The relationship between non-communicable disease occurrence and poverty—evidence from demographic surveillance in Matlab, Bangladesh

AQ1

Andrew J. Mirelman, [1,2,]* Sherri Rose,[3] Jahangir Khan,[4] Sayem Ahmed,[4] David H. Peters,[1] Louis Niessen,[1,5] and Antonio J. Trujillo[1]

AQ2

[1]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, [2]Centre for Health Economics, University of York, UK, [3]Harvard Medical School, Boston, MA, [4]International Center for Diarrheal Disease Research (Icddr,B), Dhaka, Bangladesh and [5]Liverpool School of Tropical Medicine, University of Liverpool, UK

AQ3

*Corresponding author. Centre for Health Economics, University of York, Alcuin 'A' Block, Heslington, York, YO10 5DD, UK. E-mail: andrew.mirelman@york.ac.uk

Accepted on 4 December 2015

Abstract

In low-income countries, a growing proportion of the disease burden is attributable to non-communicable diseases (NCDs). There is little knowledge, however, of their impact on wealth, human capital, economic growth or household poverty. This article estimates the risk of being poor after an NCD death in the rural, low-income area of Matlab, Bangladesh. In a matched cohort study, we estimated the 2-year relative risk (RR) of being poor in Matlab households with an NCD death in 2010. Three separate measures of household economic status were used as outcomes: an asset-based index, self-rated household economic condition and total household landholding. Several estimation methods were used including contingency tables, log-binomial regression and regression standardization and machine learning. Households with an NCD death had a large and significant risk of being poor. The unadjusted RR of being poor after death was 1.19, 1.14 and 1.10 for the asset quintile, self-rated condition and landholding outcomes. Adjusting for household and individual level independent variables with log-binomial regression gave RRs of 1.19 [standard error (SE) 0.09], 1.16 (SE 0.07) and 1.14 (SE 0.06), which were found to be exactly the same using regression standardization (SE: 0.09, 0.05, 0.03). Machine learning-based standardization produced slightly smaller RRs though still in the same order of magnitude. The findings show that efforts to address the burden of NCD may also combat household poverty and provide a return beyond improved health. Future work should attempt to disentangle the mechanisms through which economic impacts from an NCD death occur.

Key words: Non-communicable disease, poverty, Bangladesh

Key Messages

- Despite a growing global awareness of the emerging burden of non-communicable diseases (NCDs) in low and middle-income countries, there is little evidence of the microeconomic household impact.
- In rural Bangladesh, information from a demographic surveillance site allows the identification of households with NCD deaths and the use of multiple economic outcomes in a longitudinal cohort analysis.
- We find that the presence of an NCD death leads to an increased 2-year risk that households are poor according to asset score, self-report as poor or are land-poor.

Introduction

Non-communicable diseases (NCDs) are a growing portion of the disease burden in low- and middle-income countries. Troublingly, in addition to a growing burden, these countries also experience more premature mortality from NCD than high-income countries. A 2010 report by the World Health Organization (WHO) found that 29% of NCD deaths in low- and middle-income countries were in people younger than 60 years, whereas in high-income countries, this same statistic was 13% (WHO 2010b). It is generally understood that reducing premature mortality comes from the prevention of NCD-related risk factors such as smoking, sedentary lifestyles and poor diets, among others. In addition, recent research has also examined how poor health from NCD in low and middle-income countries may interact with conditions of poverty (Pandian *et al.* 2007; Schneider *et al.* 2009). In a macroeconomic study, Bloom *et al.* (2011) report the global economic production loss from NCDs such as cardiovascular disease, chronic respiratory disease, cancer, diabetes and mental health are estimated to be as high as $47 trillion.

Previous assessments of the effects of poor health on economic outcomes of households and individuals have focused on 'health shocks' or adverse health events of either death or disease (Wagstaff and Lindelow 2010; Alam and Mahal 2014). Gertler and Gruber (2002) found that households in Indonesia may not be able to fully recover economically after health shocks. In studies of health shocks where the shock is mortality, some of the most devastating impacts have been to households experiencing deaths of prime working-age adult members and the resulting loss of human capital (Beegle *et al.* 2008; Mather and Donovan 2008).

The analysis of health shocks has been concerned with removing the a priori relationship between socioeconomic status and health to isolate economic impact of health. In the USA, seminal work by Smith (1999) showed that adverse health could lead to a worsening economic condition. The framework for understanding how adverse health events effect economic outcomes was proposed by Russell (2004) and establishes that poor health can impact household wealth through direct and indirect costs. A similar framework published 2 years after Russell's was proposed by McIntyre *et al.* (2006). This framework included the concept of direct and indirect costs but also described specific coping mechanisms that households use in response to these costs (McIntyre *et al.* 2006). We draw on this latter framework and hypothesize that the economic impacts from NCD health shocks may result from direct costs incurred from health expenditure or indirect costs from the time cost associated with illness. There may also be long-term economic consequences through channels such as the loss of employment, reduced education of household members or the drawing down of household assets (Russell 2004; McIntyre *et al.* 2006).

