Socioeconomic risk factors for adult mortality in a middle-income country: The Costa Rican Longitudinal Mortality Study 1984-2007

William H. Dow, University of California, Berkeley
Luis Rosero-Bixby, University of Costa Rica

September 2009

Abstract
The Costa Rican Longitudinal Mortality Study (CR-LMS) is a unique new database that prospectively tracks 24-year mortality in a national sample of approximately 20,000 Costa Ricans who were aged 30 and over in the 1984 census. About 5,000 deaths had occurred by December 2008, mirroring death rates from national life tables. Numerous checks suggest that the death registry is extremely high quality; for example, the cohort that started over age 85 is now extinct, an indication that the death registry is complete. We estimate socioeconomic and demographic gradients in all-cause and cardiovascular adult mortality hazard regressions, something rarely accomplished in a developing country population. The data indicate that SES gradients are weak throughout this 24-year period, even among the younger adults. College education is protective but rare, while completing primary or secondary education carries little mortality advantage. Other characteristics such as region and birth month also yield intriguing results.

Introduction and Background
Little is known about adult mortality in developing countries, either levels or risk factors, because of the lack of appropriate data. Registration of deaths in vital statistics systems is often incomplete, and biased by socio-demographic differentials in registration. Population data, providing the denominator for computing rates, is also problematic. Particularly at very old ages, in which a cross-over with the mortality rates of developed countries often appears in the data, measurement problems such as age exaggeration may swamp true patterns. Vital statistics databases also usually do not keep records of the socioeconomic status (SES) of the deceased; when they do, the information is often useless because of SES changes just before death or because of the lack of compatible population information for computing rates.

Published evidence suggests a strong SES gradient in mortality across many populations, with low-SES individuals having the highest death rates. In developing countries, however, this evidence is confined almost exclusively to children, and primarily relies on studies using the indirect methods developed by William Brass or using retrospective birth histories from demographic and health surveys. Almost all of the evidence for adult SES gradients comes from developed countries. The few studies based on data from developing countries suggest that SES gradients in adult (mostly old-adult) mortality are weaker than among children (Crimmins 2005). Moreover, earlier studies of small samples of elderly Costa Ricans suggest that the SES gradients in this country are weak if any (Rosero-Bixby & Dow 2009, Rosero-Bixby, Dow & Lacle 2004).

Although Costa Rica is a middle-income country, with per capita income and health expenditures about one-fifth and one-tenth the levels of the United States, its life expectancy is equivalent to many developed countries and higher than that of the USA. Moreover, elderly male Costa Ricans seems to have one of the highest life expectancies in the world (Rosero-Bixby 2008). To further explore this apparently exceptional longevity and bring some light to the study of adult mortality in developing countries, we have recently completed the Costa Rican Longitudinal Mortality Study (CR-LMS), which prospectively tracks mortality events in a national sample of approximately 20,000 Costa Ricans drawn from the 1984 census. These are among the first comprehensive data of their type from a setting outside high-income countries.

In this paper we first detail the methodology by which we link individual Census records to the Costa Rican National Death Registry and validate our ability to comprehensively identify mortality events and thus accurately estimate mortality rates. Next we show that our data closely track aggregate mortality patterns from previously analyzed vital statistics data, further confirming Costa Rica's remarkably low mortality rates.

We then estimate hazard regression models to characterize socioeconomic patterns of mortality, prospectively analyzing mortality in relation to 1984 census SES characteristics in addition to well-documented demographic characteristics such as exact age, month of birth and whether the individual's birth was timely registered.

Preliminary results

From an original clustered, multistage sample of 21,100 Costa Ricans aged 30 and over in the 1984 census, taken with probabilistic procedures, 87% were satisfactorily linked to the national population registry. Each person's identity was fully established with the unique identification number that all Costa Ricans have from the moment their birth is recorded in uniquely and sequentially numbered ledgers (from which the identification number is taken). In some groups, such as the oldest old, illiterate, and some rural areas, the rate of successful linkage was lower, which necessitated introducing sample weights that correct for linkage success rates. The weighted sample reproduces census distributions by age, sex, education, and urbanicity. The sample weights ranged from 0.55 to 2.55, with mean 1.0 and SD of 0.14.

The 18,400 Costa Ricans in the sample were followed until December 2008 in the computer files of the civil registration system and 5,076 deaths were identified. For 419 individuals follow-up was lost when they failed to appear in the voting lists for the February 2006 presidential election. These losses were treated as censored observations with a randomly assigned date between their last appearance in the voting list and the date of the first voting list they disappeared. Table 1 shows a description of the sample by age at the beginning of the observation. The sample rendered 373,000 person-years of observation. No person aged 85 and over in 1984 is alive at the end of the observation. The oldest individuals alive in the sample in December 2008 are one 104 year old and two 102 year olds. The existence of these three individuals was verified in the field. The extinction of the oldest cohorts and the absence of super-centenarians at the end of the observation period suggests that the civil registry of deaths is highly complete—itself a unique feat for a middle-income country.

