

# A Model to Estimate the Impact of Thresholds and Caps on Coverage Levels in Community-Based Health Insurance Schemes in Low-Income Countries

Erika Binnendijk<sup>1</sup>, Ruth Koren<sup>2</sup>, David M. Dror<sup>1,3\*</sup>

<sup>1</sup>Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

<sup>2</sup>Felsenstein Medical Research Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

<sup>3</sup>Micro Insurance Academy, New Delhi, India

Email: \*[daviddror@socialre.org](mailto:daviddror@socialre.org)

Received 7 January 2014; revised 10 February 2014; accepted 19 February 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

**Background:** Community-based health insurance (CBHI) schemes are increasingly implemented in low-income settings. These schemes limit the coverage they offer both by the types of care considered, and by applying thresholds and/or caps to costs reimbursed. The consequences of these thresholds and/or caps on insurance coverage have hitherto been usually ignored, for lack of data on the distributions of healthcare costs or understanding of their impact on effective coverage levels. This article describes a theoretical model to obtain the distributions even without data collection in the field, and demonstrates the quantitative impact of thresholds and/or caps on claim reimbursements. **Methods:** This model applies to applications on healthcare expenditures in low-income settings, following research methods examined in the Western world. We looked at hospitalizations and tests; we compared the simulated distributions to empirical data obtained through 11 household surveys conducted between 2008 and 2010 in rural locations (9 in India and 2 in Nepal). **Results:** We found that the shape of the distributions was very similar in all locations for both benefits, and could be represented by a model based on a lognormal distribution. The agreement between theoretical and empirical results was satisfactory (mostly within 10% difference). **Conclusions:** The model makes it possible to simulate the expected performance of the CBHI (represented by the percentage of costs or bills covered). The aim is to match costs with local levels of willingness-to-pay for health insurance. This model makes it possible to determine at the stage of package-design the optimal levels of thresholds and/or caps for each benefit-type included.

---

\*Corresponding author.

## Keywords

**Distributions; Healthcare Costs; Community-Based Health Insurance; Micro Health Insurance; Thresholds; Caps**

## 1. Introduction

Most healthcare spending in developing countries is borne by healthcare-seekers out-of-pocket (OOP). India is a good example: 70% of health spending is private, 86% of which is OOP [1] [2]. Households in India frequently finance such OOP spending, not only for inpatient care but also for outpatient care and even for maternity-related costs, by borrowing money with interest or selling assets [3]. This inequitable and inefficient health financing situation persists in other low-income countries as well [4]. The solution proposed by WHO and other international bodies has been to strive toward universal health coverage, notably through prepayment and risk pooling mechanisms in lieu of payments at the point and time of service delivery [5] [6]. Yet, penetration of health insurance in most low-income countries remains very low [7]; in India, insurance uptake is below 5% [8]. One possible explanation for low insurance uptake is that poorer individuals in the informal sector doubt their own ability to enforce contracts with insurance companies. A solution to the problem is for people to own and run community-based health insurance (CBHI) schemes at village level [9]-[13].

Typically, CBHI schemes do not benefit from premium subsidy. Therefore, these stand-alone schemes must ensure that benefit expenditures would be limited to premium income (reflecting willingness-to-pay levels). This can be done through the application of a threshold (the predetermined amount above which the insurance reimburses the rest of the bill) and/or a cap (the predetermined amount up to which the insurance reimburses the bill). For premium calculations, information is needed on the probability of healthcare utilization by insured persons and the cost of that utilization. The consequences of applying thresholds and/or caps on the expected average pay-out of insurance, and thus on the premium, can be calculated only when the distribution of the costs of healthcare utilization is known. It is self-explanatory that when low caps are applied, the effective protection of insured persons for costs above the cap is diminished. The lower the cap, the more likely the situation of reduced insurance cover, which could lead to a perception that the insurance will not provide the protection the clients seek. Such perception could greatly dampen willingness to affiliate. It is recalled that the actual assessment of the effects of caps on the protection provided by the insurance can be known in advance only if the distribution of costs is known.

Our study builds on previous research, which confirmed the imperative need to obtain information locally [14] because locations differ significantly in the number and type of illness episodes [15], cost of healthcare [16] and the willingness-to-pay (WTP) for health insurance [17]-[19]. Location-specific data cannot be extracted accurately from national data; one source would be actual claims data (fixing premiums based on such information is called “experience rating”), but this can apply only to schemes already in operation. Projections for new schemes must rely on household surveys. Indeed, a number of experiments with CBHI have relied on baseline data collection (mainly household surveys) to obtain reliable local estimates required for package design and pricing [20]-[25]. However, the time and money required to obtain information through household surveys prior to launching CBHI impede the replication of new CBHIs. To reduce these impediments preceding the launch of CBHI, we have developed faster and cheaper methods to estimate morbidity and healthcare utilization through a quick intervention method we called “Illness Mapping” [26], as well as estimates of WTP [18]. This article deals with estimating the distributions of healthcare costs, which is the third piece of information needed.

The purpose of this paper is to develop a method for premium calculation of locally relevant health insurance packages that apply a threshold and/or a cap to benefits, and to show that this can be done even without cumbersome and expensive data collection efforts in each location where CBHI is implemented.

## 2. Data and Methods

### 2.1. Setting and Sampling

The study is based on data obtained from 11 household surveys conducted in the following periods and locations:

2008, in Warangal and Karimnagar districts, Andhra Pradesh, India; 2009, in Kalahandi, Khorda and Malkangiri districts, Odisha, India; and in Dhading and Banke districts, Nepal; and 2010, in Kanpur Dehat and Pratapgarh districts, Uttar Pradesh, India; and Vaishali and Gaya districts, Bihar, India.

All locations were selected purposively, in agreement with local Non-Government Organizations (NGOs, listed in the acknowledgement) implementing a development project among their members, for which these surveys formed part of the baseline study.

