Shortened Lifespan:
A Legacy of Exposure to Malaria Risk in Early Life

Sok Chul Hong
Seoul National University
Sokchul.hong@snu.ac.kr

SJE International Symposium
Human Capital and Economic Development

June 22, 2016

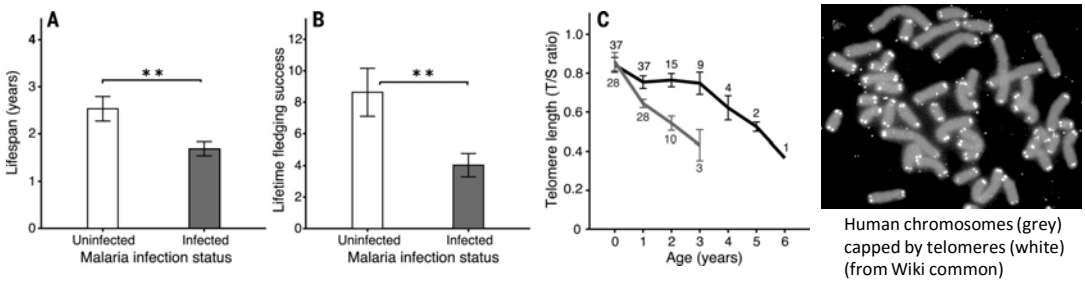
Significance of Malaria

- Significance of malaria:
  - Malaria is the world's important parasitic infectious disease, which is transmitted by mosquitoes.
  - In 2015, an estimated 214 million cases of malaria occurred worldwide and 438,000 people died, mostly children in the African region (WHO).
  - 3.2 billion people live in areas at risk of malaria transmission in 106 countries and territories (CDC).
- Short- and Long-term impacts of malaria:
  - High mortality, severe anemia, brain damage, ARDS, kidney failure, and so on.
  - Loss of school days and working days.
  - Infections during pregnancy can increase the probability of having LBW babies.
  - Infections in early life can deteriorate cognitive ability in growing year, lower labor productivity and income in adulthood, and accelerate the onset of chronic conditions in old ages [Barreca 2010; Bleakley 2010; Chang et al. 2011; Cutler et al. 2010; Hong 2007, 2011, 2013; Lucas 2010].
  - Those studies have well supported the priority of global efforts to fight malaria.

Malaria and Lifespan

- Unexplored question:
  - Can malaria especially in early life shorten lifespan among survivors?
  - Lifespan is a longer-term outcome. Thus, the question will provide new insight for measuring the potential benefit of malaria eradication.
- Malaria and lifespan:
  - Existing studies---which examined the long-term impact of early-life malaria infections on later outcomes---strongly suggest that it can shorten lifespan through indirect mechanisms.
  - However, a recent scientific study on wild birds suggests that malaria can shorten lifespan due to a (relatively direct) biological mechanism.
  - Asghar et al. (2015), "Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds," *Science* 347: 436-438.

Great Reed Warbler
(from Wikipedia)



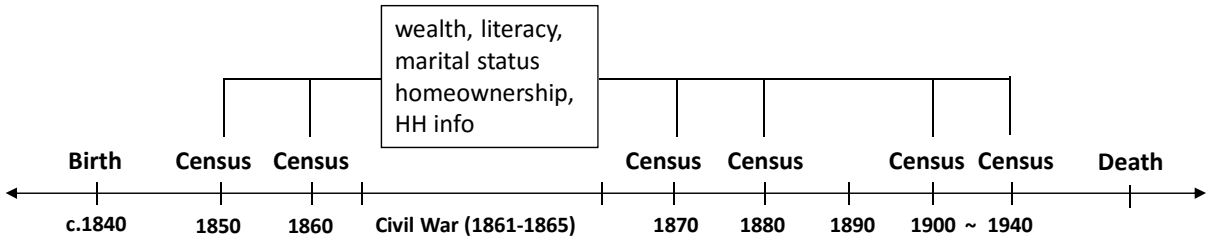
Human chromosomes (grey)
capped by telomeres (white)
(from Wiki common)

