$\leq 5\%$	$>5\%$				
Non-negative status	50,570	16,176				
Negative status	13,191	2,950				
Design #4	Average age, years					
	≤ 35	36-40	>40			
Non-negative status	12,954	51,959	1,833			
Negative status	3,570	12,239	332			
Design #5	Percent female					
	0%	1-40%	41-50%	51-60%	61-99%	100%
Non-negative status	189	3,360	4,766	19,255	38,949	277
Negative status	20	630	1,144	5,248	9,081	18
Design #6	Percent WHO stage 1 or 2/Private clinic					
	$\leq 5\%$/No	$>5\%$/No	$\leq 5\%$Yes	$>5\%$Yes		
Non-negative status	49,160	15,491	1,410	685		
Negative status	12,796	2,863	215	87		

Six possible phase I stratification schemes that use readily-available group-level information collected by the current M&E systems in Malawi

Design #1	Private Clinic					
	No	Yes				
Non-negative status	64,651	2,095				
Negative status	15,839	302				
Design #2	Year of registration/Private Clinic					
	2005/No	2005/Yes	2006/No	2006/Yes	2007/No	2007/Yes
Non-negative status	11,991	247	22,887	1,006	29,773	842
Negative status	3,492	22	6,104	167	6,333	113
Design #3	Percent WHO stage 1 or 2					
	$\leq 5\%$	$>5\%$				
Non-negative status	50,570	16,176				
Negative status	13,191	2,950				
Design #4	Average age, years					
	≤ 35	36-40	>40			
Non-negative status	12,954	51,959	1,833			
Negative status	3,570	12,239	332			
Design #5	Percent female					
	0%	1-40%	41-50%	51-60%	61-99%	100%
Non-negative status	189	3,360	4,766	19,255	38,949	277
Negative status	20	630	1,144	5,248	9,081	18
Design #6	Percent WHO stage 1 or 2/Private clinic					
	$\leq 5\%$/No	$>5\%$/No	$\leq 5\%$Yes	$>5\%$Yes		
Non-negative status	49,160	15,491	1,410	685		
Negative status	12,796	2,863	215	87		

Phase II

Six possible phase I stratification schemes that use readily-available group-level information collected by the current M&E systems in Malawi

Design #1	Private Clinic					
	No	Yes				
Non-negative status	64,651	2,095				
Negative status	15,839	302				
Design #2	Year of registration/Private Clinic					
	2005/No	2005/Yes	2006/No	2006/Yes	2007/No	2007/Yes
Non-negative status	11,991	247	22,887	1,006	29,773	842
Negative status	3,492	22	6,104	167	6,333	113
Design #3	Percent WHO stage 1 or 2					
	$\leq 5\%$	$>5\%$				
Non-negative status	50,570	16,176				
Negative status	13,191	2,950				
Design #4	Average age, years					
	≤ 35	36-40	>40			
Non-negative status	12,954	51,959	1,833			
Negative status	3,570	12,239	332			
Design #5	Percent female					
	0%	1-40%	41-50%	51-60%	61-99%	100%
Non-negative status	189	3,360	4,766	19,255	38,949	277
Negative status	20	630	1,144	5,248	9,081	18
Design #6	Percent WHO stage 1 or 2/Private clinic					
	$\leq 5\%$/No	$>5\%$/No	$\leq 5\%$Yes	$>5\%$Yes		
Non-negative status	49,160	15,491	1,410	685		
Negative status	12,796	2,863	215	87		

Six possible phase I stratification schemes that use readily-available group-level information collected by the current M&E systems in Malawi

Design #1	Private Clinic					
	No	Yes				
Non-negative status	64,651	2,095				
Negative status	15,839	302				
Design #2	Year of registration/Private Clinic					
	2005/No	2005/Yes	2006/No	2006/Yes	2007/No	2007/Yes
Non-negative status	11,991	247	22,887	1,006	29,773	842
Negative status	3,492	22	6,104	167	6,333	113
Design #3	Percent WHO stage 1 or 2					
	$\leq 5\%$	$>5\%$				
Non-negative status	50,570	16,176				
Negative status	13,191	2,950				
Design #4	Average age, years					
	≤ 35	36-40	>40			
Non-negative status	12,954	51,959	1,833			
Negative status	3,570	12,239	332			
Design #5	Percent female					
	0%	1-40%	41-50%	51-60%	61-99%	100%
Non-negative status	189	3,360	4,766	19,255	38,949	277
Negative status	20	630	1,144	5,248	9,081	18
Design #6	Percent WHO stage 1 or 2/Private clinic					
	$\leq 5\%$/No	$>5\%$/No	$\leq 5\%$Yes	$>5\%$Yes		
Non-negative status	49,160	15,491	1,410	685		
Negative status	12,796	2,863	215	87		

Phase II

- Sub-samples from each of the phase I strata
- The number of patients fixed and their resources allocated across the phase I data.

