Determinants of Childhood Diarrhea in a Highly Heterogeneous Urban Context: the Case of Dakar – Senegal

Iulia Rautu, *Aspirant FNRS – Université catholique de Louvain*
Stéphanie Dos Santos, *Institut de recherche pour le développement*
Bruno Schoumaker, *Université catholique de Louvain*

Abstract

In Senegal, diarrheal infections constitute one of the most significant urban health issues, affecting primarily children. In the capital, Dakar, the rapid urbanization of recent decades lead to high population densities in a very heterogeneous urban space. Thus, children from Dakar are exposed to different risk factors for diarrheal diseases, depending on the area they live in. In this context, this paper analyzes the relative role of the determinants of childhood diarrhea in the highly heterogeneous urban context of Dakar. It uses a three-level statistical analysis, in order to differentiate the part of factors relating to the individual, the family and the immediate environment. We also illustrate a relatively new statistical approach in demography – latent class analysis (LCA) – which allows to synthesize a large number of indicators, in order to estimate these multi-dimensional determinants. This helps to create a comprehensive causal image, while using parsimonious statistical models.

Keywords: urban health, diarrheal diseases, child health, Subsaharan Africa

Introduction

Worldwide, diarrhea continues to represent a significant cause of morbidity for all ages (Pond et al., 2004, cited by Regassa et al., 2008). For children in developing countries, it is one of the main sources of disease burden, being responsible for almost 12% of total disability-adjusted life years in children aged 1 to 4 and for 5.71% in children aged 5 to 9 years old (Institute for Health Metrics and Evaluation, 2013).

In Senegal, diarrheal infections constitute one of the most significant urban health issues, affecting primarily children. In the capital, Dakar, the rapid urbanization of recent decades was often unplanned, leading to high population densities in a very heterogeneous urban space. Thus, the urban space created a varied environment: within the same city, access to education and health care or sanitation is often very diverse. This, in turn, can affect the health of individuals differently, putting certain categories more at risk than others. This means that children from Dakar are exposed to different risk factors for diarrheal diseases, depending on the area they live in.

Moreover, a significant proportion of the population lives in slums, in overcrowded dwellings lacking access to basic sanitation. This situation, in the context of sometimes suboptimal behaviors in terms of hygiene, poses additional threats in terms of urban health in general and of diarrheal diseases in particular.

From this perspective, the first main objective of this paper is to analyze the relative role of the determinants of childhood diarrhea in the highly heterogeneous urban context of a large city, namely Dakar, the capital of Senegal. The paper also aims to distinguish the part of factors relating to the individual, the family and the immediate environment. In doing so, we illustrate a relatively new statistical approach in demography, latent class analysis (LCA). This approach allows to synthesize the network of these determinants, in order to create a comprehensive causal image, while avoiding the inclusion of too many explanatory variables and thus yielding statistical models that are parsimonious and easily interpretable.

Background

Diarrhea refers to passing loose or liquid stools 3 or more times a day (or more frequently than normal for the individual), lasting from a few days to several weeks and it is a common symptom of several possible gastrointestinal infections (WHO, 2009). These include bacteria, viruses, protozoa and helminths (Rego et al., 2005).

Diarrheas belong to the waterborne diseases category, meaning that they are infectious diseases caused by pathogenic microbes that can be directly spread through contaminated water (Centers for Disease Control and Prevention, 2012).

Infectious diseases in general, and diarrheal diseases in particular, are often studied within the so-called "disease triangle" paradigm, which can be used to predict epidemiological outcomes in public health. It illustrates the interactions between the environment, the host and an infectious agent (Fig. 1):