We followed a two-stage sampling procedure: in stage one, villages were selected; 8 were selected in Warangal, 12 in Karimnagar and 42, 15 and 34 villages respectively in Kanpur Dehat, Pratapgarh and Vaishali (from a list given by the local NGO); 27, 22, 31, 9 and 17 villages were randomly sampled in Kalahandi, Khorda, Malkangiri, Dhading and Banke respectively (from lists provided by the local NGOs); and 50 villages were selected randomly in Gaya (from lists provided by the local NGO) plus 50 villages that were selected randomly from the Census 2001 registry of villages. In stage two, households were sampled randomly in each selected village by applying the “four winds” (or “line sampling”) technique [27] (except in Kanpur Dehat, Pratapgarh and Vaishali). The households were sampled in two cohorts of equal size: member and non-member households. Households were defined as “Members” if at least one person participated in a Self-Help-Group linked to partner-NGOs; other households were “Non-members”. We aggregated the two sub-cohorts for the purpose of the analysis reported here. In Kanpur Dehat, Pratapgarh and Vaishali stage two entailed inclusion of all member households, rather than sampling.

Sample sizes were 625 households (2464 individuals) in Warangal, 1089 (4384) in Karimnagar, 1805 (8258) in Kalahandi, 1758 (9110) in Khorda, 1597 (7448) in Malkangiri, 1000 (5275) in Dhading, 1008 (5741) in Banke, 1751 (10,220) in Kanpur Dehat, 1541 (9374) in Pratapgarh, 1922 (10,286) in Vaishali and 2006 households (13,320 individuals) in Gaya. It is noted that sampling was intended to represent the target population for the CBHI programme rather than the entire population in either state or district. 100% of the sampled households were rural. The overall response rate was 100% because there was always a willing adult respondent in the household.

The survey questionnaire was translated from English into the local languages (Telugu, Oriya, Hindi and Nepali, respectively for AP, Odisha, UP and Bihar, and Nepal), back translated for verification, and pre-tested among 80 households per language. Surveyors fluent in local dialects conducted the interviews. These surveyors were hired through local survey companies. The interviews were conducted orally at the houses of the sampled respondents as the respondents mostly were illiterate. The answers provided were recorded by the interviewers. Informed consent of the respondents was obtained and recorded prior to interviews. Participants’ names were kept confidential in data recording and analysis. The research tools were reviewed for ethical compliance by an ad-hoc advisory committee composed of scholars and senior scholars from India and other countries that held a two-day workshop in New Delhi in 2006 in preparation of the study, and discussed ethical compliance of the study. It is recalled that at the time of the rollout of the survey there was no local ethics committee in place in India or Nepal.

## 2.2. Data

We queried about general demographics and socio-economic status of households. Respondents to the household survey were asked, among others, to report the age, gender, education and economic activity of every household member. For our “income-proxy” we followed the method as adopted by the Indian National Sample Survey Organization [28] to note expenditures on many items of household consumption. Our income-proxy does not include health expenditure as recommended in previous studies related to determinants of health expenditure [29] [30].

Besides demographics and socio-economic status, we also queried about hospital admissions (exceeding 24 hours) in the year preceding the survey, as well as the direct medical costs related to each admission. These expenditures included the hospital bill, and other consumables from outside sources used during the hospitalization and directly related to it e.g. expenditures for medicines or tests. We also queried about costs of diagnostic tests performed in an outpatient setting in the month preceding the survey.

## 2.3. Method

We define the following variables:

- $X$  = the costs per event of an insurance benefit, such as hospitalization or tests;  
 $f(x)$  = the density function of  $X$ ;  
 $t$  = threshold value,  $c$  = cap value, we assume  $t < c$ ;  
 $E(X)$  = the expected costs of an insurance benefit;  
 $E_{0,c}(X)$  = the expected pay-out by an insurer for a policy covering the costs of an insurance benefit up to cap  $c$ ;  
 $E_{t,\infty}(X)$  = the expected pay-out by an insurer for a policy covering the costs of an insurance benefit above threshold  $t$ ;  
 $E_{t,c}(X)$  = the expected pay-out by an insurer for a policy covering the costs of an insurance benefit above threshold  $t$  and below cap  $c$ .

We assume the density function of healthcare expenditures to be a continuous distribution describing the healthcare expenditures per cost generating event. Previous research has shown that the distribution of positive healthcare expenditures per person per year can be approximated by a lognormal model [31]-[34]. We therefore investigate in this article to what extent a lognormal model can be used to approximate the distribution of the healthcare costs per event. The density function of a lognormal distribution  $f(x)$  is:

$$1) f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}$$

If  $X$  follows a lognormal distribution with parameters  $\mu$  and  $\sigma$ , the natural logarithm of  $X$  ( $\ln(X)$ ) follows a normal distribution with parameters  $\mu$  and  $\sigma$  and consequently, the variable  $Y = (\ln(X) - \mu)/\sigma$  follows the standard normal distribution.

Given a certain sample we first estimate the values for parameters  $\mu$  and  $\sigma$  of the lognormal distribution. For this, we only need the mean costs of the benefit and its standard deviation. We denote the mean costs by  $\bar{X}$  and the standard deviation by  $S$ . Then the parameters  $\mu$  and  $\sigma$  of the lognormal distribution are estimated as:

$$2) \sigma = \sqrt{\ln\left(\frac{S}{\bar{X}}\right)^2 + 1}$$

$$3) \mu = \ln(\bar{X}) - \frac{1}{2}\sigma^2$$

Insurance companies commonly introduce thresholds and/or caps for reimbursement to assure their long-term survival. We can also calculate the expected pay-out by the insurer with a certain cap or threshold (or both), after we have estimated the parameters  $\mu$  and  $\sigma$ . Subsequently,  $E_{0,c}(X)$ ,  $E_{t,\infty}(X)$  and  $E_{t,c}(X)$  would be calculated as follows:

$$4) E_{0,c}(X) = E(X) \cdot \Phi(d - \sigma) + c \cdot (1 - \Phi(d)),$$

$$5) E_{t,\infty}(X) = E(X) \cdot (1 - \Phi(e - \sigma)) - t \cdot (1 - \Phi(e)),$$

where  $d = \frac{\ln(c) - \mu}{\sigma}$ ,  $e = \frac{\ln(t) - \mu}{\sigma}$  and  $\Phi(z)$  is the cumulative density function of the standard normal distribution [33] [34].

In Formula (4), the first part equals the expected pay-out by the insurer when the costs of the insurance benefit was lower than the cap value; the second part equals the expected pay-out when the costs of the benefit was equal to the cap, or above it. The first part of Formula (5) equals the expected pay-out by the insurer when the benefit costs exceed the threshold. The second part of the formula equals the costs that are excluded from the insurance policy.