Health shocks may result from either NCD or infectious disease; however, NCD explains a large proportion of the adult disease burden, and risk factor reduction may play an important role in preventing premature mortality. Public health and policy options for mitigating NCD and related risk factors may also be different than those for infectious diseases and isolating the economic impact is important for designing better health and social protection systems. Previous analyses of health shocks have identified health shocks from NCD to be of particular importance in low resource settings, though there is also a gap in the literature on the economic impacts of NCD (Gertler and Gruber 2002; Alam and Mahal 2014).

In Bangladesh, NCDs are the leading cause of mortality and the overall proportion of deaths from NCDs has grown over time (Ahsan Karar *et al.* 2009; IHME 2013). A WHO survey in Bangladesh found that nearly every adult, 98.7% of the population over age 25, had at least one NCD risk factor (WHO 2010a). It has also been found that several risk factors such as hypertension and NCD mortality cluster among the poor in Bangladesh (Razzaque *et al.* 2011). A recently published study looking across 24 years in rural Bangladesh also finds that the burden of mortality overall in rural Bangladesh is becoming more pro-poor, with communicable disease death rates staying consistently higher in the poor, while NCD death rates have shifted from being higher for the wealthier to being higher for the poor (Khan *et al.* 2015).

The objective of this study is to evaluate whether households in rural Bangladesh experiencing health shocks from NCD mortality have a higher risk of being poor after death. NCD health shocks in terms of adult mortality are examined in Matlab, Bangladesh, a rural area that is unique because of its long-running demographic surveillance program.

Methods

Data and study design

The two data sources for this study are the ongoing continuous health and demographic surveillance system conducted by the International Center for Diarrheal Disease Research in Bangladesh (icddr,b) and a specific survey designed to estimate the household economic impact of NCD mortality. The first source is a yearly census of vital statistics that also includes a periodic socioeconomic status census, approximately every 10 years that collects information on the ownership of household assets. The economic impact survey was collected in Matlab in the year 2012, up to 2 years post NCD death, and included separate modules on the demographic characteristics and the economic impact from the NCD death.

A matched cohort design was used where the exposure of interest was an adult death from an NCD (Sjolander *et al.* 2012). Matched cohort designs are relatively rare, but it was used here because of the ability to identify comparison households using the census surveillance data. The matching procedure established a comparison group of households that balanced confounding variables related to the household's economic status prior to the exposure to an NCD death and the longitudinal cohort design accounts for reverse causation (Mahal *et al.* 2010).

The study population consisted of all adult NCD deaths in those aged over 15 years in the calendar year 2010. A total of 909 adult NCD deaths were identified and 856 of the households were surveyed in 2012, representing a 6% attrition rate. This is the same as the annual percentage of individual outmigration in Matlab in 2010 (Icddr,b 2012). Our attrition seems plausible and low given that household dissolution may be higher after a death (Hosegood *et al.* 2004).

From the households with an NCD death, the deceased was matched to another individual in a comparison household with no deaths in the year 2010. The direct matching procedure found households with comparable individuals based on the age, sex and village of the deceased individual. Comparison individuals were matched exactly on sex, in 5-year age bands and in the same or nearest village. The logic of the matching procedure was that comparison should occur between households with a similar individual, as a contributor to the household economic production and consumption. While the geographic region of Matlab is relatively homogenous, matching on this variable also provides a control household that has similar economic condition and production opportunities. Two matching households were identified for each household with

an NCD death. A diagram of the study design is provided in Supplementary Appendix S1.

Outcome variables

Three measures of economic status were used: an asset-based wealth index, self-reported economic condition and the total amount of land that a household owns (Ravallion and Lokshin 2000; Filmer and Pritchett 2001; Bhuiya et al. 2005; Wagstaff and Lindelow 2010). Each of the three measures were assessed in the 2012 survey; however, only self-reported economic status was asked prior to the death. This was because the most recent socioeconomic status census in 2005 was too far in the past to reliably assess asset quintile or land-owning prior to the death.

The three outcomes provide different but complementary pictures of the socioeconomic condition of the household. The asset-based index, based on durable household items provides an estimate for the distribution of wealth based on ownership of durable items but lacks information about price. Self-reported economic condition accounts for a subjective component of wealth but may capture only relative economic information. Finally, the measure of total land-holding, commonly used in rural Bangladesh, may be limited because of the shifting importance of agriculture in the economy. Detailed information regarding the definitions, calculations and advantages and disadvantages for each of the three economic outcomes is provided in Supplementary Appendix S2.

Primary independent variable—NCD mortality

The 909 adult deaths from NCDs were identified by International Classification of Disease version 10 (ICD-10) codes that are assigned by the surveillance team in Matlab through a dual physician review verbal autopsy. Deaths from injuries including unintentional injury such as accident and drowning, and intentional injury such as suicide and homicide were excluded (Icddr,b 2012). The causes of NCD death included cancer, COPD, diabetes, cardiovascular disease (including hypertensive disorders, ischemic heart disease and stroke), blood disorders, metabolic disease, mental disorders, neurological disease as well as other respiratory and digestive diseases. The total numbers of deaths for each cause is provided in Supplementary Appendix S3. Stroke contributes the most deaths, which is consistent with studies of NCD burden in South Asia (Wasay et al. 2014).