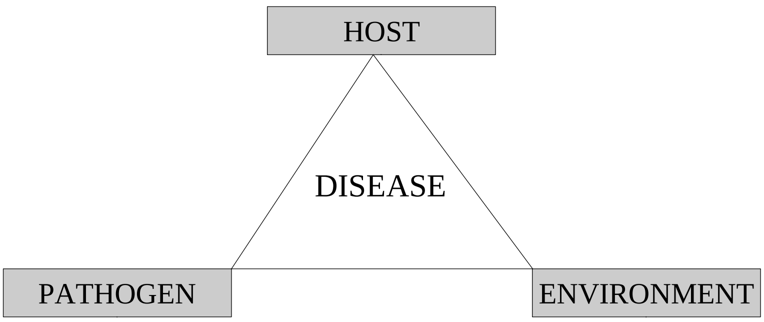

Figure 1: The Epidemiologic Triad

The disease triangle is a conceptual model that conditions the occurrence of a disease to three variables: a host organism (in our case, the child), the pathogen (a bacteria, virus, parasite, etc., that cause diarrheal diseases) and a conducive environment, which favors the infection. However, the third dimension is, although assumed, often overlooked (Scholthof, 2007).

Many papers focus only on the relationship between the first two components, the pathogen and the host. Moreover, the studies which do look at the environmental component, usually do not control for other factors, such as behaviors or socio-economic characteristics.

However, due to their waterborne dimension, diarrheal diseases are susceptible to be influenced by a diverse range of determinants from different spheres: characteristics of the individual, the family/household and particularities of the immediate environment (Fewtrell and Colford, 2004; Masangwi et al., 2009; Woldemicael, 2011).

Certain types of behavior, particularly related to personal hygiene, can influence exposure to pathogens (Waddington et al., 2009). These behavioral aspects become particularly important in the case of highly endemic environments. Nevertheless, assessing behavioral compliance often proves to be a challenging task: on the one hand, behaviors are generally self-reported, which is conducive to bias; on the other hand, implementing certain protective behaviors depends of other factors, such as education or wealth (Jalan and Ravallion, 2003).

Aspects like access to clean water and proper sanitation can provide barriers to transmission (Waddington et al., 2009). However, cities in developing countries often face a wide diversity in terms of living conditions and a particularly underprivileged situation is that of slums. This creates further inequity between categories of population living in the same larger area. Thus, within the same city, cases are repeatedly reported in the same areas, which indicates a relatively high level of clustering of the disease (Myaux et al., 1997).

The city Dakar is the capital of Senegal and it characterized by tropical climate with two seasons: the dry season and the wet season. Although Senegal is one of the poorest countries in the world and, in 2014 ranked 163 of 187 on the Human Development Index, the Dakar region tends to be relatively wealthy compared to the rest of the country, constituting by far the main area of economic development in the country and attracting businesses and workers from throughout the West African region. This contributes to a relatively high rate of population growth, which between 2002 and 2008 was on average of 2,8% (Service Régional de la Statistique et de la Démographie de Dakar, 2009). In 2008, the population of Dakar was estimated at 2.482.294 inhabitants, which represents almost a quarter of the country's entire population (Service Régional de la Statistique et de la Démographie de Dakar, 2009). This population is highly heterogeneous, both in terms of socio-economic characteristics and in terms of living conditions. This, added to the considerable diversity from one area to another within the city, creates a wide and varied panorama of potențial risk and protective factors with regard to diarrheal diseases.

Data and Methods

The research uses the hierarchical quantitative data of the project ACTU-PALU, conducted in 50 districts of Dakar and its suburbs, and coordinated by the Institute for Research and Development (IRD) in France, in collaboration with the Cheikh Anta Diop University of Dakar. Data collection took place between 15th of September and the 22nd of December of

2008, by means of questionnaires. The representative sample of 2952 households include 7416 children aged 2 to 10 (Lalou, 2008).

The cross-sectional survey was conducted during and just after the rainy season of 2008, a period with floods and thus with higher transmission of infectious diseases. The questionnaire covers aspects of behavior, socio-demographic characteristics, characteristics of the dwelling and of the area, as well as the occurrence of an episode of diarrhea in children aged 2 to 10[1], living in the household. The survey project was designed on three levels: the individual level, parent / guardian and household level and neighborhood level.

This hierarchical data structure allows for a multilevel statistical approach, which takes into account the non-independence of individuals within the same cluster and permits to distinguish the effect of neighborhood and household variables from that of micro level ones. For this purpose, the paper employs multilevel logistic regression, with an outcome variable referring to whether or not the child had a diarrheal episode during the previous 15 days. As covariates, indicators from different spheres (demographic, socio-economic, behavioral and environmental) and belonging to several levels of statistical analysis (child, household and neighborhood) are considered.