Subsequently we can calculate  $E_{t,c}(X)$  as:

$$6) E_{t,c}(X) = (E_{0,c}(X) + E_{t,\infty}(X)) - E(X)$$

The design of an insurance product with a cap and/or threshold requires understanding of the quantitative consequences of these measures. We use two outcome measures to assess the quantitative consequences of caps

and/or thresholds. The first is the percentage of the total costs that is reimbursed by the insurance (Formula (7)). This indicates the income required by the insurance company, assuming an actuarially fair premium, in order to break even.

7) Percentage of costs covered:

a) Situation with only a cap:  $\frac{E_{0,c}(X)}{E(X)}$ ;

b) Situation with only a threshold:  $\frac{E_{t,\infty}(X)}{E(X)}$ ;

c) Situation with both a threshold and a cap:  $\frac{E_{t,c}(X)}{E(X)}$ .

The second outcome measure is the percentage of bills covered by the insurance (*i.e.* percentage of bills below cap, above threshold and between threshold and cap in the situation with only a cap, only a threshold and both a threshold and a cap respectively, Formula (8)). This second outcome measure can serve as an indicator of the satisfaction of prospective clients.

8) Percentage of bills covered:

a) Situation with only a cap:  $\Phi(d)$ ;

b) Situation with only a threshold:  $1 - \Phi(e)$ ;

c) Situation with both a threshold and a cap:  $\Phi(d) - \Phi(e)$ .

## 2.4. Analysis

The unit of analysis is individual cases of hospitalizations and tests.

The analysis is based on hospitalizations with non-zero costs (95% of all hospitalizations reported) as those are the ones relevant for insurance payouts. For the analysis of tests, all outpatient tests had costs and were included.

We used STATA version 11 and MS Excel (version 2010).

All amounts, reported in Indian Rupee (INR) during the surveys were converted into international dollars (purchasing power parity, PPP\$) using the exchange rate of PPP\$1 = INR 16.130 for the survey conducted in 2008, INR 16.692 for the 2009 survey and 18.073 for the 2010 survey [35]. The amounts reported in Nepalese Rupee (NPR) were converted into PPP\$ using the exchange rate of PPP\$1 = NPR 29.222 [35].

## 3. Results

### 3.1. Socio-Economic Status of Studied Populations

The studied populations are characterized by low income and low educational levels (Table 1). The income-proxy of the population varies from PPP\$ 1.0 to PPP\$ 2.6 per person per day; and more than 40% of persons aged 15 years or more had no or very little education (below Class 5). At the same time, the percentage of people with higher education differed markedly across locations. Most people were engaged in self-employment in agriculture, daily wage labour and self-employment in business/trade.

### 3.2. Reported Costs for Hospitalizations and Tests

Information about expenditures for hospitalizations and tests is self-reported (Table 2). The incidence of hospitalizations ranges from 0.02 to 0.15, and for tests the range was between 0.02 and 0.10. The mean costs of hospitalizations and tests also differ markedly across the locations studied: a 6-fold difference in the mean costs of hospitalizations (from PPP\$109 in Malkangiri to PPP\$685 in Karimnagar) and a 5-fold difference in mean costs of tests (ranging from PPP\$6 in Malkangiri to PPP\$29 in Karimnagar). This large variation in healthcare costs across locations is in line with our previous findings in five other rural locations in India [16].

In this study, we want to examine whether the distributions of costs, not just cost levels, differ across locations. This could be done when the mean of the distributions could be standardized while maintaining the shape of each distribution. We “standardized” the costs by dividing reported costs in each location by that location’s av-

**Table 1.** Socio-economic profile of the studied populations.

	Warangal	Karimnagar	Kalahandi	Khorda	Malkangiri	Kanpur Dehat	Pratapgarh	Vaishali	Gaya	Dhading	Banke
Income-proxy per person per month (PPP\$) <sup>a</sup>											
Mean	77.99	74.49	27.79	33.41	27.98	59.76	47.28	56.58	43.84	70.58	56.87
Standard error	1.21	0.97	0.22	0.22	0.22	0.67	0.53	0.40	0.28	0.70	0.48
Education of adult population (15 years and older)											
No education	40.1	42.5	51.7	27.4	68.6	31.5	35.2	45.5	46.3	37.5	34.3
Class 1 - 5	9.1	8.4	16.8	20.3	13.5	11.8	12.5	14.5	12.3	18.8	13.8
Class 6 - 10	28.3	26.8	25.5	42.3	15.9	41.1	37.7	31.2	33.0	28.9	35.6
Class 11 and higher	22.5	22.2	6.1	10.1	2.0	15.6	14.6	8.9	8.4	14.8	16.2
Economic activity of adult population (15 years and older)											
Self-employed in agriculture	9.8	19.9	20.6	10.6	27.9	31.0	10.9	12.9	12.3	31.8	24.06
Self-employed in business/trade	10.0	7.5	5.4	17.1	3.7	4.0	7.7	9.8	3.6	14.17	8.35
Regular salaried employee	7.0	4.5	2.2	4.2	2.6	2.7	3.8	1.6	2.9	8.39	11.2
Casual wage labourer	39.7	31.5	26.3	12.5	27.5	8.1	15.2	19.5	33.0	4.21	8.91
Domestic duties	10.9	10.6	29.5	38.3	28.4	31.5	32.7	35.5	28.4	13.6	21.66
Not working	22.5	25.9	15.9	17.4	9.9	22.7	29.7	20.8	19.9	27.83	25.82

<sup>a</sup>Income is proxied as monthly per capita consumer expenditure through questions on many items of household expenditure and expressed in Purchasing Power Parity International Dollar.