Table 4

	Complete patient data (N=82,887)		Aggregated quartely-clinic data (N=1,518)		Case–Control Study I (N=5,000)		Two-phase Study (N=5,000)	
	OR	95 % CI	OR	95 % CI	OR	95 % CI	OR	95 % CI
Year of registration								
2005	REF		REF		REF		REF	
2006	0.91	(0.87, 0.95)	0.89	(0.82, 0.96)	0.96	(0.82, 1.13)	0.83	(0.68, 1.01)
2007	0.76	(0.73, 0.80)	0.74	(0.69, 0.80)	0.81	(0.69, 0.94)	0.75	(0.62, 0.91)
Clinic type								
Public	REF		REF		REF		REF	
Private	0.31	(0.19, 0.47)	0.33	(0.15, 0.64)	0.29	(0.08, 0.85)	0.31	(0.19, 0.50)
Clinic/year interaction								
Private/2006	2.16	(1.38, 3.55)	2.03	(1.00, 4.66)	2.26	(0.68, 8.93)	2.18	(1.30, 3.67)
Private/2007	2.05	(1.29, 3.39)	1.98	(0.95, 4.61)	2.32	(0.68, 9.40)	2.07	(1.21, 3.52)

Table 4

	Complete patient data (N=82,887)		Aggregated quartely-clinic data (N=1,518)		Case–Control Study I (N=5,000)		Two-phase Study (N=5,000)	
	OR	95 % CI	OR	95 % CI	OR	95 % CI	OR	95 % CI
Year of registration								
2005	REF		REF		REF		REF	
2006	0.91	(0.87, 0.95)	0.89	(0.82, 0.96)	0.96	(0.82, 1.13)	0.83	(0.68, 1.01)
2007	0.76	(0.73, 0.80)	0.74	(0.69, 0.80)	0.81	(0.69, 0.94)	0.75	(0.62, 0.91)
Clinic type								
Public	REF		REF		REF		REF	
Private	0.31	(0.19, 0.47)	0.33	(0.15, 0.64)	0.29	(0.08, 0.85)	0.31	(0.19, 0.50)
Clinic/year interaction								
Private/2006	2.16	(1.38, 3.55)	2.03	(1.00, 4.66)	2.26	(0.68, 8.93)	2.18	(1.30, 3.67)
Private/2007	2.05	(1.29, 3.39)	1.98	(0.95, 4.61)	2.32	(0.68, 9.40)	2.07	(1.21, 3.52)

Table 4

	Complete patient data (N=82,887)		Aggregated quartely-clinic data (N=1,518)		Case–Control Study I (N=5,000)		Two-phase Study (N=5,000)	
	OR	95 % CI	OR	95 % CI	OR	95 % CI	OR	95 % CI
Year of registration								
2005	REF		REF		REF		REF	
2006	0.91	(0.87, 0.95)	0.89	(0.82, 0.96)	0.96	(0.82, 1.13)	0.83	(0.68, 1.01)
2007	0.76	(0.73, 0.80)	0.74	(0.69, 0.80)	0.81	(0.69, 0.94)	0.75	(0.62, 0.91)
Clinic type								
Public	REF		REF		REF		REF	
Private	0.31	(0.19, 0.47)	0.33	(0.15, 0.64)	0.29	(0.08, 0.85)	0.31	(0.19, 0.50)
Clinic/year interaction								
Private/2006	2.16	(1.38, 3.55)	2.03	(1.00, 4.66)	2.26	(0.68, 8.93)	2.18	(1.30, 3.67)
Private/2007	2.05	(1.29, 3.39)	1.98	(0.95, 4.61)	2.32	(0.68, 9.40)	2.07	(1.21, 3.52)

Benefits of Two-Phase Design

- Uses aggregated data and sub-samples of patient-level data
- Lesser degrees of uncertainty (confidence intervals)
- N = 5,000 ⟶ 80% power
- Cost-efficient

Shortcomings

- Trade-Off:
  - Standard Error estimates of covariates increases by 20% compared to complete patient-level data - imbalancing
  - Interpretation of aggregated (group-level) data
  - Forces a balance

	Complete patient data (N=82,887)		Aggregated quartely-clinic data (N=1,518)		Case–Control Study I (N=5,000)		Two-phase Study (N=5,000)	
	OR	**95 % CI**	**OR**	**95 % CI**	**OR**	**95 % CI**	**OR**	**95 % CI**
Age[a]								
Linear	0.96	(0.94, 0.98)	0.35	(0.19, 0.66)	1.00	(0.99, 1.00)	0.94	(0.87, 1.01)
Quadratic	1.06	(1.05, 1.07)	0.77	(0.53, 1.08)	1.05	(1.01, 1.09)	1.04	(0.99, 1.08)
Clinic type								
Public	REF		REF		REF		REF	
Private	0.31	(0.19, 0.47)	0.33	(0.15, 0.64)	0.29	(0.08, 0.85)	0.31	(0.19, 0.50)

[a]Quartely-clinic cohort data model uses average age(in years).

Conclusion and Future Work

- There is a need for innovative strategies that are robust to ecological bias and that bypass the financial impasse of M&E of patient-level outcomes
- Two-Phase Design addresses the relationship between the patient's outcome and a particular variable (i.e. clinic type: public/private)
- Two-Phase design potentially useful at the national and local scale, but isn't reliable in all situations

Acknowledgements

Faculty/Research Advisor

Dr. Sebastien Haneuse

Postdoctoral Student

Dr. Claudia Rivera

Director of Biostatistics and Computational Biology Summer Program

Dr. Rebecca Betensky

Senior Project Coordinator

Jessica Boyle

Instructors

Heather Mattie
Olivia Orta
Sarah Anoke

Peers

Thank You!