



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Juvakoski, Anni; Rantanen, Henrik; Mulas, Michela; Corona, Francesco; Vahala, Riku; Varis, Olli; Mellin, Ilkka

Evidence of waste management impacting severe diarrhea prevalence more than WASH : An exhaustive analysis with Brazilian municipal-level data

Published in: Water Research

DOI: 10.1016/j.watres.2023.120805

Published: 01/12/2023

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:

Juvakoski, A., Rantanen, H., Mulas, M., Corona, F., Vahala, R., Varis, O., & Mellin, I. (2023). Evidence of waste management impacting severe diarrhea prevalence more than WASH : An exhaustive analysis with Brazilian municipal-level data. *Water Research*, *247*, Article 120805. https://doi.org/10.1016/j.watres.2023.120805

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres



Anni Juvakoski^{a,*}, Henrik Rantanen^e, Michela Mulas^{a,b}, Francesco Corona^c, Riku Vahala^a, Olli Varis^a, Ilkka Mellin^d

^a Department of Built Environment, Aalto University, PO Box 15200, FI-00076, Aalto, Finland

^b Department of Teleinformatics Engineering, Federal University of Ceará, Campus do Pici, Fortaleza, Brazil

^c Department of Chemical and Metallurgical Engineering, Aalto University, PO Box 16100, Espoo, FI-00076, Aalto, Finland

^d Department of Mathematics and Systems Analysis, Aalto University, PO Box 11100, FI-00076, Aalto, Finland

^e Independent

ARTICLE INFO

Keywords: Waste management Housing Public health Low- and middle-income countries (LMIC) Water, sanitation, and hygiene (WASH) Municipalities

ABSTRACT

Adequate housing protects from diarrhea, which is a substantial health concern in low- and middle-income countries. The purpose of this study was to quantify the relationship between severe diarrhea and housing features at the municipal level to help in public health planning.

Regression analyses were performed on annual (2000–2012) datasets on Brazilian municipalities (5570) in six household feature categories (e.g., waste management) and four severe diarrhea outcomes (e.g., diarrhea deaths of under-5 children). Household data were not available elsewhere of this magnitude and granularity, highlighting the scientific value-add of this study. Municipalities were clustered prior to regression analysis because of data heterogeneity. The compositional household feature data were also subjected to principal component analysis to diminish feature variable multicollinearity.

The highest explanatory power was found for diarrhea deaths of under-5 children ($R^2 = 10-22$ %), while those in the over-5 population were the least best explained ($R^2 = 0.3-7$ %). Household features predicted diarrhea outcomes more accurately in the "advanced" housing municipality cluster ($R^2 = 16-22$ %) than in the "midlevel" ($R^2 = 7-20$ %) and "basic" ($R^2 = 6-12$ %) ones (over-5 diarrhea deaths excluded). Under-5 children's diarrhea death prevalence was three times higher in the "basic" cluster than in the "advanced" cluster. Importantly, the impact of waste management was overall the largest of all household features, even larger than those of WASH, i.e., water supply, sanitation, and household drinking water treatment. This is surprising in the context of existing literature because WASH is generally regarded as the most important household factor affecting gastrointestinal health.

In conclusion, public health interventions could benefit from customizing interventions for diarrhea outcomes, municipality types, and household features. Waste management's identified stronger association with diarrhea compared to WASH may have important implications beyond the water field and Brazil.

1. Introduction

Although diarrhea-related hospitalizations and diarrhea mortalities have been decreasing significantly for the last few decades around the world, they remain a substantial health concern in low- and middle-income countries (UNICEF, 2022). More specifically, global diarrhea mortality of under-5 children has dropped by 70 % between 1990 and 2017. This decrease is largely due to improved healthcare (see Section 3.7) as well as improvements in housing, particularly access to water,

sanitation, and hygiene (WASH) (Troeger et al., 2020). As countries and municipalities become more prosperous, they can apply large-scale policies to improve insufficient housing. Household features that effectively reduce contact with fecal pathogens potentially causing diarrhea are achievable for more and more households in the world. These improvements include appropriate sanitation, solid waste management, and a safe drinking water source. In low-income areas, inadequate housing conditions remain a major cause of diarrhea (e.g., Adane et al., 2017; WHO, 2017).