AQ4
Statistical analysis
Contingency tables and log-binomial regression

Aggregate and stratified contingency tables were initially used to examine the relative risk (RR) of being poor 2 years after death given an NCD health shock (Greenland and Morgenstern 1990; Cummings et al. 2003). Multivariate analyses used a log-binomial regression model, mathematically similar to Poisson regression, to provide estimates of RR conditional on the independent variables. This approach is similar to the discrete poverty approach using a multinomial regression model used previously for understanding movements into and out of poverty in two time periods (Glewwe et al. 2002; Justino et al. 2008). The model is specified as follows:

$$\log Y_i(t) = \alpha(t) + \beta_1 NCD_i + \beta_k X_i \qquad (1)$$

Where Y_i is an indicator of whether household i is poor or not according to one of the three measures for economic status, thus accounting for households that both remain in a state of poverty and households that move into a state of poverty after an NCD death. β_1 is the coefficient on the indicator for whether a household had an NCD adult death in 2010. Taking the exponent β_1 gives the RR of being poor for households with an NCD death. β_k represents a vector of coefficients corresponding to a vector of independent variables X_i. The model is run separately for each of the three economic outcomes of interest and standard errors are clustered at household level. A model with an interaction term for prime-age status of the deceased was also used to the impact of premature NCD mortality.

Marginal effect estimation with regression standardization and machine learning

An estimator of marginal effect was calculated with regression standardization, an extension of epidemiologic standardization methods (Rothman et al. 2012; Sjolander and Greenland 2013). This estimator uses the marginal distribution of the baseline covariates and matching variables in the exposed subjects to estimate the familiar marginal exposure effect on the exposed parameter. This parameter is commonly referred to as the average treatment effect on the treated (ATT) in the health economics and statistics literature. The standardized estimator is calculated through a process of averaging over observed covariate distributions in the exposed subjects after obtaining counterfactual outcomes. The regression standardized version of the estimator uses the parametric regression model to calculate the expected counterfactual scenarios and obtain a marginal effect by averaging over all of the covariates (Snowden et al. 2011; Sjolander and Greenland 2013).

One of the limitations of the regression standardization approach, as well as the log-binomial regressions, regards fitting the model. If the model is constructed incorrectly, then bias could result. To explore this further, a machine learning algorithm was used for specifying the regression function and calculating a subsequent standardized estimator. The machine learning algorithm implemented here is a 'super learner', which is a type of machine-learning algorithm called an 'ensembling' algorithm (van der Laan et al. 2007; van der Laan and Rose 2011). The super learner algorithm is implemented in R programming language using the *SuperLearner* package and bootstrapping is used to obtain standard errors and confidence intervals (CIs, Polley and van der Laan 2013).

More details on the regression standardization and machine learning approach are provided in Supplementary Appendix S4.

Ethical approval

This research has been approved by the Institutional Review Board of the authors' institutions.

Results

Relation between the three economic outcomes and descriptive information

The percentage of the population classified as poor in 2012 differed depending on which measure was used (Table 1). For the asset index, 22% of the study population was classified as poor at follow up and for self-reported economic condition and landholding this figure was 42% and 63%. The largest correlation is between the asset index and self-rated economic condition, with a Spearman rank coefficient of 0·39. The asset index and total landholding have the lowest correlation, with a coefficient of 0·26 and the correlation between self-rated economic condition and landholding has a coefficient of 0·32.

Descriptive statistics for the independent variables are in Table 2. The distribution of variables used for matching: age, sex and village show that there is balance among the groups. Household position is the only independent variable that shows a significant difference between groups; fewer of the NCD group are either household heads or

AQ14 **Table 1**. Comparison and correlation of three economic outcomes at follow-up

Economic outcome, 2012	Description	Percentage poor (SE)	Correlation		
			Asset Index	Self-rated condition	Landholding
1. Asset index	Wealth index based on a list of 26 durable household items and classified into 5 quintiles.	0.22 (0.01) Threshold: 3rd quintile	NA	0.39	0.26
2. Self-rated condition	Perceived ranking of household economic condition by household representative on a scale of 1 = poorest to 5 = richest.	0.42 (0.01) Threshold: 3rd ladder step	0.39	NA	0.32
3. Landholding	Total amount of land area for homestead and agriculture that a household reports owning.	0.63 (0.01) Threshold: 50 dm total land	0.26	0.32	NA

Notes: Percentage poor is for the pooled sample group with NCD death and comparison group. Thresholds for being poor include being in the 1st two quintiles for the asset index, being on the first two ladder rungs for self-rated condition and owning <50 decimals of land. Correlation for each measure is measured with Spearman's rank coefficient. NA is not applicable because of perfect correlation. The asset index and associated wealth quintile are calculated using polyconic principal component analysis (PPCA) with eigenvalue weights from the Matlab socioeconomic census in 2005. Self-rated condition asks respondents to rank the household's economic condition on a 5-step ladder. Landholding is measured in decimals, which is equivalent to 1/100 of an acre or 40.46 square meters.