However, including a large set of variables pertaining to different levels may lead to statistical models which are difficult to interpret or even to estimate. Moreover, certain factors are multidimensional notions, which cannot be directly measured. For example, the surrounding environment, socio-economic status or health-related behavior are complex concepts and no single indicator can express them correctly and reflect the risks they may pose to health. They can however be approximated, by clustering several relevant manifest indicators. To achieve this, we employ latent variables in the form of latent class analysis (LCA).

The principle at the basis of latent variables is local / conditional independence, which postulates that the observed association between manifest variables is not due to hidden causality, but to the fact that these variables are, actually, indicators indicators of the same underlying concept: the latent variable (McCutcheon, 1987). This implies that, once the controlled for this latent variable, the association between the manifest variables disappears (statistical independence which is conditional upon the latent variable). In the case of categorical data, the result are latent classes.

One of the main advantages of using latent class analysis is that this procedure yields parsimonious models: it synthesizes a large quantity of data (manifest variables) into a pre-defined number of categories (called "classes") and subsequently assigns individuals to these classes, based on their respective conditional probabilities.

For the purpose of the present analysis, four latent variables were estimated in R, using the poLCA package (Linzer and Lewis, 2011). They refer to health-related behaviors in the household, to the household wealth, to the domestic sanitation and to neighborhood sanitary

1 The lowest age was limited to two years old, in order to avoid the potential bias of breastfeeding

infrastructures. These latent variables were subsequently used as explanatory part in a three level (child, household and neighborhood) logistic regression, to analyze their impact on the dependent variable of interest: the occurrence of a diarrheal episode in children.

The final multilevel models were elaborated using MLwiN Version 2.1 (Rasbash et al., 2009). They include demographic characteristics of the child and of the mother, as well as the four multi-dimensional concepts, approximated by means of LCA. Additionally, a synthetic variable was included in the model, which proxies the perceived insalubrity of the neighborhood, aggregated at neighborhood level. It represents the proportion of households that consider garbage to constitute a problem of the neighborhood.

Results

Three types of results are of interest. Firstly, the latent class analysis results, which allow to characterize the study population. Secondly, certain descriptive results and thirdly, the regression fixed coefficients and random effects.

Characteristics of the Survey Population (LCA Results)

LCA yields two types of results. First type of results are the conditional probabilities, which link each manifest indicator to its latent variable. Analyzing the conditional probabilities of the four latent variables estimated allows to characterize each of the classes estimated, with regard to the risk for diarrheal diseases (see Annex).

At the household level, mothers from households assigned to class A are less likely to give their children unknown and unprescribed medicine or to buy drugs outside of established locations, compared to mothers from households assigned to class B. Therefore, class A would refer to households which tend to have adequate medical behaviors, compared to class B, which is comprised of households which tend to have erroneous medical behaviors. In the same line of reasoning, households assigned to class C tend to be wealthier than those assigned to class D, whereas households in class D have the lowest level of wealth. Households assigned to class G tend to have improved domestic sanitation, as opposed to those assigned to class F. At the neighborhood level, those neighborhoods assigned to the Class I are characterized by proper infrastructure, compared to those assigned to Class H.

As the latent variables estimated are not connected to each other in any way, belonging to one class in one latent variable is not an indicator of class membership in another latent variable. Thus, for example, a household can belong to class A (adequate medical behaviors), but also to class F (improper domestic sanitation).

The second type of results, the mixing proportions, are the share of each of the latent classes in the population. This distribution of the classes for each latent variable creates a "portrait" of the units (households and neighborhoods) of the sample (Fig. 2 to 5).