* Corresponding author.

https://doi.org/10.1016/j.watres.2023.120805

Received 13 July 2023; Received in revised form 26 October 2023; Accepted 27 October 2023 Available online 28 October 2023

0043-1354/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).





E-mail address: anni.juvakoski@aalto.fi (A. Juvakoski).

In Brazil and most other countries, municipalities are the main decision-making units in control of providing key public health services (Machoski and de Araujo, 2020; Neves, 2012; Lima et al., 2020). Brazilian municipalities are responsible for managing the water supply, sewage disposal, domestic solid waste, urban cleaning, and drainage (Neves, 2012). As municipalities are the main duty-bearers of these activities, the effects of their housing policies on health should thus be assessed on a relevant scale. Municipality-level analyses of diseases and their determinants have been found to be relevant in the epidemiology of several other health topics, such as the health program impact on infant mortality (Aquino et al., 2009), obesity (Reeve et al., 2015), and cancer mortality (Roquette et al., 2018).

As the majority of such decision-making is done by the municipalities, it is instrumental in understanding how successful municipalities are in fostering public health, especially in low- and middle-income countries. However, the existing literature on diarrhea determinants consists mostly of experimental studies looking at intervention efficacy in small communities (e.g., Degebasa et al., 2018), reviews of those experimental studies (e.g., Darvesh et al., 2017), and studies on individuals/families (e.g., Soboksa, 2021). Controlled experimental studies on intervention efficacy on the community, household, or individual scales can prove causality between diarrhea and its determinants, yet they provide little evidence on whether such interventions are effective when implemented in a different context or on a larger scale. Findings from small-scale case studies may not be generalizable to the municipal level. Additionally, experimental studies conducted in communities usually have short follow-up periods, thus providing only snapshots of determinant effects (Darvesh et al., 2017). Therefore, these research approaches have limited utility in designing effective public health policies. Nevertheless, the need to widen the research scope from individuals and households to neighborhoods and municipalities has been recognized in recent studies. For example, in their systematic review, Jung et al. (2017) found that poor neighborhood sanitation (including basic public sewerage infrastructure, open drainage, and open defecation) poses a risk of diarrhea almost equal to poor household sanitation. Similarly, Fuller and Eisenberg (2016) found that herd protection from WASH interventions within communities is a key factor in defending against diarrhea.

Not only has the studied scale of the housing-diarrhea relationship been somewhat limited in the existing literature, but studies have also mostly concentrated solely on WASH. There is a water sector field consensus that the betterment of drinking water and sanitation services are the main ways to combat diarrhea (e.g., WHO, 2017). For that reason, the effects of solid waste management practices, housing type and materials, and electrification are often left out of studies. However, some papers have demonstrated that these factors, not only WASH, effectively reduce contact with diarrhea-causing pathogens (e.g., Bühler et al., 2014; Randremanana et al., 2016; Yaya et al., 2018). Ahmed et al. (2020) recently found an association between increased diarrhea hospitalizations and improper waste management. See more discussion about literature on different household factors' connection with diarrhea in Section 3. Furthermore, many case studies find that household WASH interventions in fact produce varying outcomes on inhabitants' health (e.g., Sahiledengle et al., 2021; Carlton et al., 2014). Similarly, a recent review of sanitation intervention effectiveness found that, overall, sanitation interventions rarely affected diarrhea prevalence (Contreras and Eisenberg, 2020). If these issues exist on a larger scale, municipalities could save millions and increasingly boost their inhabitants' health through re-prioritizing efficient interventions in their policies.