**Table 2.** Socio-economic profile of the studied populations.

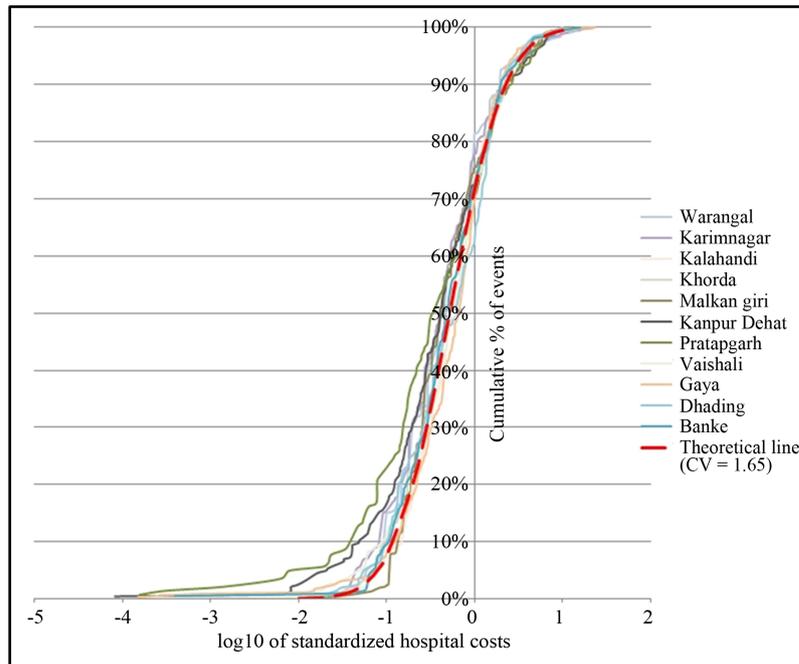
	Hospitalizations				Diagnostic tests			
	Incidence of cases with cost > 0 (n)	Mean costs (PPP\$ <sup>a</sup> )	Standard deviation (PPP\$ <sup>a</sup> )	Coefficient of variation (CV)	Incidence of cases with cost > 0 (n)	Mean costs (PPP\$ <sup>a</sup> )	Standard deviation (PPP\$ <sup>a</sup> )	Coefficient of variation (CV)
Warangal	0.15 (377)	625.49	1267.03	2.03	0.05 (124)	23.12	43.98	1.90
Karimnagar	0.15 (654)	685.52	1414.20	2.06	0.06 (271)	28.86	41.57	1.44
Kalahandi	0.06 (469)	183.34	268.93	1.47	0.10 (799)	5.76	12.41	2.15
Khorda	0.06 (523)	345.71	563.96	1.63	0.05 (490)	9.67	13.77	1.42
Malkangiri	0.06 (468)	108.94	185.62	1.70	0.04 (304)	5.53	5.28	0.95
Kanpur Dehat	0.02 (245)	681.91	1129.76	1.66	0.03 (328)	20.52	31.88	1.55
Pratapgarh	0.02 (143)	356.44	588.56	1.65	0.02 (196)	14.50	24.95	1.72
Vaishali	0.03 (331)	272.92	369.14	1.35	0.06 (605)	18.84	30.01	1.59
Gaya	0.03 (353)	372.10	610.53	1.64	0.02 (272)	13.75	18.00	1.31
Dhading	0.04 (204)	502.31	657.07	1.31	0.03 (139)	23.28	30.40	1.31
Banke	0.04 (218)	437.58	717.58	1.64	0.04 (255)	14.04	21.37	1.52

<sup>a</sup>PPP\$ = Purchasing Power Parity International Dollar.

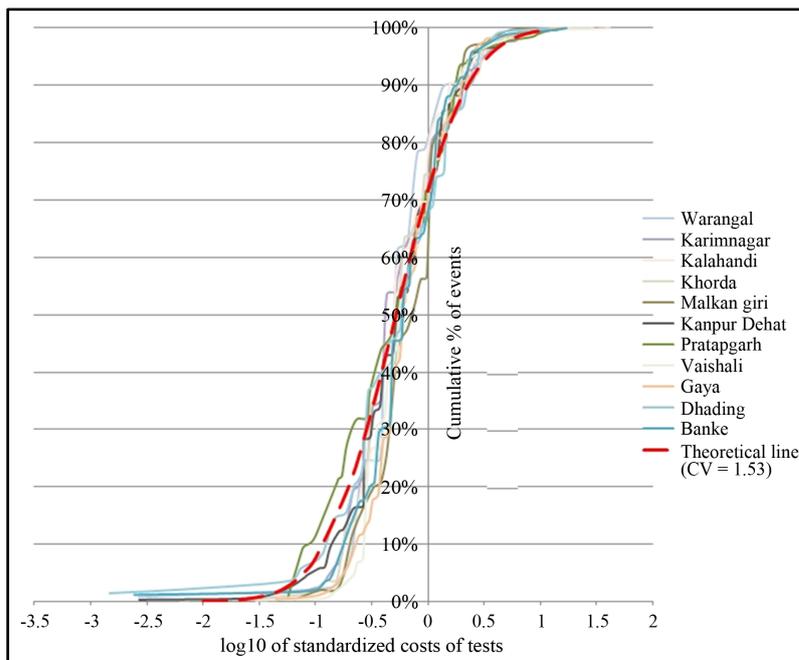
erage cost, obtaining the same mean cost of one in all locations, while the Coefficient of Variation (CV) of each distribution remained unchanged. These standardized distributions are shown in **Figure 1** (for hospital expenditures) and in **Figure 2** (for expenditures on tests).

To our surprise, the standardized distributions of hospital expenditures in the 11 locations were really similar; the same clear similarity across locations was also observed with the distributions of expenditures on tests.

We then proceeded to elaborate a model describing the best-fit for all distributions. It is recalled that previous research has successfully used the lognormal distribution as a good approximation for empirical distributions of



**Figure 1.** Distributions of hospital costs.



**Figure 2.** Distributions of costs on diagnostic tests.

healthcare costs [31]–[34]. We therefore created a theoretical lognormal curve (Equation (1)) with mean of one and a CV which is the arithmetic average for the 11 locations (1.65 for hospital costs and 1.53 for tests, **Table 2**). The resulting theoretical curve is plotted in **Figure 1** for hospitalization expenditures and in **Figure 2** for expenditures on tests.

### 3.3. Outcome Measures as Predicted by the Theoretical Model and Compared with Empirical Data

The model needs to predict the financial performance of CBHI schemes functioning with different levels of caps and thresholds. We observed the consequences of different levels of caps and thresholds on 1) the percentage of over all costs that the insurance would reimburse; and 2) the percentage of bills covered by the insurance.

As a reminder, all bills and all costs are reimbursable fully only when neither threshold nor cap apply. Moreover, most insured persons do not claim at all (and these would not be shown in the distributions); and many bills entail low costs, with only very few very high cost bills. Both a threshold and a cap will be effective in reducing the average reimbursement per insured. In many cases, the insurance product is limited both by a threshold and by a cap.