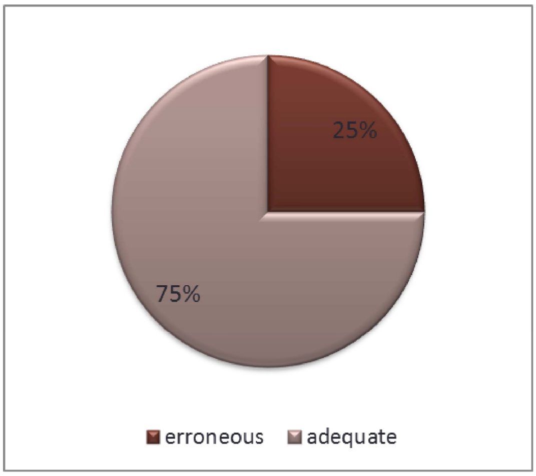

Figure 2: Distribution of households by Medical Behavior

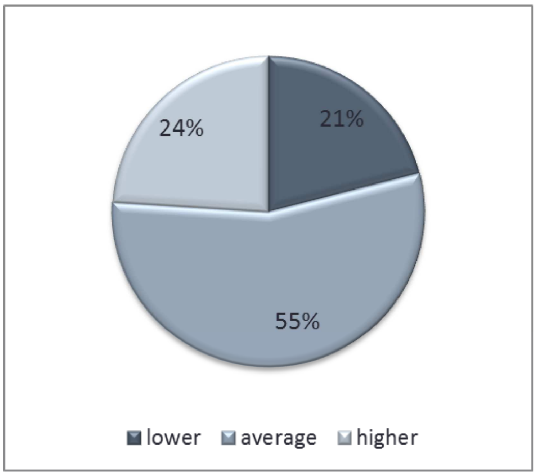

Figure 3: Distribution of households by Wealth

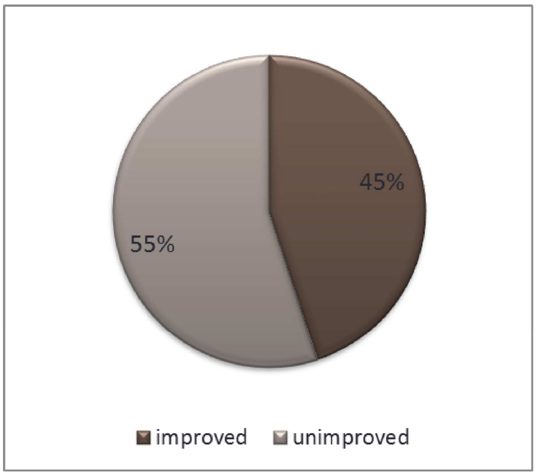

Figure 5: Distribution of households by Domestic Sanitation

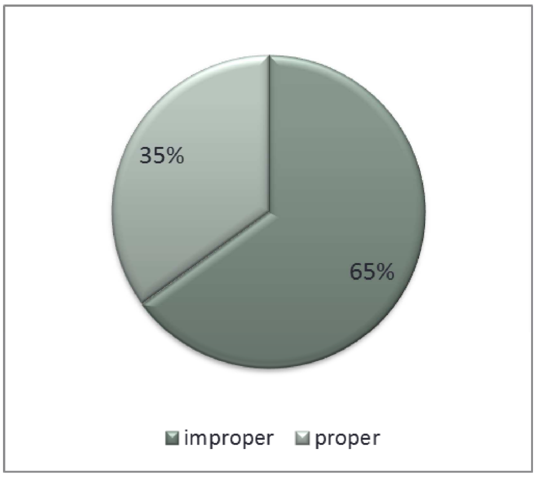

Figure 4: Distribution of neighborhoods by Sanitary Infrastructure

Thus, three quarters (75%) of households are characterized by adequate health behaviors and one-fourth (25%) by erroneous behavior. In terms of wealth, more than half (55%) of households belong to the average class, with 24% being richer and 21% being poorer than the average. Regarding domestic sanitation, the distribution is more or less equal, with 45% of households having access to improved domestic sanitation, while 55% of households do not. At the neighborhood level, a little over one third (35%) are distinguished by proper sanitation systems, contrary to 65%, with improper infrastructure.

Descriptive Results

The statistical analysis was performed on 7416 children aged 2 to 10, living in 2952 households, which are located in 50 neighborhoods of Dakar and its suburbs. Out of the 7416 children, 3665 were boys and 3751 were girls (Fig. 6).