A Historical Study

- Limitation of modern studies:
  - Studies on lifespan require data containing lifetime variables.
  - Various confounding factors would obstruct the identification of the effect of malaria in early life.
- Malaria in mid-19th century America:
  - Malarial fevers had been very prevalent across most areas of the United States throughout the mid-19th century. [Malaria death rate per 100,000: Southern US in 1850 = 111, Top 10 malarial countries in 2012 = 113.5]
  - Its transmission via mosquitoes was not revealed until 1898.
  - Eradication: 1880s in the mid-West, 1920s in the South, 1940s using DDT.
  - This implies that those in the areas with climate suitable for mosquitoes were exposed to high risk of malarial fevers without effective prevention and treatment.
- Union Army sample:
  - This study uses the sample of Union Army soldiers who attended the American Civil War (1861-1865) with lifetime variables.
  - Many of them were exposed to high risk of malarial fevers in early life.

UA Dataset

- UA dataset:
  - Number of observations: 39,517 from 351 randomly selected companies.
  - Lifetime variables from various historical records, including Complied Military Service Records (CMSR), Pension Records (PEN), Carded Medical Records (CMR), Surgeons' Certificates (SCRTS), and US Federal Censuses in 1850-1940 (CEN).



Wartime Data

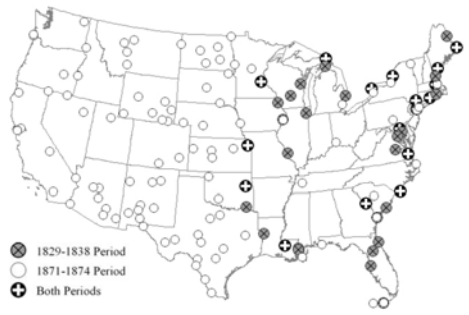
CMSR: birth year & place, enlistment, battle, rank, POW
CMR: wartime diseases & injuries

Post-War Data

PEN: death year, place & cause
SCRTS: physical exam, onset of chronic diseases

Measure of Malaria Risk

- Measuring the exposure to malaria risk:
  - No information on early-life infections. A variable measured at a certain point of time may not capture chronic infections and prevailing risk of malarial fevers.
  - Instead, I estimate the county-level annual probability of contracting malarial fevers, using malaria incidence rate in the US Army forts (*M*) and risk factors such as temperature, rainfall, and elevation (*E*).
  - Hong (2007, *J Econ History*):



APPENDIX FIGURE 1
LOCATIONS OF U.S. ARMY FORTS IN THE STUDY

corr(*M*, *E*)

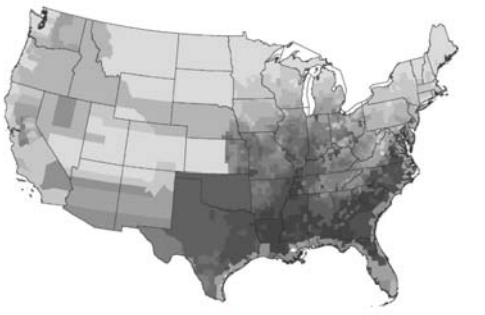


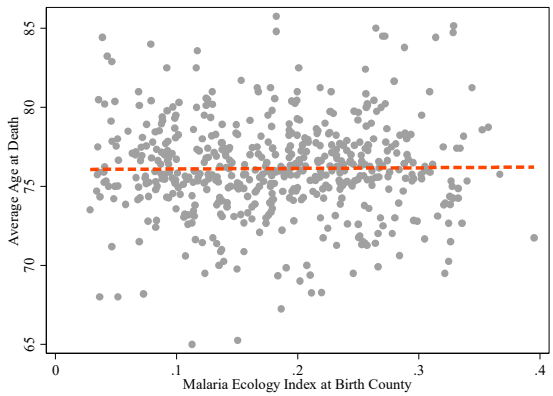
Fig. A1. Estimated pre-eradication malaria ecology. *Note*: The risk index by 1860 county boundary. Darker areas are more malarial.