**Tables 3** and **4** contain both the results obtained by applying our theoretical model and those obtained by applying the same caps/thresholds to the empirical data derived from the household surveys in all 11 locations. It should be noted that these caps/thresholds are per event (*i.e.* hospitalization or test).

We observe marked similarity between the theoretical model and the empirical data in the prediction of the percentage of total costs that the insurance needs to cover. For instance, when the cap is assumed to equal 100% of the mean, the predicted outcome was less than 10% different from the empirical outcome (“good agreement”) in 8 out of the 11 locations (**Table 3**). In two more examples (when the cap was assumed to be 0.5 times the mean, and when the cap was assumed to be 2 times the mean), we observed good agreement in 9 out of the 11 locations. And, the theoretical and empirical values were in good agreement in all locations when comparing the effect of introducing a threshold without a cap (two examples shown: thresholds of 10% or 20% of the mean).

A comparison of predicted versus empirical data on the % of bills also yielded a marked similarity when the caps were set at 100% or 200% of the mean (good agreement in 9 and 11 cases out of 11, respectively). On the other hand, the agreement was lower, with only 5 out of 11 locations in good agreement, when the cap was assumed to be 50% of the mean. With a threshold of 10% of the mean, 10 out of 11 cases were in good agreement and a threshold of 20% of the mean 7 out of 11 were in good agreement.

A similar pattern was observed when comparing the theoretical model with the empirical data relating to tests (**Table 4**).

## 4. Discussion

This article deals with developing a model to estimate insurance coverage under CBHI. If a threshold and/or a cap limits the payout of any benefit then the evaluation of coverage can be done only when the distribution of the expected costs is known. Because thresholds and caps are not rare but a standard feature of benefit packages, we are keen to develop a method to predict the distribution in each case. A predictive method has particularly important ramifications in the context of introducing health insurance to rural poor persons in low income countries, where data is usually missing and collection of the required data entails high cost, which often cannot be borne as part of implementation costs in these settings.

In this study, we examine the distributions of the costs of two benefits (hospitalizations and diagnostic tests) derived from eleven household surveys conducted prior to launching CBHI schemes (nine in India and two in Nepal). As the mean cost of the benefits differed significantly across the locations, we wished to study the shape of the distributions independently of the mean costs. The shape of the distribution is determined by the relative frequency of small bills versus high/outlier bills: for two data series with the same mean costs but with very different ratios of small bills to large bills, the consequences of applying the same threshold and/or cap could be completely different in terms of insurance coverage.

We were not surprised to find in this study large differences across locations in incidence/prevalence of illness, utilization of healthcare, its cost and different levels of WTP for health insurance (**Table 2**), as this corroborated previous relevant research [15]–[17]. On the other hand, when—in our search for a theoretical model of the distributions—we “standardized” the eleven data series by dividing costs reported in each location by that loca-

**Table 3.** Theoretical model vs. empirical data for hospitalizations.

		Cap			Threshold	
Theoretical model	Fraction of mean	0.5	1.0	2.0	0.2	0.1
(Assumed CV <sup>a</sup> = 1.65)	% of cost covered	38%	57%	75%	82%	90%
	% of bills covered <sup>b</sup>	49%	72%	88%	80%	92%
	Warangal	Cap/threshold (PPP\$ <sup>c</sup> )	313	625	1250	125
	% of cost covered	35%	53%	67%	82%	90%
	% of bills covered <sup>b</sup>	57%	81%	93%	69%	85%
Karimnagar	Cap/threshold (PPP\$ <sup>c</sup> )	343	685	1371	137	69
	% of cost covered	34%	50%	66%	83%	90%
	% of bills covered <sup>b</sup>	58%	78%	89%	69%	85%
Kalahandi	Cap/threshold (PPP\$ <sup>c</sup> )	92	183	366	36	18
	% of cost covered	39%	59%	77%	82%	90%
	% of bills covered <sup>b</sup>	46%	74%	89%	82%	92%
Khorda	Cap/threshold (PPP\$ <sup>c</sup> )	173	346	691	69	35
	% of cost covered	37%	56%	75%	82%	90%
	% of bills covered <sup>b</sup>	46%	71%	88%	78%	90%
Malkangiri	Cap/threshold (PPP\$ <sup>c</sup> )	54	109	218	22	11
	% of cost covered	36%	53%	71%	81%	90%
	% of bills covered <sup>b</sup>	56%	75%	87%	78%	98%
Kanpur Dehat	Cap/threshold (PPP\$ <sup>c</sup> )	341	682	1364	136	68
	% of cost covered	33%	50%	69%	83%	91%
	% of bills covered <sup>b</sup>	57%	72%	89%	69%	84%
Pratapgarh	Cap/threshold (PPP\$ <sup>c</sup> )	178	356	713	71	36
	% of cost covered	31%	50%	69%	84%	91%
	% of bills covered <sup>b</sup>	57%	70%	87%	62%	77%
Vaishali	Cap/threshold (PPP\$ <sup>c</sup> )	136	273	546	55	27
	% of cost covered	38%	58%	77%	82%	91%
	% of bills covered <sup>b</sup>	46%	70%	87%	80%	89%
Gaya	Cap/threshold (PPP\$ <sup>c</sup> )	186	372	744	74	37
	% of cost covered	39%	61%	79%	82%	90%
	% of bills covered <sup>b</sup>	42%	70%	90%	82%	92%
Dhading	Cap/threshold (PPP\$ <sup>c</sup> )	251	502	1005	100	50
	% of cost covered	36%	59%	80%	82%	90%
	% of bills covered <sup>b</sup>	48%	63%	87%	75%	91%
Banke	Cap/threshold (PPP\$ <sup>c</sup> )	219	438	875	88	44
	% of cost covered	37%	56%	74%	82%	90%
	% of bills covered <sup>b</sup>	51%	72%	90%	77%	91%

<sup>a</sup>CV = Coefficient of Variation; <sup>b</sup>Percentage of bills covered by the insurance: percentage of bills below cap or above threshold in the situation with only a cap or only a threshold respectively; <sup>c</sup>PPP\$ = Purchasing Power Parity International Dollar.