Table 2. Characteristics of the study population prior to death for NCD group and comparison group

Variable	Pooled	NCD death	Comparison	Test for diff. (NCD vs comparison)
	Mean (SD)/proportion	Mean (SD)/proportion	Mean (SD)/proportion	(P value or X^2)
Matching variables (deceased individual and matched comparison)				
Age	67.46 (12.46)	67.71 (12.72)	67.33 (12.33)	0.48
Prime age (% 15–59)	0.21 (0.01)	0.21 (0.01)	0.21 (0.01)	0.94
Female	0.45 (0.01)	0.45 (0.02)	0.45 (0.01)	NA
Number of villages	145	136	145	NA
Individual characteristics (deceased individual and matched comparison)				
Education				
None	0.61 (0.01)	0.62 (0.02)	0.60 (0.01)	0.20
1-5 years	0.25 (0.01)	0.23 (0.01)	0.26 (0.01)	
6+ years	0.15 (0.01)	0.16 (0.01)	0.14 (0.01)	
Marital status				
Single/unmarried	0.02 (0.00)	0.02 (0.00)	0.01 (0.00)	0.11
Married	0.62 (0.01)	0.64 (0.02)	0.61 (0.01)	
Divorced	0.01 (0.00)	0.01 (0.00)	0.00 (0.00)	
Widowed	0.36 (0.01)	0.34 (0.02)	0.38 (0.01)	
Head or Spouse of Head	0.71 (0.01)	0.67 (0.02)	0.73 (0.01)	0.00
Poor	0.43 (0.50)	0.44 (0.50)	0.43 (0.49)	0.43
Household characteristics				
Muslim	0.86 (0.01)	0.85 (0.01)	0.869 (0.01)	0.34
Household size	6.21 (2.92)	6.18 (2.70)	6.23 (3.02)	0.68
N	2585	856	1729	

Notes: NA (not applicable) applies to the test for difference for sex because of exact matching. For villages, the test of differences was not applicable. Student's *t*-test P values are calculated for continuous measures of age and household size. X^2 P values are calculated for categorical and binary variables. All values are estimated in the baseline year prior to death, calendar year 2009. The 856 NCD households represent all of the identified 909 households that had deaths in 2010. This means there was an attrition rate of 6% for the study. Poor is assessed by self-reported economic condition prior to death for NCD households and in 2009 for comparison households (asked retrospectively in the 2012 NCD and Economics survey). Methods note: The differences between descriptive characteristics for the NCD group and the comparison group at baseline are first assessed. There is ambiguity in the literature about whether accounting for matched variables in a matched cohort analysis is needed. We use a *t*-test statistic for continuous variables and assuming independence between the groups. A paired *t*-test may also be used in matched cohort study and we find similar result when using a paired or unpaired test but only report the unpaired results (Cummings et al. 2003, Sjolander and Greenland 2013). A chi-squared test for independence is used for categorical variables (Faraway 2006).

spouses of the household head. The percentage of households who retrospectively report as poor according to the outcome of self-reported wealth prior to the death is 44% for the NCD households and 43% for the comparison households. In sum, the overall descriptive information shows that the groups are neither statistically nor substantially different prior to the NCD death and position in the household is controlled for in subsequent models to account for potential bias.

Results from contingency tables and log-binomial regression

The RRs of being poor and 95% CIs are listed in Table 3. For the pooled group, a significant RR of 1.19 is found using the asset quintile. This RR is 1.14 for self-rated condition and 1.10 for landholding and both are significant. The stratified contingency table also showed that different variables of the deceased individual lead to a significant risk of being poor depending on the outcome that was

Table 3. RR estimates for the NCD vs the comparison group for three economic outcomes

	Asset quintile			Self-rated economic condition			Landholding		
	RR	Lower	Upper	RR	Lower	Upper	RR	Lower	Upper
Full sample	1.19	1.02	1.39	1.14	1.03	1.26	1.10	1.03	1.17
Sex									
Male	1.22	0.98	1.52	1.18	1.04	1.34	1.18	1.08	1.29
Female	1.16	0.91	1.48	1.09	0.94	1.27	1.01	0.93	1.10
Age									
Prime age	1.96	1.39	2.76	1.49	1.24	1.79	1.07	0.94	1.22
Old age	1.04	0.83	1.30	1.05	0.93	1.19	1.11	1.03	1.19
Education									
None	1.13	0.96	1.34	1.16	1.04	1.29	1.04	0.97	1.11
1–5 years	1.28	0.85	1.94	0.99	0.00	INF	1.30	1.12	1.51
6+ years	1.55	0.66	3.64	1.36	0.91	2.03	1.12	0.89	1.40
Marital status									
Single/widow/divorce	1.08	0.81	1.45	1.01	0.71	1.43	1.09	0.99	1.20
Married	1.26	1.03	1.54	1.22	1.08	1.38	1.12	1.03	1.22
Position									
Head/spouse	1.18	0.97	1.43	1.16	1.03	1.3	1.14	1.06	1.23
Non-head/spouse	1.18	0.86	1.63	1.05	0.86	1.28	1.03	0.90	1.18

Notes: In the table, 'Asset quintile' refers to being in the bottom two quintiles as measured by asset-based principal component analysis. 'Self-rated economic condition' refers to being in the poor or very poor group and 'Landholding' refers to owning <50 decimals of land. A decimal of land refers to 1/100th of an acre or 40.46m^2. 'Prime Age' here refers to those deaths to individuals aged 15–59 years and 'Old Age' refers to deaths to those aged 60 years and older.