Variable of Lifespan

- Definition of early life:
  - I estimate the malaria ecology index at birth, in 1850 (age 10 on average), in 1860 (age 20), and in 1900 (age 60).
  - The term of early life denotes the years of age from birth up to age 10.
  - The in-utero period is also contained into the range of early life
  - Dataset does not allow to test the significance for other ages b/w birth and 10.
- Lifespan = Death Year – Birth Year:
  - 14.8% died in the War. It is thought that those deaths were less associated with malaria infections in early life. So I exclude wartime deaths for the analyses.
  - On the other hand, most information on death year is obtained from Pension Records. Its availability is related with the change of Civil War Pension law.
  - Prior to 1890 (The Dependent and Disability Pension Act), the pension was given to only those who could prove time spent in the military and had a disability incurred while in service.
  - If veterans died before 1890, their death years were not reported in general. [26.5% out of native-born war survivors]
  - To avoid potential selection biases, I examine only those who survived up to 1900.

Baseline Sample and Correlation

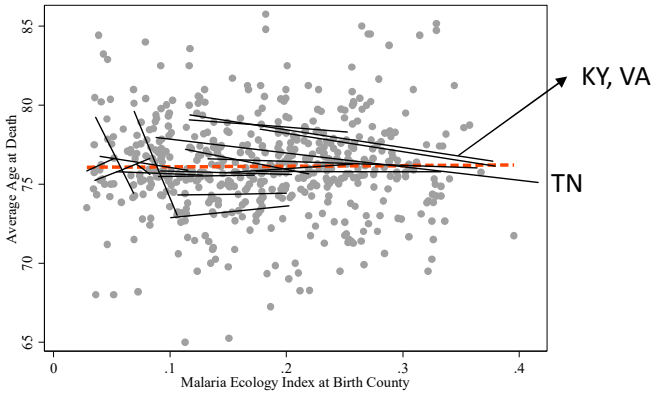
- Sample for Baseline Analyses:
  - 11,006 native-born veterans who survived up to 1900 and so could apply for the pension.
  - I release this constraint later.
- Correlation b/w lifespan and malaria ecology index at birth county:



  - Scatter plot for 917 groups born in the same county.
  - The dashed line suggests that lifespan was not related with malaria risk at birth.

Baseline Sample and Correlation

- Sample for Baseline Analyses:
  - 11,006 native-born veterans who survived up to 1900 and so could apply for the pension.
  - I release this constraint later.
- Correlation b/w lifespan and malaria ecology index at birth county:



  - Within states of birth, negative correlations are observed.
  - This strongly suggests the use of state-of-birth fixed effects.

Identification Strategy: Baseline Estimation

- Baseline estimation:
  - $L_{ics} = \alpha + \beta M_c + \delta_s + Z_c\Gamma + \sum_t X_{i,t}\Pi_t + \varepsilon_{ics}$
  - L_{ics}: veteran i who was born in county c at state s.
  - M_c: malaria ecology index at the county of birth
    estimated with 1825-1859 average climate variables / normalized for simple interpretation.
  - δ_s: state-of-birth FE
  - Z_c: early-life local confounding factors measured as of 1850.
    crude death rate, cause-specific mortality rate (9 major diseases, using the 1850 mortality census manuscripts), infant mortality rate, population density, fertility rate, school enrollment rate, average farm value, % of slave populations, type of transportation at the county of birth.
  - $X_{i,t}$: individual characteristics at various points of lifetime
    - Birth year.
    - Wartime experience: year of enlistment, diseases, injuries, rank, POW, company mortality rate.
    - Onset of chronic diseases by 1899: for 18 types of diseases.
  - In later analyses, I utilize the information in 1850, 1860 and 1900 census records to control for:
    - Malaria ecology index in 1900, 1860, and 1850, estimated using decadal climate variables.
    - SES c.1900: marital status, literacy, home ownership;
    - SES c.1860 and c.1850: HH wealth, HH size.