Brazil was chosen as the study area for the following three reasons. First, Brazil is in the top 20 of the most economically unequal countries in the world, based on the GINI index (World Bank, 2020). Brazil is home to some of the most privileged as well as underprivileged communities in the world, therefore providing a diverse study sample. Its five geographic regions divided into 27 states (including the federal district) are distinguished by varying wealth and education levels (Szwarcwald et al., 2016). Second, Brazil has been a meticulous collector of data (see Section 2.1). As a result, there are vast household data available for this study. Third, Brazil is a huge upper-middle-income country with its 211 million population and 5570 municipalities (in 2019). Middle-income countries are an important research focus point, as more and more low-income countries are becoming part of this group, especially in the coming years. With this study, the authors aim to provide a highly relevant point-of-view for combating diarrhea in municipal and other large-scale public health planning-a clear knowledge gap in today's literature. To the best of the authors' knowledge, this granularity of household features and health data are not available elsewhere on this magnitude, and a study on this scale has not been conducted before. This municipal-scale analysis matches the intervention scale, thus providing a comprehensive outlook on the determinants of severe diarrhea in low- and middle-income countries. The goal of the study is to help equip decision-makers in the field with practical takeaways for tackling municipal-level health challenges.

The research question the study aims to answer is: How do housing features relate to severe diarrhea prevalence in different types of municipalities? In practice, the dependence of diarrhea hospitalizations and deaths on household features in Brazilian municipalities was examined using linear regression. The set of explanatory variables included the prevalence of types of water supply, sanitation, waste management, household drinking water treatment, house walls, and electrification. The dependent variables were prevalence of diarrhea hospitalizations and diarrhea deaths of under-5 children and the rest of the population. Before the regression analysis, observations (2000-2012) from municipalities (5570) were grouped into clusters by using the k-means method. This was done based on municipalities' overall housing level to explore whether it influences the way housing affects diarrhea prevalence. Clustering was followed by principal component analysis of household features to remove collinearity and reduce the dimensionality of the problem by decreasing the number of explanatory variables used in the regression models. See the Materials and Methods section for further reasoning for the methodology.

2. Materials and methods

Fig. 1 presents the overall workflow of the used methods. All code was written and executed in R and is available upon request. Household feature data (years 2000–2012) were available in the form of the number of families in each municipality utilizing a certain household feature each year. The annual diarrhea data from the same years indicate the number of diarrhea hospitalizations and diarrhea deaths in each municipality. Both data types were acquired from all 5570 municipalities in Brazil. To the best of the authors' knowledge, this granularity of household features and health data is not available elsewhere to this magnitude. The 2019–20 OECD Survey of Health Data Development (OECD iLibrary 2021) notes that in terms of development and the use of data within key national health datasets, Brazil "compares favorably" to other countries.

2.1. Household feature data

The data were acquired from the Brazilian Primary Care Information System (*Sistema de Informação da Atenção Básica*, SIAB), (SIAB, 2023), which collected the data through household visits. Data were collected from 15.6 to 35.4 million households annually, depending on the year. The number of counted families increased over the years, but the slightly varying sample sizes did not critically affect the proportions of families using each feature. The monthly collected data were later consolidated by the IT department of the Brazilian Unified Health System (*Departamento de Informática do Sistema Único de Saúde*, DATASUS) working under the Brazilian Ministry of Health with the following alterations. Municipalities not supplying data monthly were excluded from the data.



Fig. 1. Workflow of used methods. SIAB = the Brazilian Primary Care Information System (Sistema de Informação da Atenção Básica). DATASUS = IT department of the Brazilian Unified Health System (Departamento de Informática do Sistema Único de Saúde).

In the case of families fitting into several categories within one feature category, the more commonly used one was recorded in the data. In most cases, not all households in each municipality were visited.

To be able to compare municipalities, the number of households in each feature category was divided by the number of total sampled households to get the feature prevalence (%), i.e., count data were transformed into proportions. Outlier detection and removal are described in the supplementary material. Household feature categories are described according to SIAB data technical notes in Table 1. To help discuss the overall change in household features in Brazil over the years, prevalence maps (1998–2015) were created in addition to the actual statistical analysis.

2.2. Diarrhea hospitalization and diarrhea death data

Diarrhea hospitalizations and diarrhea deaths of under-5 children and the over-5 population were acquired from the database of DATA-SUS. The data indicate the number of diarrhea hospitalizations and diarrhea deaths in each municipality in each year (2000–2012). According to the data source, it includes all diarrhea hospitalizations from all public and most private health facilities as well as diarrhea deaths from all possible registering facilities.