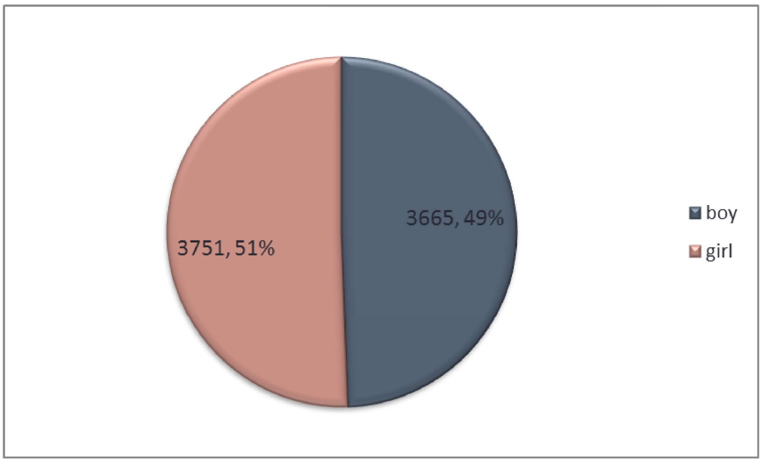

Figure 6: Gender distribution

Among the 1151 children who were declared to have experienced a diarrheal episode in the two weeks preceding the survey, the distribution by gender is relatively equal (Fig. 7).

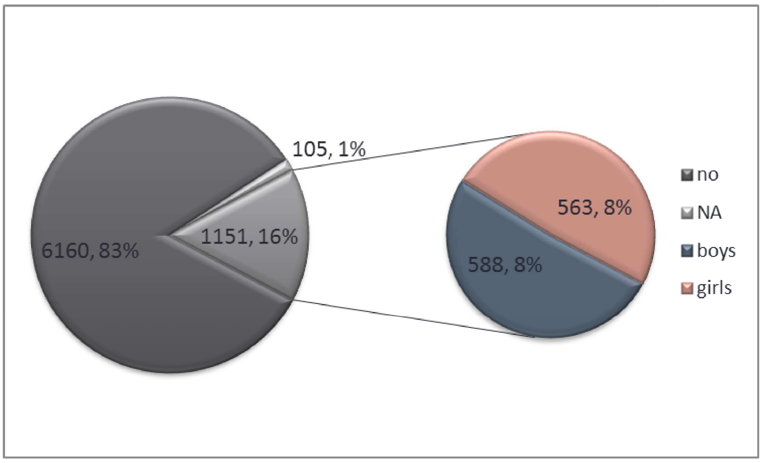

Figure 7: Incidence by gender

With, regard to age, 40% (2934) children were aged two to four, 33% (2454) were aged five to seven and 27% (2028) were aged eight to ten at the moment of the survey (Fig. 8).

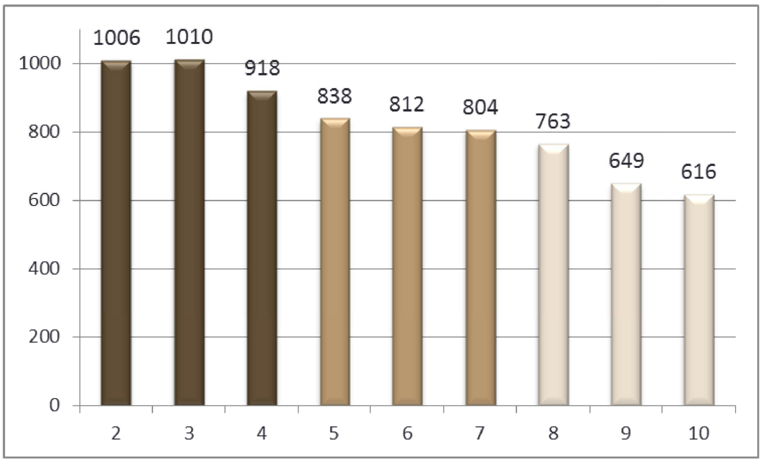

Figure 8: Age distribution

Although only 40% of the children in the sample were aged two to four, more than half (647) of the 1151 children who experienced a diarrheal episode were aged two to four. Comparatively, only 17% (195) were aged 8 to 10 (Fig. 9). This suggests that older children tend to be less vulnerable or less exposed to diarrheal diseases. However, this observation needs to be tested by means of a multivariate regression.