Identification Strategy: Selection and Endogeneity

- Selection issue:
  - The pathogen of malaria was not known in the mid-19th century, and so people (veterans' parents) seems to have selected counties without knowing the risk of malaria risk.
  - In addition, the malaria ecology index is based on exogenous environmental factors. This is a kind of instrumental variable.
  - However, people might partially recognize the risk of infectious diseases in some ways. This can cause a problem of selection.
  - In addition, omitted variables that might be correlated with malaria risk can cause another endogeneity problem.
- County-of-birth FE estimation:
  - Using malaria ecology index estimated with annual climate variables at the year of birth / very strict specification.
- IV estimation:
  - Instrumental variable: ratio of birth county's adult population from father's birth state or country.
  - People in the early-19th century much relied on their friends and ethnic groups to get information on places to migrate [Costa and Kahn 2006].

Effects of Early-Life Exposure to Malaria Risk on Lifespan

Table 1. Estimated Effect of Early-Life Exposure to Malaria on Age at Death

Dependent variable: Age at death

	Baseline estimation					County-of-birth FE	IV estimation		OLS with the sample used for IV estimation
				Baseline 1	Baseline 2		First (Malaria)	Second	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Malaria ecology at the county of birth	0.1877	-0.8426***	-0.9250***	-0.8641***	-0.8373**	-0.7977*		-0.9062	-0.8796**
	(0.1384)	(0.3032)	(0.3390)	(0.3324)	(0.3319)	(0.4170)		(3.4991)	(0.4224)
Ratio of birth county's adult populations from father's birth state or country							0.4173***		
							(0.0307)		
Endogeneity Test (p-value)								0.9939	
Controls:									
Year of birth	Y	Y	Y	Y	Y	Y	Y	Y	Y
State-of-birth FE		Y	Y	Y	Y		Y	Y	Y
Early-life county conditions			Y	Y	Y		Y	Y	Y
Wartime experience				Y	Y	Y	Y	Y	Y
Dummy of chronic conditions by 1899					Y	Y	Y	Y	Y
County-of-birth FE						Y			
Observations	11006	11006	11006	11006	11006	10927	6695	6695	6695

Note : Standard errors are clustered on birth county. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

- The effect of early-life exposure to malaria risk on lifespan is estimated significant across various specifications.
- The result is robust even if the onset of chronic conditions are controlled for.
- Magnitude: Born in a county with 1-SD higher malaria risk ➔ 0.84-year lower lifespan.
- Note that the baseline estimation uses veterans who survived up to 1900 or age 60.
- The problems of omitted variables and selection seem to be little substantial.

Late-Life Exposure & Role of SES

Table 2. Significance of the Exposure to Malaria in Early Life
Dependent variable: Age at death

Sample	1900 links			1860-1900 links				1850-1860-1900 links	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Malaria ecology:									
at the county of birth	-0.8986**	-0.9116**	-1.0241**	-1.2491**	-1.2374**	-1.3453**	-1.4099**	-1.4211**	-1.9568**
	(0.3988)	(0.3987)	(0.4065)	(0.5716)	(0.5724)	(0.6051)	(0.6048)	(0.6845)	(0.9518)
at the county in 1900			0.2017				0.1977		0.2056
			(0.1981)				(0.3135)		(0.3810)
at the county in 1860						0.1883	0.1082		0.5162
						(0.3267)	(0.3588)		(0.5020)
at the county in 1850									0.1228
									(0.8234)
Baseline 2's controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Additional controls:									
SES in 1900		Y	Y	Y	Y	Y	Y	Y	Y
SES in 1860					Y	Y	Y	Y	Y
SES in 1850									Y
Observations	7592	7592	7592	3894	3894	3894	3894	2621	2621

Note : Standard errors are clustered on birth county. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

- I utilize the 1900, 1860 and 1850 census links to see the effects of exposure to malaria risk and SES at age 10, age 20 and age 60. The sample size decreases as additional links are utilized.
- The effect to exposure to malaria risk after age 10 is estimated insignificant.
- The results are robust even if SES variables are controlled for.
- Why is the effect in (8) and (9) larger? Selection.