Upon collecting data from the database, the diarrhea data was aggregated into two categories: children four years and younger (=under-5 children) and people older than that (=over-5 population). All diarrhea data were listed in the form of the patient's home municipality, even if the patient was admitted to a health facility somewhere

Table 1

Household feature categories. Household feature data (2000–2012) and their descriptions were acquired from the Brazilian Primary Care Information System (SIAB). Note that the household (point-of-use) water treatment category does not consider possible water treatment carried out by a water supplier. The household features presented in this table were acquired as count data (number of families using a feature), which were transformed into proportions (%).

Household feature category	Household feature				
1. Water supply	1.1 Public supply at premises or proximity				
	1.2 Well or spring				
	1.3 Other: rainwater, water truck or public				
	fountain, well, or tap				
2. Sanitation	2.1 Public sewage system				
	2.2 Any type of septic pit				
	2.3 Open defecation				
Waste management	3.1 Collected by a public or private company				
	3.2 Burned/buried				
	3.3 Dumped/not disposed				
House wall materials	4.1 Brick/adobe				
	4.2 Rammed earth coated with plaster or similar				
	4.3 Uncoated rammed earth				
	4.4 Wood				
	4.5 Reused/inappropriate materials, including				
	cardboard, plastic, canvas, and straw				
	4.6 Other materials such as stone or concrete				
5. Household water treatment	5.1 Filtering				
(point-of-use treatment)	5.2 Boiling				
	5.3 Chlorinating				
	5.4 No household water treatment				
6. Electrification	6.1 Electrification (including non-continuous				
	supply or non-regularized installation)				

else. Hospitalizations (DATASUS, 2023a) listed under all available fitting categories describing diarrhea were included to compile the diarrhea hospitalization data. These categories included "cholera," "typhoid and paratyphoid fever," "shigellosis," "amebiasis," "diarrhea and gastroenteritis of infectious origin," and "other infectious gastro-intestinal diseases." Diarrhea mortality was also compiled from the same source—the mortality database in DATASUS (DATASUS, 2023b). Diarrhea deaths were likewise assembled from all the available categories describing diarrhea mortality and include "intestinal infectious diseases," "cholera," "diarrhea and gastroenteritis of infectious origin," and "other infectious gastro-intestinal diseases."

To obtain proportional data, the count data were transformed into occurrences per 100,000 inhabitants by dividing the diarrhea hospitalizations and deaths by the population of the municipality and multiplying that by 100,000. It was assumed that the diarrhea data contained no errors, as diarrhea prevalence can change drastically from one year to the next due to sporadic epidemics. To help analyze the overall change in diarrhea hospitalizations and diarrhea deaths in Brazil over the years, prevalence maps (for the years 2000, 2006, and 2012) were created in addition to the statistical analysis (Fig. 2 and supplementary material).

2.3. Clustering municipalities into groups

The authors suspected that relationships between severe diarrhea and household features may be inappropriately masked by the heterogeneity of the vastly different types of Brazilian municipalities. Analysis of aggregated countrywide data therefore would not help in designing targeted public health policies. Therefore, all observations in the data (from any year and municipality) were clustered using the k-means



Fig. 2. Change in example variables during the study period. Top row: diarrhea hospitalization prevalence of under-5 children in 2000, 2006, and 2012. Bottom row: trash dumping (i.e., not disposing of trash) prevalence in 2000, 2006, and 2012. Hosp. = hospitalizations.

algorithm (MacQueen, 1967). All analyses were done with proportional data (occurrences in each municipality per municipal population, i.e., all values were between 0 and 1), which is why no additional normalization or standardization was necessary. In *k*-means clustering, similar observations are assigned to groups based on which group has the nearest mean. All household features were used as the clustering basis data. If observations from a single municipality appeared in several different clusters, the municipality was assigned to the cluster where it appeared the most often. Geographical information was not used in clustering. The optimal number of clusters was determined using the elbow (Thorndike, 1953) and silhouette (Rousseeuw, 1987) methods.