**Table 4.** Theoretical model vs. empirical data for outpatient diagnostic tests.

		Cap			Threshold	
Theoretical model (Assumed CV <sup>a</sup> = 1.53)	Fraction of mean	0.5	1.0	2.0	0.2	0.1
	% of cost covered	39%	58%	77%	80%	90%
	% of bills covered <sup>b</sup>	47%	71%	88%	80%	94%
Warangal	Cap/threshold (PPP\$ <sup>c</sup> )	11.6	23.1	46.2	4.6	2.3
	% of cost covered	42%	58%	72%	80%	90%
	% of bills covered <sup>b</sup>	41%	79%	90%	94%	99%
Karimnagar	Cap/threshold (PPP\$ <sup>c</sup> )	14.4	28.9	57.7	5.8	2.9
	% of cost covered	40%	58%	76%	80%	90%
	% of bills covered <sup>b</sup>	54%	70%	86%	92%	99%
Kalahandi	Cap/threshold (PPP\$ <sup>c</sup> )	2.9	5.8	11.5	1.2	0.6
	% of cost covered	44%	62%	75%	79%	90%
	% of bills covered <sup>b</sup>	30%	71%	91%	88%	98%
Khorda	Cap/threshold (PPP\$ <sup>c</sup> )	4.9	9.7	19.3	1.9	1
	% of cost covered	41%	60%	77%	81%	90%
	% of bills covered <sup>b</sup>	43%	74%	90%	87%	98%
Malkangiri	Cap/threshold (PPP\$ <sup>c</sup> )	2.8	5.5	11	1.1	0.6
	% of cost covered	45%	70%	89%	80%	89%
	% of bills covered <sup>b</sup>	26%	57%	88%	97%	98%
Kanpur Dehat	Cap/threshold (PPP\$ <sup>c</sup> )	10.3	20.5	41	4.1	2.1
	% of cost covered	40%	60%	77%	81%	90%
	% of bills covered <sup>b</sup>	43%	70%	89%	86%	95%
Pratapgarh	Cap/threshold (PPP\$ <sup>c</sup> )	7.3	14.5	29	2.9	1.4
	% of cost covered	37%	57%	75%	82%	91%
	% of bills covered <sup>b</sup>	48%	72%	93%	72%	89%
Vaishali	Cap/threshold (PPP\$ <sup>c</sup> )	9.4	18.8	37.7	3.8	1.9
	% of cost covered	42%	64%	81%	80%	90%
	% of bills covered <sup>b</sup>	40%	69%	91%	97%	100%
Gaya	Cap/threshold (PPP\$ <sup>c</sup> )	6.9	13.7	27.5	2.7	1.4
	% of cost covered	44%	65%	84%	81%	90%
	% of bills covered <sup>b</sup>	39%	68%	85%	96%	99%
Dhading	Cap/threshold (PPP\$ <sup>c</sup> )	11.6	23.3	46.6	4.7	2.3
	% of cost covered	38%	60%	81%	81%	90%
	% of bills covered <sup>b</sup>	45%	66%	86%	84%	94%
Banke	Cap/threshold (PPP\$ <sup>c</sup> )	7	14	28	2.8	1.4
	% of cost covered	43%	64%	80%	81%	90%
	% of bills covered <sup>b</sup>	45%	67%	91%	95%	98%

<sup>a</sup>CV = Coefficient of Variation; <sup>b</sup>Percentage of bills covered by the insurance: percentage of bills below cap or above threshold in the situation with only a cap or only a threshold respectively; <sup>c</sup>PPP\$ = Purchasing Power Parity International Dollar.

tion's average cost (a practice common in natural sciences) we were surprised to see that the shape of the distributions was remarkably similar. This led us to develop a theoretical model based on the assumption that the distributions are lognormal [31]-[34] and its shape is determined by the Coefficient of Variation (CV = standard deviation/mean). The CV of the theoretical model was assumed to be the arithmetic average of the CVs of the eleven locations.

Previous research has shown that the distribution of positive healthcare expenditures per person per year can be fitted by a lognormal model [31]-[34]. Several researchers have investigated the influence of different covariates on healthcare costs in order to find ways to make an even better fit of the lognormal model [36] [37]. The purpose of this article is to estimate the distribution of locally relevant healthcare costs without cumbersome and expensive data collection. In the context of this article we therefore don't make use of that knowledge as this would require more data inputs which are not (easily) at hand in rural India and would therefore diminish the advantage we seek in the modelling.

When we compared the theoretical model against the predications based on empirical data we obtained satisfactory predictions of the consequences of applying thresholds and/or caps on the expected financial performance of CBHI, both in the case of hospitalizations and tests (Tables 3 and 4). The outcome measures were the percentage of bills covered and the percentage of costs reimbursed by the insurance under different assumptions for caps and/or thresholds. Given this satisfactory comparison between empirical data and the theoretical model we were strengthened in our belief that the lognormal model can provide a good approximation for this data.

The arithmetic averages of the CVs of the two benefits under study were remarkably similar even though one benefit was related to inpatient care and the other to a component of outpatient care. We do not know if this similarity might apply also to other benefits, or is a coincidence. This would have to be examined against the distributions of additional cost components of outpatient care.

Applying the theoretical model requires an estimate of the mean cost of the benefit(s) in question. We believe that it should be easy enough to obtain mean costs in each new location as part of the investigation to estimate the probability of healthcare utilization based on facilitated interactive group discussions which we developed, called "Illness Mapping" [26].

In designing health insurance there is always the tension between the desire to include more benefits and the wish to keep the premium amount affordable. The reasons for introducing a threshold and/or a cap seem rather straight forward: thresholds can reduce the administrative burden of many small claims, caps can protect the insurance from very high outlier bills. However, both measures could reduce the willingness to join, especially when insured persons must still borrow money or sell assets to pay for healthcare bills [3]. Setting caps and/or thresholds is therefore a delicate issue. The percentage of bills covered can be used as an outcome measure that could serve as an indicator of the expected satisfaction of prospective clients.

The theoretical model developed in this paper is based on nine Indian and two Nepalese locations. These eleven locations are very different in mean cost and probability of healthcare utilization of the studied benefits (Table 2). The satisfactory fit of the empirical distributions with the distributions derived on the basis of the theoretical model supports the notion that the model could be applicable for all Indian and Nepalese rural poor cohorts. More generally, the theoretical model presented here makes it possible to simulate the expected performance of the CBHI in the local context at the stage of package design, and determine not only which components should be included but also the optimal level of threshold and cap that would meet the expressed levels of WTP best. To the best of our knowledge, this has never before been done in development projects aiming to develop CBHI with suitable health insurance coverage under severe rationing rules.