Larger Effects in Younger Ages

- In (a), I run the baseline regressions for those who survived by different years in 1866-1925. In (b), I run the Cox proportional hazard estimations. In both, I plot the estimated coefficients and 90% confidence intervals.
- The negative impact was more substantial in younger ages. This can be much larger because many didn't enroll in pension system before 1890.
- Then, the size of effect declines as affected veterans died off. The significant effect is observed up to 1910 or age 70.
- Hazard model provides the same implication.

Larger Effects in Younger Ages

Table 3. Early Exposure to Malaria Risk and Cause of Death

Depedent variables: Dummy of specific cause of death

Sample	All veterans		War survivors		Those who survived by 1900		
	Mean	Baseline 1	Mean	Baseline 2	Mean	Baseline 1	Baseline 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cause of death:							
Infectious diseases	0.174	0.0146	0.127	0.0234	0.077	0.0193	0.0172
		(0.0163)		(0.0159)		(0.0148)	(0.0148)
Heart diseases	0.170	-0.0084	0.212	-0.0104	0.242	-0.0208	-0.0222
		(0.0153)		(0.0190)		(0.0234)	(0.0237)
Senility	0.147	-0.0432***	0.126	-0.0463***	0.120	-0.0507***	-0.0520***
		(0.0160)		(0.0163)		(0.0192)	(0.0191)
Respiratory diseases	0.121	0.0563***	0.128	0.0732***	0.113	0.0507***	0.0516***
		(0.0135)		(0.0154)		(0.0175)	(0.0175)
Accident or external causes	0.093	0.0114	0.035	-0.0003	0.022	0.0023	0.0023
		(0.0131)		(0.0075)		(0.0076)	(0.0076)
Genito-urinary diseases	0.090	-0.0023	0.110	-0.0000	0.127	0.0059	0.0092
		(0.0129)		(0.0155)		(0.0197)	(0.0199)
Diseases of blood-forming organs	0.068	-0.0034	0.084	-0.0058	0.097	-0.0059	-0.0047
		(0.0101)		(0.0127)		(0.0165)	(0.0166)
Nervous system	0.054	-0.0091	0.063	-0.0089	0.066	-0.0043	-0.0046
		(0.0091)		(0.0111)		(0.0137)	(0.0137)
Intestinal diseases	0.052	0.0004	0.062	0.0053	0.058	0.0108	0.0102
		(0.0098)		(0.0116)		(0.0132)	(0.0132)
Others	0.117	-0.0139	0.139	-0.0185	0.158	-0.0023	-0.0013
		(0.0134)		(0.0153)		(0.0201)	(0.0202)
Observations	9583		7673		5393		

Note : Standard errors are clustered on birth county. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

- People born in higher malaria-risk counties less likely died of 'old age', i.e. senility.
- They more likely died of respiratory diseases.

Mechanism & Implication

- Key finding:
  - UA veterans born in higher malarial counties died earlier than otherwise.
- Mechanism:
  - Indirect channels: early exposure ➔ deteriorate human capital accumulation and lifetime health ➔ shorten lifespan.
  - The estimation in this study is robust even if lifetime health and SES variables are included as controls. This may suggest that indirect channels can be less significant. Of course, the specification is not enough to capture the channels well.
  - Biological channels seem more plausible as suggested in Asghar et al. (2015, *Science*).
- Implication:
  - Global health aspect: The potential benefit of malaria eradication would be larger than discussed so far.
  - Economic history aspect: This study well accounts for why people in the past died earlier.