2.4. Principal component analysis of household features and regression within clusters

Each year, each family can belong to only one of the features within each feature category in the data. For example, a family can utilize (1) the public sewage system, (2) any type of septic pit, or (3) open defecation within the household sanitation category. This means the data are compositional, i.e., the sums of families within each feature type add up to 100 % within each feature category. The features within a category are thus highly negatively correlated with each other. Due to extensive collinearity, a principal component analysis (PCA) (Pearson, 1901) was used for the compositional data within each variable category to obtain uncorrelated covariates. PCA works by linearly transforming the data into a new coordinate system where the variation in the data is projected with fewer dimensions than the original data. This way, maximum variability in the data was also retained while reducing dimensionality. PCA was performed for each household feature category within each cluster. The composition of principal components is presented in the supplementary material, with the exception of waste management principal components, which are presented in the Results and Discussion section.

The principal components that explained more than 10 % of the variance within a feature category were used as independent variables in linear multivariate regression analysis performed within clusters (Table 3). Backward selection was used to create optimized models with the principal component variables. The model with the lowest Akaike Information Criterion (AIC) value was selected in each case. One of the four scaled diarrhea variables (hospitalizations and deaths of the over-5 population and under-5 children) was used as the dependent variable in each model. All observations were weighted by the population of municipalities to give more emphasis to larger municipalities, where diarrhea prevalence data is assumed to be less noisy due to larger sample sizes. As the electrification feature did not fit into any of the feature categories, it was used in the regression analysis as an independent variable without applying PCA. Pairwise correlations between the principal components were below 60 %, with two exceptions where the correlations were 67 % and 78 %.

2.5. Ethics statement

The used health and household data are anonymous and openly available in public databases. Therefore, no ethical concerns are declared.

3. Results and discussion

This opening overview on diarrhea and household features is kept short because diarrhea prevalence in Brazil has been covered in other works (e.g., Patrícia et al., 2013). All figures referred to in this section are shown in the supplementary material, except for two example figures of under-5 children's diarrhea hospitalizations and trash dumping (Fig. 2). In any given year, correlations between any of the diarrhea and any of the household features in municipalities were weak (pairwise Pearson correlation coefficients <0.2), although their explanatory power in a multivariate setting was good. A similar trend between diarrhea variables and household features can nevertheless be observed from maps depicting their geographical distribution in Brazil during the years 1998–2015 (see the supplementary material).

Between 2000 and 2019, the total global annual number of diarrhea deaths of under-5 children decreased by 61 % (UNICEF, 2022). A similar decreasing trend can be observed in Brazil's under-5 children's diarrhea hospitalizations (Fig. 2, top row). A high prevalence of diarrhea hospitalization and death has occurred mostly in the North and Northeast regions, which are the poorest in Brazil (Szwarcwald et al., 2016). With open defecation, a drastic change has occurred over the years in the North and Northeast regions. Open defecation has decreased a lot since the 1990s in the Northeast region. In the data, housing made with recycled/inappropriate materials has almost disappeared from Brazil, apart from a few clustered municipalities in the Northeast region (see the supplementary material). Trash dumping has also decreased strongly over time (Fig. 2, bottom row). The practice was mostly restricted to the Northeast region and a few municipalities in the westernmost corner of the North region in the 2000s. Also, WASH features follow similar trends as the other household features (see the supplementary material).

3.1. Clustering results

The data were clustered to unmask spatial relationships between severe diarrhea and household features. Based on the silhouette and elbow methods, the optimal number of clusters was three. The geographical distribution of each cluster is presented in Fig. 3. The clusters were given labels ("advanced," "mid-level," and "basic") based on the level of housing in the three clusters (Fig. 4). By and large, most municipalities belonging to the "advanced" cluster are located in the Southeast region (Fig. 3). "Mid-level" status municipalities are found in the South and Northwest regions, and "basic" municipalities are located mostly in the Northeast region. The Central-West region shows up as a mix of all three cluster types. These spatial patterns are clear even if geographic coordinates were not used in clustering. This is an expected cluster division, as the order of prosperity from richest to poorest region is Southeast, South, Central-West, North, and Northeast (Szwarcwald et al., 2016). According to Szwarcwald et al. (2016), the first three are thought to roughly resemble high-income nations, while the last two resemble low-income nations. These trends can also be seen in maps in the supplementary material depicting illiteracy and extreme poverty levels in Brazil (created from 2010 census data).