Table 4. Log-binomial regression for the effect of an adult NCD death on measures of household economic condition

	Asset quintile	Asset quintile - PA interaction	Self-rated condition	Self-rated condition - PA interaction	Landholding	Landholding - PA interaction
Intercept	0.85	1.60	1.94*	1.89	1.34	1.55
NCD death	1.19*	1.07	1.16**	1.09	1.14**	1.15**
Prime age	0.74**	0.74**	0.86*	0.86	0.84**	0.84**
Female	0.99**	0.98***	0.99***	0.99**	0.99***	0.99***
Age	0.55***	0.55***	0.83**	0.83**	0.76***	0.76***
Education						
1–5 years (ref. none)	0.26***	0.27***	0.43***	0.44***	0.65***	0.65***
6+ years (ref. none)	1.69	2.06	1.30	1.34	2.09***	2.14***
Marital status						
Married (ref. unmarried)	2.94	3.40	1.23	1.18	1.17	1.23
Divorced (ref. unmarried)	1.86	2.25	1.41	1.44	2.41***	2.47***
Widowed (ref. unmarried)	1.01	1.00	0.93	0.93	0.94	0.94
HH head/spouse	0.87	0.87	0.88	0.88	0.78***	0.78***
Muslim	0.92***	0.92***	0.95***	0.95***	0.97***	0.97***
Household size	1.31	1.38	0.80	0.81	0.42	0.42
NCD x prime age		1.70***		1.33*		0.94
log likelihood	−1216.67	−1212.17	−1863.08	−1861.32	−2203.54	−2203.18
Deviance	1355.33	1346.34	1582.16	1578.63	1255.09	1254.35
Num. obs.	2491	2491	2491	2491	2490	2490

Notes: All standard errors are clustered at the household level. PA stands for prime-age, which is age 15–59 years. Significance: ***P ign.01, **P 1, .05, *P 5, .1.

used. For the asset quintile, this was prime age status and marital status (RRs: 1.96 and 1.26). For self-rated condition, deaths to male members, prime age members, uneducated members, married members and household heads lead to significant increased risks of being poor. With landholding, the death of a male member, old age member, member with primary education, married member or household head shows a significant increased risk of being poor. In the latter two outcomes, the RR of being poor given a prime age death for self-reported economic condition and the RR of being poor given the death of someone with primary education for landholding had some of the highest values (RR: 1.49 and 1.30).

Table 4 lists the results from the log-binomial regression. The RRs shown are the coefficients from the regression equation taken as exponents. After adjusting for all independent variables, the RRs are significant at the 5% level for self-reported wealth and landholding and at the 10% level for asset quintile. The RR of being poor after an NCD death with the asset quintile is 1.19 (SE 0.09), with self-rated economic condition this is 1·16 (SE 0.07) and with landholding this becomes 1.14 (SE 0.06). The death of a prime age member shown by an interaction term in a separate column is also shown to significantly increase the risk of being poor when the asset quintile and self-rated condition are the outcomes of interest (RR: 1.70,

Table 5. Regression standardization marginal effect of an adult NCD death on measures of household economic condition using parametric regression and machine learning

	RR	SE	Lower	Upper
Parametric regression standardization				
Asset quintile	1.19	0.09	1.01	1.37
Self-rated condition	1.16	0.05	1.05	1.27
Landholding	1.14	0.03	1.08	1.20
Machine-learning (super learner) estimation				
Asset quintile	1.15	0.09	0.98	1.32
Self-rated condition	1.13	0.05	1.03	1.23
Landholding	1.11	0.03	1.05	1.17

Notes: All results are calculated with the RR equation: $\hat{\phi}RR = \dfrac{\frac{1}{N}\sum_{i=1}^{N}\left[\hat{E}(Y|A=1,W)\right]}{\frac{1}{N}\sum_{i=1}^{N}\left[\hat{E}(Y|A=0,W)\right]}$, using parametric regression standardization and

machine learning. The machine learning uses an ensembling super-learner algorithm described in the references. The super learner algorithm is implemented in R programming language using the SuperLearner package (Polley and van der Laan 2013). A collection of three algorithms were used in this analysis: logistic regression implemented with the generalized linear models (glm) package, the arithmetic mean where the marginal probability of being poor in each cross-validation fold is assigned to each household and a final package (glmnet) for penalized regression using the least absolute shrinkage and selection operator (LASSO).

SE 0.22 and 1.33, SE 0.15). For landholding, the effect of an NCD death to a prime age member does not significantly modify the risk of being poor (RR: 0.94, SE 0.13).