A look at the distributions in the eleven locations studied (both of hospitalizations and tests) makes clear that the distributions are less similar at the left tail. This scattering in the range of the small bills may reflect a variation in health seeking behaviour in less severe illnesses (due to healthcare availability/proximity and/or education). It may also be caused by recall bias which can be expected to be more pronounced in the case of small bills [38]. This problem could be reduced by applying a threshold.

It would be interesting to replicate this research in other developing countries where CBHI schemes are implemented or considered, in order to establish whether the distributions of costs follow a similar pattern.

## 5. Conclusion

The analysis in this paper has shown that a theoretical model of healthcare costs based on a lognormal distribution

can be used to calculate the impact of thresholds and caps on the share of costs and bills that would be reimbursable. This information, hitherto ignored, can be used to refine premium calculations of community-based health insurance schemes in low income countries and foretell expected client satisfaction with locally relevant health insurance packages that define a threshold and/or a cap. Application of the model makes it possible to obtain the local data required faster and cheaper than by conducting household surveys.

## Acknowledgements

The authors gratefully acknowledge funding support from the Netherlands Organization for Scientific Research (NWO), under WOTRO Integrated Programme grant No. W01.65.309.00 to Prof. Dror. This grant funded the household surveys in Warangal, Karimnagar, and Gaya in full, plus partial funding of the household survey in Malkangiri, Kalahandi and Khorda, for which additional funding was obtained from the German Federal Ministry for Economic Cooperation and Development (through a contract between AWO International, Madhyam Foundation Odisha, and the Micro Insurance Academy New Delhi). The household survey in Nepal (Dhading and Banke districts) was funded through a grant by the Rockefeller Foundation (2008 PSH 229) to Prof. Dror plus another grant by an anonymous donor to Save the Children USA/Nepal. The household surveys in Kanpur Dehat, Pratapgarh and Vaishali were funded by the European Commission 7th Framework Program, grant ID HEALTH-F2-2009-223518. The sponsors had no influence or role in study design, in the collection, analysis and interpretation of data; in the writing of the article; and in the decision where to submit the article for publication. We benefited from logistical and research support from the Micro Insurance Academy, New Delhi, India. The following organizations provided logistical support at the research sites: CDF in Warangal and Karimnagar; Madhyam Foundation (comprising 11 grassroots NGOs: Parivartan, PUSPAC, SOMKS, SDS, ODC (Malkangiri); Mahashakti Foundation, DAPTA, Lok Yojana, Sanginee (Kalahandi); MVPS, DSS (Khorda)); Shramik Bharti in Kanpur Dehat; BAIF in Pratapgarh; Nidan in Vaishali; BASIX in Gaya; DEPROSC in Dhading; and NIRDHAN in Banke. Last but not least, we acknowledge all the surveyors and the respondents for their participation in the study.

## References

- [1] World Development Indicators (2012) <http://data.worldbank.org/data-catalog/world-development-indicators>
- [2] Karan, A.K. and Selvaraj, S. (2012) Why Publicly-Financed Health Insurance Schemes Are Ineffective in Providing Financial Risk Protection. *Economic and Political Weekly*, **47**.
- [3] Binnendijk, E., Koren, R. and Dror, D.M. (2012) Hardship Financing of Healthcare among Rural Poor in Orissa, India. *BMC Health Services Research*, **12**, 23.
- [4] Kruk, M.E., Goldmann, E. and Galea, S. (2009) Borrowing and Selling to Pay for Health Care in Low- and Middle-Income Countries. *Health Affairs (Millwood)*, **28**, 1056-1066. <http://dx.doi.org/10.1377/hlthaff.28.4.1056>
- [5] James, C. and Savedoff, W. (2010) Risk Pooling and Redistribution in Health Care: An Empirical Analysis of Attitudes toward Solidarity. World Health Report 2010, World Health Organization, Geneva.
- [6] World Health Organization (2010) Health Systems Financing: The Path to Universal Coverage. The World Health Report 2010, World Health Organization, Geneva.
- [7] Staib, D. and Bevere, L. (2011) World Insurance in 2010: Premiums Back to Growth, Capital Increases. *Swiss Re Sigma*, **2011**, 1-44.
- [8] Ma, S. and Sood, N. (2008) A Comparison of the Health Systems in China and India. Rand Corporation, Center for Asia Pacific Policy, Santa Monica.
- [9] Ahuja, R. (2005) Health Insurance for the Poor in India: An Analytical Study. Indian Council for Research on International Economic Relations (ICRIER), New Delhi.
- [10] Bhat, R. and Jain, N. (2006) Factoring Affecting the Demand for Health Insurance in a Micro Insurance Scheme. Indian Institute of Management, Ahmedabad.
- [11] Dror, D.M., Radermacher, R., Khadilkar, S.B., Schout, P., Hay, F.X., Singh, A. and Koren, R. (2009) Microinsurance: Innovations in Low-Cost Health Insurance. *Health Affairs (Millwood)*, **28**, 1788-1798. <http://dx.doi.org/10.1377/hlthaff.28.6.1788>
- [12] National Commission on Macroeconomics and Health (2005) Financing and Delivery of Healthcare Services in India. National Commission on Macroeconomics and Health Background Papers, Ministry of Health and Family Welfare,

Government of India, New Delhi.