To establish whether this sample is representative and properly reproduces the mortality levels of the population, Figure 1 compares the observed mortality rates by five-year age brackets and robust 95% confidence intervals (that include clustering design effects) with two series of rates from official life tables that approximately correspond to the beginning and end of the observation period (1985-90 and 2000-05). The sample rates do not differ significantly from those from the life tables. Although mortality rate estimates are noisy for younger adults, the sample confidence intervals at younger ages always overlap the 1985-89 population rates, and the confidence intervals for older ages overlap the most recent life table.

Figure 1 also shows that a Gompertz function (a line for the rate's logarithms) fits well the series of age specific death rates, especially after age 45. At younger ages the observed rates tend to be above the Gompertz line. This is partially caused by the fact that mortality at these ages correspond mostly to an earlier period when mortality was higher (for example, all of the morality under 40 is for the period 1984-94, since after 1994 all observed individuals are older than 40 years).

Preliminary results of SES gradients in mortality, estimated with hazard regression models, show generally minor SES effects (table 2). College educated individuals have lower mortality, but completion of primary or secondary school has little protective effect either for all-cause or for cardiovascular mortality. This lack of education gradient is similar in both older and younger adults (not shown), in contrast to developed country population gradients which tend to be larger at younger ages. The gradient is also consistent throughout the 24 year follow-up period, suggesting that the lack of gradient is not just a short-run phenomenon due to factors such as SES differences in nutritional transition phase.

A variety of other intriguing results merit further exploration as well. Those living in rural areas in 1984 have lower mortality, as do residents of the Chorotega region which has been previously identified as a high longevity area. Health insurance associations indicate adverse selection of sicker individuals into health insurance coverage. Males have, as expected, higher mortality risks. There are no prominent occupational effects, and higher wealth is associated with lower all-cause mortality after controlling for education. Intriguingly, those born in March and April show 17% higher cardiovascular mortality. Overall, this unique new database provides insight into a wide array of relationships about which our understanding has previously been largely confined to speculation.

Table 1. Persons, exposure and deaths in the CR-LMS sample

Age in 1984	Persons	Years exposure	Deaths 1984-2005	Lost	Alive Dec. 2008
30-34	3,671	81,842	231	110	3,330
35-44	5,588	124,693	573	83	4,932
45-54	3,907	83,132	873	77	2,957
55-64	2,736	51,681	1,279	65	1,392
65-74	1,605	23,475	1,264	50	291
75-84	749	7,295	708	25	16
85+	151	871	148	3	0
Total	18,407	372,990	5,076	413	12,918

Table 2. Relative death rates by SES characteristics, estimated with hazard regression models

Predictor	All causes of death		Cardiovascular deaths	
	Relative rate	(95% CI)	Relative rate	(95% CI)
March-April born	1.06	(0.98 - 1.14)	1.17	(1.04 - 1.32)*
Gender				
Non working women	1.00	Ref.	1.00	Ref
Working women	1.05	(0.94 - 1.18)	0.87	(0.72 - 1.05)
Male	2.04	(1.64 - 2.53)*	1.71	(1.12 - 2.60)*
Occupation (high in household)				
Professional, managerial	0.94	(0.87 - 1.02)	1.04	(0.90 - 1.19)
Other	1.00	Ref	1.00	Ref
Agriculture, extractive	1.02	(0.91 - 1.15)	0.97	(0.80 - 1.17)
Education				
None	1.01	(0.93 - 1.09)	0.96	(0.84 - 1.10)
Incomplete elementary	1.00	Ref	1.00	Ref
Complete elementary	0.97	(0.89 - 1.05)	0.95	(0.82 - 1.10)
High school drop out	0.85	(0.74 - 0.96)*	0.81	(0.64 - 1.04)
High school diploma	1.01	(0.88 - 1.17)	1.10	(0.87 - 1.38)
College	0.77	(0.67 - 0.88)*	0.68	(0.51 - 0.90)*
Wealth index (1-14)	0.98	(0.97 - 1.00)*	0.99	(0.97 - 1.02)
Health insurance				
Direct from job	1.00	Ref	1.00	Ref
Paid voluntarily	0.97	(0.87 - 1.08)	0.97	(0.79 - 1.19)
From Relative's job	1.02	(0.92 - 1.12)	0.95	(0.80 - 1.12)
From government	1.21	(1.10 - 1.33)*	1.03	(0.87 - 1.21)
None	0.97	(0.88 - 1.06)	0.80	(0.68 - 0.93)*
Rural (reference urban)	0.87	(0.80 - 0.95)*	0.88	(0.76 - 1.02)
Region				
Metropolitan San Jose	1.00	Ref	1.00	Ref
Other Central	1.01	(0.94 - 1.08)	1.11	(0.98 - 1.27)
Chorotega	0.81	(0.70 - 0.95)*	0.67	(0.53 - 0.86)*
Central Pacific	0.85	(0.71 - 1.02)	0.83	(0.57 - 1.19)
Brunca	0.83	(0.73 - 0.94)*	0.90	(0.70 - 1.16)
Huetar North	1.02	(0.89 - 1.16)	1.31	(1.07 - 1.60)
Caribbean	0.79	(0.67 - 0.94)*	0.98	(0.74 - 1.30)*

Hazard regressions also control for time trend, linkage quality, and timely birth registration. Models include interaction terms (not shown) of sex with time and with age (reported sex effects are for people age 30 in year 1984).

Figure 1. Age specific death rates compared to two Costa Rican life tables