Results from the regression standardization and machine learning for marginal effects

The results of the regression standardization estimator using parametric regression are listed in Table 5. The results represent the marginal effects for the risk of being poor in 2012 given an adult NCD death in 2010. For the asset-based quintile, self-rated status and landholding the RR was positive and significant with values of 1.19 (SE 0.09), 1.16 (SE 0.05) and 1.14 (0.03), respectively, and 95% CIs above 1. This means that having a death leads to 19%, 16% or 14% greater risk of being poor in 2012. The results with the super-learner machine learning approach show similar results to the regression standardization approach but overall have slightly smaller RRs. Interestingly, the RR using the asset-based quintile, 1.15 (SE 0.09), is no longer statistically significant at a 5% level since the CI crosses 1 (95% CI: 0.98–1.32). The results for self-rated condition (RR = 1.13, SE 0.05) and land-holding (RR = 1.11, SE 0.03), however, do find significantly elevated risks of being poor.

For relative understanding, our findings of increased RR of being poor after having adult NCD death is higher than the increased risk of being poor of 8%, 5% and 3% (for asset quintile, self-reported wealth and landholding) found for non-Muslim households, which are generally acknowledged as being more disadvantaged in rural regions of Bangladesh. While there are no comparable studies in Bangladesh to compare our findings, researchers in Vietnam have assessed the risk of moving into a poor state for those that are non-poor before an injury and found a RR of 1.21 (Thanh *et al.* 2006). Several studies have also used a self-reported measure of economic well-being to assess the impact of a health shock. In Tanzania, researchers found that 20% of households reported having a year of 'very bad' living conditions specifically due to the death of a household member (Beegle *et al.* 2008).

Discussion

This work adds to an understanding of the impact of an NCD death in terms of remaining or becoming poor and finds that there is an increased incident risk of being poor following NCD death. This is found to be robust for different economic outcomes and estimation methods. The results provide evidence that there may be a pathway from NCD health shock to economic outcome that could result from either direct and indirect cost of illness or the longer term impact of coping strategies.

Previous work in Matlab has only examined the cross-sectional relationship between mortality and poverty (Razzaque *et al.* 2009). Our study highlights that there are significant economic impacts from NCDs, which represent the majority of adult mortality in Matlab. The RR of staying or becoming poor is largest when measured with asset quintile, followed by self-rated condition and then landholding. A prime age death is also found to have a large impact on the risk of poverty when it is measured by asset quintile. Measuring poverty according to landholding, the significant protective effect found for a death of a household head or associated spouse may be due to the redistribution of land after a head dies. Overall, given the low correlation between the three economic outcomes, it is interesting that we find NCD mortality increases the risk of poverty in each of them.

Another strength of this study is the use of census surveillance data to identify all households with an NCD death. Using mortality mitigates the potential upward bias found in studies that use self-reported health (Grimm 2010; Islam and Maitra 2012). The inclusion of mortality, though, also means the exclusion of relevant information about morbidity, such as the length of illness and the severity prior to the death. There may also be unobserved health shocks from morbidity, where a sick person recovers, in the comparison group. This deserves further exploration in a setting where morbidity is observed, and shocks from morbidity and subsequent mortality can be separated. Another limitation is the aggregation of the causes of death into a broad category for NCD. This may mask important heterogeneities due to differences in illness prior to death. Factors such as the length and intensity of illness and the impact that these have on household economic condition should be an area of future research and will be important for developing NCD policies that target specific illness. While this study was not designed to look at individual causes of death, we have provided log-binomial regression results for seven broad NCD causes in Supplementary Appendix S5.

Finally, the generalizability of these findings to other settings may be limited given the long-running health and surveillance program in Matlab. The NCD burden may be higher than other rural, low-income areas and this could bias the economic impact of NCD mortality upward. On the other hand, if residents of Matlab have healthier behaviours and greater health knowledge, the impact of an NCD death on households could be diminished. Finally, causal inferences from our study may still be limited because of unobservable factors such as health behaviours and inter-household preferences that differ between the two groups in our study population, leading to misspecification of the model.

This work looks at the economic impacts of deaths and whether there are negative consequences for households. There are several policy instruments that may be warranted and policies to provide improved financial risk protection should follow from further evidence on the importance of direct costs, indirect costs and coping strategies that the households incur. One approach would be to help households adjust economically after the mortality through better access to financial protection tools and risk-pooled insurance. Micro-lending in Bangladesh is ubiquitous and a recent study has established stronger evidence for its poverty alleviation impact over the long-term (Khandker and Samad 2014). Formal health insurance mechanisms are rare in rural Bangladesh and establishing these types of programs in an area with a large informal sector such as rural Bangladesh poses many challenges (Acharya *et al.* 2013; Bannerjee *et al.* 2014). More broadly, measures should be taken to address the burden of NCDs through more health services for prevention. Such services could be provided by the government or a non-profit organization in much the same way that basic maternal and child health services have been provided in Matlab.