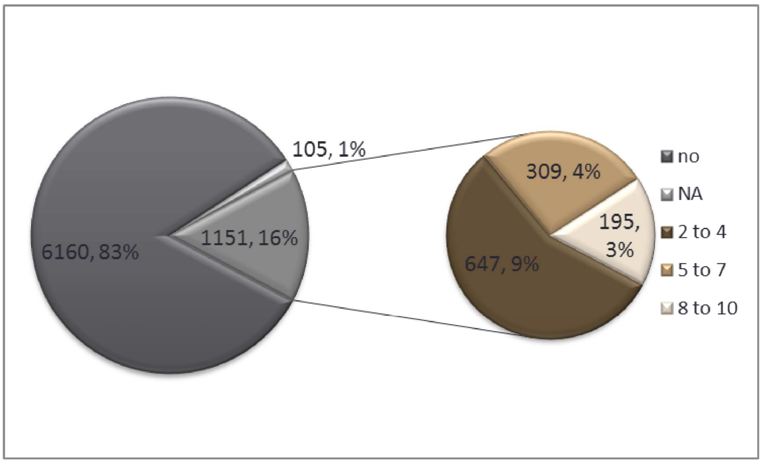

Figure 9: Incidence by age

Risk Factors for Diarrheal Diseases (Regression Results)

Table 1 shows the risk factors (odds ratios are presented) associated with an episode of diarrhea in children of 2 to 10 years old, which is used as the response variable in this work. The analysis consists of a three-level logistic regression. The explanatory variables used are continuous (age of the child and of the mother, as well as the proportion of households per neighborhood that judged their neighborhood to be insalubrious) and categorical (the child's gender, the mother's literacy and occupation, as well as the four latent variables previously estimated).

Model 1 does not contain any explanatory variable (i.e. is an empty model), and was used for calculating the intra-class correlation (see below).

Model 2 includes variables at the individual level: the age (included as a continuous variable[2]) and the gender of the child. Subsequently, in Model 3, we included variables pertaining to the mother/tutor of the child and to the household or dwelling. Here, aside from variables related to the mother's/tutor's age and socio-economic status, three of the estimated latent variables were added: medical behaviors, household wealth and domestic sanitation.

Finally, Model 4 includes all previous variables, to which we added two neighborhood-level variables: the latent variable related to sanitary infrastructures and the synthetic indicator regarding perception of insalubrity.

2 Models using age-squared and broad categories of age were also tested, but results did not differ significantly

		M1	M2	M3	M4
	Intercept	0.19***	0.46***	0.32***	0.20***
child	girl		0.97	0.97	0.97
	age		0.85***	0.84***	0.84***
mother	age			1	1
	illiterate			0.95	0.92
	-employed			1.09	1.1
	-other status			1.47**	1.50**
household	erroneous behavior			1.29***	1.28***
	-average wealth			1.18*	1.16
	-low wealth			1.43***	1.38**
	unimproved domestic sanitation			1.04***	1.05**
neighborhood	improper infrastructure				1.23**
	% discontent garbage				1.78**

Table 1: Estimated fixed effects of three-level logistic regression

CI: *** 99%; ** 95%

The fixed effects and their significance levels remain rather stable from one model to another. Results show that a wide array of determinants are involved in the outcome variable. As expected, age constitutes a protective factor for children, each additional year bringing a 15-16% reduction in odds for diarrheal diseases. Behavioral factors have a strongly significant influence on diarrheal diseases, independent of wealth or social status: living in a household where medical behaviors tend to be erroneous is associated with a 28% increase in the odds of diarrheal diseases. Similarly, living in a poorer household also tends to be associated with higher odds of diarrheal diseases. While the impact of domestic sanitation seems rather modest, neighborhood sanitation plays an important role, both in terms of sanitary infrastructure and of perceived quality of waste management.