- [13] Dror, D.M. and Jacquier, C. (1999) Micro-Insurance: Extending Health Insurance to the Excluded. *International Social Security Review*, **52**, 71-97. <http://dx.doi.org/10.1111/1468-246X.00034>
- [14] Dror, D.M. (2007) Why “One-Size-Fits-All” Health Insurance Products Are Unsuitable for Low-Income Persons in the Informal Economy in India. *Asian Economic Review*, **49**, 47-56.
- [15] Dror, D.M., van Putten-Rademacher, O. and Koren, R. (2009) Incidence of Illness Among Resource-Poor Households: Evidence from Five Locations in India. *Indian Journal of Medical Research*, **130**, 146-154.
- [16] Dror, D.M., van Putten-Rademacher, O. and Koren, R. (2008) Cost of Illness: Evidence from a Study in Five Resource-Poor Locations in India. *Indian Journal of Medical Research*, **127**, 347-361.
- [17] Dror, D.M., Rademacher, R. and Koren, R. (2007) Willingness to Pay for Health Insurance among Rural and Poor Persons: Field Evidence from Seven Micro Health Insurance Units in India. *Health Policy*, **82**, 12-27. <http://dx.doi.org/10.1016/j.healthpol.2006.07.011>
- [18] Binnendijk, E., Dror, D.M., Gerelle, E. and Koren, R. (2013) Estimating Willingness-to-Pay for Health Insurance among Rural Poor in India by Reference to Engel’s Law. *Social Science & Medicine*, **76**, 67-73. <http://dx.doi.org/10.1016/j.socscimed.2012.10.006>
- [19] Dror, D.M. and Koren, R. (2012) The Elusive Quest for Estimates of Willingness to Pay for Micro Health Insurance among the Poor in Low-Income Countries. In: Churchill, C. and Matul, M., Eds., *Protecting the Poor—A Micro Insurance Compendium Volume 2. Volume Chapter 7*. International Labour Organization, Geneva, 156-173.
- [20] Doyle, C., Panda, P., Van de Poel, E., Rademacher, R. and Dror, D.M. (2011) Reconciling Research and Implementation in Micro Health Insurance Experiments in India: Study Protocol for a Randomized Controlled Trial. *Trials*, **12**, 224. <http://dx.doi.org/10.1186/1745-6215-12-224>
- [21] Dong, H., Mugisha, F., Gbangou, A., Kouyate, B. and Sauerborn, R. (2004) The Feasibility of Community-Based Health Insurance in Burkina Faso. *Health Policy*, **69**, 45-53. <http://dx.doi.org/10.1016/j.healthpol.2003.12.001>
- [22] AC Nielsen ORG-MARG Pvt Ltd. (2001) Needs and Demands for Healthcare and Health Insurance among Landless Agriculture Laborers in Burdwan District, West Bengal. Final Report Submitted to GTZ Technical Assistance Team, Health Sector Support, AC Nielsen ORG-MARG Pvt Ltd., Kolkata.
- [23] De Allegri, M., Pokhrel, S., Becher, H., Dong, H., Mansmann, U., Kouyate, B., Kynast-Wolf, G., Gbangou, A., Sanon, M., Bridges, J. and Sauerborn, R. (2008) Step-Wedge Cluster-Randomised Community-Based Trials: An Application to the Study of the Impact of Community Health Insurance. *Health Research Policy and Systems*, **6**. <http://dx.doi.org/10.1186/1478-4505-6-10>
- [24] Mwaura, J.W. and Pongpanich, S. (2012) Access to Health Care: The Role of a Community Based Health Insurance in Kenya. *The Pan African Medical Journal*, **12**. <http://www.panafrican-med-journal.com/content/article/12/35/full>
- [25] Onwujekwe, O., Onoka, C.A., Uguru, N., Nnenna, T., Uzochukwu, B.S., Eze, S., Kirigia, J. and Petu, A. (2010) Preferences for Benefit Packages for Community-Based Health Insurance: An Exploratory Study in Nigeria. *BMC Health Services Research*, **10**, 162. <http://dx.doi.org/10.1186/1472-6963-10-162>
- [26] Binnendijk, E., Gautham, M., Koren, R. and Dror, D.M. (2012) Illness Mapping: A Time and Cost Effective Method to Estimate Healthcare Data Needed to Establish Community Based Health Insurance. *BMC Medical Research Methodology*, **12**, 153.
- [27] Som, R.K. (1996) *Practical Sampling Techniques*. 2nd Edition, Marcel Dekker, New York.
- [28] National Sample Survey Organization (2008) Household Consumer Expenditure in India, 2006-07. Ministry of Statistics and Programme Implementation, Government of India, New Delhi.
- [29] Flores, G., Krishnakumar, J., O’Donnell, O. and van Doorslaer, E. (2008) Coping with Health-Care Costs: Implications for the Measurement of Catastrophic Expenditures and Poverty. *Health Economics*, **17**, 1393-1412. <http://dx.doi.org/10.1002/hec.1338>
- [30] Wagstaff, A. (2008) *Measuring Financial Protection in Health*. World Bank, Washington DC.
- [31] Newhouse, J.P. and The Insurance Experiment Group (1993) *Free for All? Lessons from the RAND Health Insurance Experiment*. Harvard University Press, Cambridge.
- [32] Van der Laan, B.S. (1988) *Modelling Total Costs of Claims of Non-Life Insurances: With Applications to Health Insurance and to Motor Insurance*. Erasmus University, Rotterdam.
- [33] Van Vliet, R.C. (2000) A Statistical Analysis of Mandatory Pooling across Health Insurers. *Journal of Risk and Insurance*, **67**, 197-217. <http://dx.doi.org/10.2307/253771>
- [34] Van Vliet, R.C. (2004) Deductibles and Health Care Expenditures: Empirical Estimates of Price Sensitivity Based on Administrative Data. *International Journal of Health Care Finance and Economics*, **4**, 283-305.

<http://dx.doi.org/10.1023/B:IHFE.0000043759.93644.e0>

- [35] World Economic Outlook Database (2012)  
<http://www.imf.org/external/pubs/ft/weo/2012/01/weodata/index.aspx>
- [36] Manning, G.W., Basu, A. and Mullahy, J. (2005) Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data. *Journal of Health Economics*, **24**, 465-488. <http://dx.doi.org/10.1016/j.jhealeco.2004.09.011>
- [37] Gilleskie, D.B. and Mroz, T.A. (2004) A Flexible Approach for Estimating the Effects of Covariates on Health Expenditures. *Journal of Health Economics*, **23**, 391-418. <http://dx.doi.org/10.1016/j.jhealeco.2003.09.008>
- [38] Harel, Y., Overpeck, M.D., Jones, D.H., Scheidt, P.C., Bijur, P.E., Trumble, A.C. and Anderson, J. (1994) The Effects of Recall on Estimating Annual Nonfatal Injury Rates for Children and Adolescents. *American Journal of Public Health*, **84**, 599-605. <http://dx.doi.org/10.2105/AJPH.84.4.599>