This study proves that there is an economic argument to be made for addressing the burden of NCDs in rural, low-income settings and the methods developed here provide a model for estimating the economic impacts from health shocks in other settings as well. The increased risk of being poor in the follow-up period for households with an NCD death ranged from 14% to 19%. For two of these outcomes, asset quintile and self-rated economic condition, an NCD death to a prime age household member moderates the economic impact and increases the risk of a household being poor. Without further action, households will have higher risks of moving into or staying in poverty because of the expenses and loss of human capital imposed by NCD health shocks.

Supplementary data

Supplementary data are available at *HEAPOL* online.

Ethical approval

This research has been approved by the Institutional Review Board of the Johns Hopkins Bloomberg School of Public Health and the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b).

Conflict of interest statement. None declared.

References

Acharya A, Vellakkal S, Taylor F *et al*. 2013. The impact of health insurance schemes for the informal sector in low and middle-income countries. *Policy Research Working Paper*. The World Bank.

Ahsan Karar Z, Alam N, Kim Streatfield P. 2009. Epidemiological transition in rural Bangladesh, 1986-2006. *Global Health Action* **2**.

Alam K, Mahal A. 2014. Economic impacts of health shocks on households in low and middle income countries: a review of the literature. *Global Health* **10**: 21

Bannerjee A, Duflo E, Hornbeck R. 2014. Building health insurance and microfinance in India: there cannot be adverse selection if there is no demand. *American Economic Review: Papers and Proceedings* **104**: 291–7.

Beegle K, De Weerdt J, Dercon S. 2008. Adult mortality and consumption growth in the age of HIV/AIDS. *Economic Development and Cultural Change* **56**: 299–326.

Bhuiya A, Mushtaque A, Chowdhury R, Momen M, Khatun M. 2005. Marital disruption: determinants and consequences on the lives of women in a rural area of Bangladesh. *Journal of Health Population and Nutrition* **23**: 82–94.

Bloom De, Cafiero ET, Jane-Llopis E *et al*. 2011. *The Global Economic Burden of Non-Communicable Diseases*. Geneva: World Economic Forum.

Cummings P, McKnight B, Weiss NS. 2003. Matched-pair cohort methods in traffic crash research. *Accident Analysis Prevention* **35**: 131–41.

Faraway J. 2006. *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*.Chapman and Hall/CRC.

Filmer D, Pritchett L. 2001. Estimating wealth effects without expenditure data-or tears: an application to educational enrollments in states of India. *Demography* **38**: 115–32.

Gertler P, Gruber J. 2002. Insuring consumption against illness. *American Economic Review* **92**: 51–70.

Glewwe P, Gragnolati M, Zaman H. 2002. Who gained from Vietnam's boom in the 1990s? *Economic Development and Cultural Change* **50**: 773–92.

Greenland S, Morgenstern H. 1990. Matching and efficiency in cohort studies. *American Journal of Epidemiology* **131**: 151–9.

Grimm M. 2010. Mortality shocks and survivors' consumption growth*. *Oxford Bulletin of Economics and Statistics* **72**: 146–71.

Hosegood V, McGrath N, Herbst K, Timaeus IM. 2004. The impact of adult mortality on household dissolution and migration in rural South Africa. *AIDS* **18**: 1585–90.

Icddr,b 2012. *Health and Demographic Surveillance System—Matlab: Volume Forty Four Registration of Health and Demographic Events 2010*. Dhaka, Bangladesh: Icddr,b.

IHME 2013. *GBD Profile: Bangladesh*. Seattle, WA: Institute for Health Metrics and Evaluation.

Islam A, Maitra P. 2012. Health shocks and consumption smoothing in rural households: does microcredit have a role to play? *Journal of Development Economics* 232–43.

Justino P, Litchfield J, Pham HT. 2008. Poverty dynamics during trade reform: evidence from Rural Vietnam. *Review of Income and Wealth* **54**: 166–92.

Khan J, Trujillo A, Ahmed S *et al*. 2015a. Distribution of non-communicable disease mortality and association with poverty in a rural area in Bangladesh—a household panel data analysis. *International Journal of Epidemiology*

Khan JA, Trujillo AJ, Ahmed S *et al*. 2015b. Distribution of chronic disease mortality and deterioration in household socioeconomic status in rural Bangladesh: an analysis over a 24-year period. *International Journal of Epidemiology*

Khandker S, Samad H. 2014. Dynamic effects of microcredit in Bangladesh. *Policy Research Working Paper*. The World Bank.

Mahal A, Karan A, Engelgau M. 2010. The economic implications of non-communicable disease for India. *HNP Discussion Paper*. Washington, DC: The World Bank.

Mather D, Donovan C. 2008. The impacts of prime-age adult mortality on rural household income, assets, and poverty in Mozambique. *Food Security Collaborative Working Papers*. Michigan State University Department of Agricultural, Food, and Resource Economics.

McIntyre D, Thiede M, Dahlgren G, Whitehead M. 2006. What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? *Social Science and Medicine* **62**: 858–65.

Pandian JD, Srikanth V, Read SJ, Thrift AG. 2007. Poverty and stroke in India: a time to act. *Stroke* **38**: 3063–9.

AQ13 Polley E, van der Laan MJ. 2013. Package 'SuperLearner'.

Ravallion M, Lokshin M. 2000. Identifying welfare effects from subjective questions. *Economica* **68**: 335–57.