Looking at the random effects however (particularly the intra-class correlation calculated on an empty model – Model 1), shows that, while only 1% was related to neighborhood factors, almost a quarter (23%) of the variance in the dependent variable was related to household factors. A possible reason for the low intra-class correlation at the neighborhood level could be related to the number of observations on the third level. With only 50 observations (neighborhoods) in the sample, this may have had an impact on the explanatory power of the variables at the neighborhood level.

Discussion

The regression results illustrate a wide spectrum of potential risk factors for childhood diarrheas, from individual characteristics (age), to family (wealth and behavior) and the immediate environment (domestic and neighborhood sanitation).

Household sanitation does not appear to play an important role, but neighborhood sanitation does. This is in line with a few other studies, which raised the issue that, in the absence of proper neighborhood sanitation, the impact of having good domestic sanitation may be cancelled: neighboring households, which do not have proper sanitation, will dispose of their wasterwater and garbage in the open, thus increasing the risk for their neighbors as well, irrespective of how good their domestic sanitation is (Heller, 1999).

Our results show that behavioral factors play a significant role, even when controlling for wealth and socio-economic factors. This suggests that relatively inexpensive, targeted interventions could potentially help reduce diarrheas in children, by improving hygiene and protective measures at the family level.

In our analysis, the mother's literacy does not appear to have any impact on the risk of diarrheal diseases. This could be due, in part, to the fact that more educated mothers tend to have occupations which demand their absence from the household. This means that the children are left with a carer (be it a member of the extended family or a person employed for this purpose) for most of the day.

A second possibility is that a certain amount of bias incurred from to the fact that diarrheal episodes were recorded on account of self-reporting. Although diarrhea is a symptom easily identifiable by the parents and the recall time does not exceed the generally accepted period, of 15 days, in order for results to be accurate (Fewtrell and Colford, 2004), it may be possible that this induced some degree of bias in the recording of the disease.

Conclusion

This paper aimed to analyze the relative role of the determinants of childhood diarrhea in a heterogeneous urban context, decanting the part relating to the individual, to the family/household and to the immediate environment.

To this end, we conducted a two-step statistical analysis: firstly four latent variables were created, using LCA; secondly, these latent variables were used as co-variates in a three-level logistic regression, in order to explain a recent diarrheal episode in children aged two to ten.

The results confirmed the hypothesis that diarrheal diseases in children are influenced by factors as belonging to the individual, the family, as well as to the immediate environment. This implies that measures aimed at reducing gastrointestinal infections need to include several components, in order to have a significant impact.

The results also show that LCA is a useful method for integrating a large number of demographic, socio-economic, behavioral and environmental factors, in order to create a comprehensive, but easily interpretable causal image.

References

Centers for Disease Control and Prevention (2012) Global WASH-Related Diseases and Contaminants. Available from: http://www.cdc.gov/healthywater/wash_diseases.html (accessed 26 August 2015).

Fewtrell L and Colford JM (2004) *Water, Sanitation and Hygiene: Interventions and Diarrhoea A Systematic Review and Meta-analysis*. Health, Nutrition and Population (HNP), Discussion paper, Washington D.C.: World Bank.

Heller L (1999) Who really benefits from environmental sanitation services in the cities? An intra-urban analysis in Betim, Brazil. *Environment and Urbanization* 11(1): 133–144.

Institute for Health Metrics and Evaluation (2013) The Global Burden of Disease. Available from: http://vizhub.healthdata.org/gbd-compare/ (accessed 9 September 2015).

Jalan J and Ravallion M (2003) Does piped water reduce diarrhea for children in rural India? *Journal of econometrics* 112(1): 153–173.

Lalou R (2008) Projet ACTU-PALU - Rapport scientifique 2008. Available from: http://en.calameo.com/read/0001329037be93e1ded73 (accessed 3 April 2014).

Linzer DA and Lewis JB (2011) poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software* 42(10): 1–29.

Masangwi SJ, Morse TD, Ferguson NS, et al. (2009) Behavioural and environmental determinants of childhood diarrhoea in Chikwawa, Malawi. *Desalination* 248(1-3): 684–691.

McCutcheon AL (1987) *Latent class analysis*. Newbury Park, Calif.: Sage Publications.