Razzaque A, Carmichael GA, Streatfield PK. 2009. Adult Mortality in Matlab, Bangladesh. *Asian Population Studies* **5**: 85–100.

Razzaque A, Nahar L, Abu Haider Mohammad Golam M *et al*. 2011. Sociodemographic differentials of selected noncommunicable diseases risk factors among adults in Matlab, Bangladesh: findings from a WHO STEPS survey. *Asia Pacific Journal of Public Health* **23**: 183–91.

Rothman KJ, Greenland S, Lash TL. 2012. *Modern Epidemiology*. Lippincott Williams and Wilkins.

Russell S. 2004. The economic burden of illness for households in developing countries: a review of studies focusing on malaria, tuberculosis, and human immunodeficiency virus/acquired immunodeficiency syndrome. *American Journal of Tropical Medicine and Hygiene* **71**: 147–55.

Schneider M, Bradshaw D, Steyn K, Norman R, Laubscher R. 2009. Poverty and non-communicable diseases in South Africa. *Scandinavian Journal of Public Health* **37**: 176–86.

Sjolander A, Greenland S. 2013. Ignoring the matching variables in cohort studies–when is it valid and why? *Statistics in Medicine* **32**: 4696–708.

Sjolander A, Johansson AL, Lundholm C *et al*. 2012. Analysis of 1:1 matched cohort studies and twin studies, with binary exposures and binary outcomes. *Statistical Sciences* **27**: 395–411.

Smith JP. 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. *The Journal of Economic Perspectives* **13**: 145–66.

Snowden JM, Rose S, Mortimer KM. 2011. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *American Journal of Epidemiology* **173**: 731–8.

Thanh NX, Hang HM, Chuc NT *et al*. 2006. Does "the injury poverty trap" exist? a longitudinal study in Bavi, Vietnam. *Health Policy* **78**: 249–57.

van der Laan MJ, Polley EC, Hubbard AE. 2007. Super learner. *Statistical Applications in Genetics and Molecular Biology* **6**: Article25.

van der Laan MJ, Rose S. 2011. *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York: Springer.

Wagstaff A, Lindelow M. 2010. Are health shocks different?: evidence from a multi-shock survey in Laos. *Policy Research Working Paper*. The World Bank.

Wasay M, Khatri IA, Kaul S. 2014. Stroke in South Asian countries. *Nature Reviews Neurology* **10**: 135–43.

WHO 2010a. *Bangladesh NCD Risk Factor Survey 2010*. Geneva: World Health Organization.

WHO 2010b. *Global Status Report on Noncommunicable Diseases 2010: Description of the Global Burden of NCDs, Their Risk Factors and Determinants*. Geneva: World Health Organization.

Author Query Form

AUTHOR QUERIES – TO BE ANSWERED BY THE CORRESPONDING AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please click on each query number and respond by indicating the change required within the text of the article. If no change is needed please add a note saying "No change."

AQ1:   Please check whether the edit made to the article title is appropriate.

AQ2:   Please check that all names have been spelled correctly and appear in the correct order. Please also check that all initials are present. Please check that the author surnames (family name) have been correctly identified by a pink background. If this is incorrect, please identify the full surname of the relevant authors. Occasionally, the distinction between surnames and forenames can be ambiguous, and this is to ensure that the authors' full surnames and forenames are tagged correctly, for accurate indexing online. Please also check all author affiliations.

AQ3:   Please provide department/division name (if any), full road and district address and Zip or postal code in all affiliations.

AQ4:   Please check whether all section heads and subheads have been set as intended.

AQ5:   Please provide a Funding statement, detailing any funding received. Remember that any funding used while completing this work should be highlighted in a separate Funding section. Please ensure that you use the full official name of the funding body, and if your paper has received funding from any institution, such as NIH, please inform us of the grant number to go into the funding section. We use the institution names to tag NIH-funded articles so they are deposited at PMC. If we already have this information, we will have tagged it and it will appear as coloured text in the funding paragraph. Please check the information is correct.

AQ6:   Please check whether the placement of Ethical Approval is appropriate.

AQ7:   Please provide publisher location in references "Wagstaff and Lindelow 2010," "Rothman et al. 2012," "Acharya et al. 2013," "Mather and Donovan 2008," "Khandker and Samad 2014" and "Faraway 2006."

AQ8:   Please provide page range in the reference "Ahsan Karar 2009".

AQ9:   Please note that the reference "Beegle et al. 2008b" seems to be an incorrect repetition of the reference "Beegle et al. 2008a." Hence, "Beegle et al. 2008b" has been deleted and the reference citations have been set appropriately.

AQ10:   Please update the volume number, year and page range in the reference Khan et al. 2015a.

AQ11:   Please note that as there are two "Khan et al. 2015" references in the reference list, they have been differentiated using "a" and "b." Please add "a" and "b" appropriately in the corresponding reference citations.

AQ12:   Please provide volume number and page range in the reference Khan et al. 2015b.

AQ13:   Please provide complete details in the reference Polley and van der Laan 2013.

AQ14:   Please check whether all tables are OK as typeset.