Myaux J, Ali M, Felsenstein A, et al. (1997) Spatial distribution of watery diarrhoea in children: identification of ‘risk areas’ in a rural community in Bangladesh. *Health & place* 3(3): 181–186.

Rasbash J, Charlton C, Browne WJ, et al. (2009) MLwiN Version 2.1. Centre for Multilevel Modelling, University of Bristol.

Regassa G, Birke W, Deboch B, et al. (2008) Environmental Determinants of Diarrhea among Under-Five Children in Nekemte Town, Western Ethiopia. Available from: http://www.ejhs.ju.edu.et/admin/Volume-18-Num2/Wonde_B.pdf (accessed 21 November 2012).

Rego RF, Moraes LRS and Dourado I (2005) Diarrhoea and garbage disposal in Salvador, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 99(1): 48–54.

Scholthof K-BG (2007) The disease triangle: pathogens, the environment and society. *Nature Reviews Microbiology* 5(2): 152–156.

Service Régional de la Statistique et de la Démographie de Dakar (2009) *Situation* *économique et sociale de la région de Dakar de l'année 2008*. Ministère de l'Economie et des Finances. Available from: http://www.ansd.sn/publications/annuelles/SES_Region/SES_Dakar_2008.pdf.

Waddington H, Snilstveit B, White H, et al. (2009) *Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries*. International Initiative for Impact Evaluation New Delhi. Available from: http://zunia.org/uploads/media/knowledge/3ie%20SR0011249554991.pdf (accessed 20 March 2013).

WHO (2009) Diarrhoeal disease. Available from: http://www.who.int/mediacentre/factsheets/fs330/en/index.html (accessed 11 January 2013).

Woldemicael G (2011) Diarrhoeal morbidity among young children in Eritrea: environmental and socioeconomic determinants. *Journal of Health, Population and Nutrition* 19(2): 83–90.

Annex

Conditional probabilities for the estimated latent variables

Medical behavior (N = 2952)		Class A		Class B
Gave the child unknown medicine, unprescribed by a health professional	no	0.77		0.49
	yes	0.23		0.51
Generally understands posology	with difficulty	0.20		0.25
	easily	0.80		0.75
Buys medicine from street hawkers	no	0.97		0.16
	yes	0.03		0.84
Buys medicine at the market	no	0.96		0.32
	yes	0.04		0.68
Household wealth (N = 2952)		Class C	Class D	Class E
Television	no	0.00	0.01	0.60
	yes	1.00	0.99	0.40
Video player (DVD or VCD)	no	0.14	0.52	1.00
	yes	0.86	0.48	0.00
Electric or gas stove	no	0.07	0.68	0.94
	yes	0.93	0.32	0.06
Living-room	no	0.69	0.97	1.00
	yes	0.31	0.03	0.00
Refrigerator	no	0.18	0.65	0.97
	yes	0.82	0.35	0.03
Ventilator	no	0.07	0.19	0.82
	yes	0.93	0.81	0.18
Car	no	0.69	0.96	0.99
	yes	0.31	0.04	0.01
Domestic sanitation (N = 2952)		Class F		Class G
Source of drinking water	improved	0.83		1
	unimproved	0.17		0
Source is private to the	yes	0		1

household	no	1	0
Type of toilet	improved	0.79	0.87
	unimproved	0.21	0.13
Toilet is private to the household	yes	0.24	0.95
	no	0.76	0.05
Wastewater management	sewer	0.26	0.37
	septic tank	0.11	0.14
	other	0.63	0.49
Garbage disposal	private bin with private collection	0.04	0.09
	other type of bin	0.84	0.82
	open disposal	0.13	0.09
Neighborhood sanitation (N = 50)		Class H	Class I
Water piping in the neighborhood	complete	0.91	0.94
	incomplete	0.09	0.06
Sewerage in the neighborhood	complete	0.00	0.85
	incomplete	1.00	0.15
Garbage collection in the neighborhood	complete	0.42	0.88
	incomplete	0.58	0.12
Water system perceived quality	good	0.00	0.25
	bad	1.00	0.75
Sewerage perceived quality	good	0.20	1.00
	bad	0.80	0.00
Garbage collection perceived quality	good	0.35	0.49
	bad	0.65	0.51