As can be seen from the boxplots of Fig. 4, the three clusters are different from each other. These differences can be seen most clearly when looking at the three top rows (=prevalence of 1. water supply, 2. sanitation, and 3. waste management), where the level of advancement of household features decreases from left to right. The boxes in the figure represent the most common cases in each cluster (lines inside boxes = medians). The "advanced" cluster includes the municipalities with the highest prevalence of the most progressive options of every household feature type. That is, most households in their municipalities use the public water supply at or near the premises, the public sewage system, waste collection by public or private companies, electrification, and brick/adobe walls. The "mid-level" cluster falls in the middle in terms of the level of its household features; it has the highest prevalence of midlevel household features (Fig. 4, second column in the top three rows). The "basic" cluster has the lowest prevalence of the most advanced household features and the highest prevalence of the most basic ones (Fig. 4, top three rows). These patterns mirror the overall levels of prosperity, education, and well-being within the clusters, as indicated in Fig. 5 (in terms of diarrhea) and maps presenting the prevalence of illiteracy and extreme poverty (in the supplementary material).

The "advanced" and "mid-level" clusters mostly have similar medians within the feature categories of water supply, sanitation, and waste management (top three rows) except in the right-most column, where the "advanced" cluster is clearly different from the other two



Fig. 3. Geographical distribution of municipalities belonging to different housing clusters. Purple = "advanced" housing cluster (1), green = "mid-level" housing cluster (2), yellow = "basic" housing cluster (3). The thick gray lines on the map represent region (macro regions) borders, and the thinner lines represent municipality borders. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Fig. 4. Distributions of household features in each of the three housing level clusters. Purple = "advanced" housing cluster (1), green = "mid-level" housing cluster (2), yellow = "basic" housing cluster (3). The values on the *x*-axis represent the distribution (%) of families using each feature in the municipalities belonging to each cluster. The lines inside the boxes represent the median, while the left and right edges of the boxes represent the first and third quartiles, respectively. The endpoints of the "whiskers" (i.e., the lines outside the boxes) represent the maximums and minimums. The black points are observations that do not fit between the first and third quartiles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

clusters. The features in the four bottom rows do not have a clear order of progressiveness. However, as filtering is the most common type of household drinking water treatment in the "advanced" cluster, it could be interpreted as being the most progressive option in that category. Fig. 4 also shows that some features are very rare. According to the data, wall types other than brick and wood are infrequent everywhere. Wooden houses are common only in the mid-level cluster. Boiling drinking water at home is also almost non-existent throughout. On average, on the municipality level, the population in the "advanced" cluster is three times the size of the population in the other two clusters (see the supplementary material).

Diarrhea prevalence in the three clusters follows a trend similar to the household features: diarrhea hospitalization prevalence of the over-5 population and under-5 children in the "basic" cluster are approximately two times more than in the advanced cluster, while the "midlevel" cluster falls somewhere in between (Fig. 5). The most drastic



Fig. 5. Diarrhea prevalence in each cluster. Purple = "advanced" housing cluster (1), green = "mid-level" housing cluster (2), yellow = "basic" housing cluster (3). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difference is observed with children's diarrhea death prevalence, which is three times higher in the "basic" cluster than in the "advanced" cluster. It is also notable that there is only a small difference in the prevalence of diarrhea hospitalizations and deaths between the very small age group of under-5 children and the rest of the population. In other words, diarrhea is a much more severe health concern among young children.

3.2. Principal component regression results

As this is an observational study on the municipal scale, the modest explanatory power of the models is expected (Table 2) because diarrhea is a complex phenomenon with many other explanations besides suboptimal housing (see Section 3.7). As can be seen in Table 2, the relationships between diarrhea and household features vary significantly within the three clusters and among the different severe diarrhea outcomes. However, the regression models cannot be directly compared to each other because each model uses data from different municipalities due to clustering. It is well-established that all included household features correlate strongly with diarrhea (see Sections 3.2-3.5). The key aim was to present the relative strengths of relationships in the different clusters to be able to recommend targeted prioritization of features for different kinds of municipalities. Although using PCA obscures the interpretation of the effect of each feature (e.g., sewerage, septic tanks, and open defecation), the effects of feature categories (e.g., sanitation and waste management) can be compared to each other. In Table 2, ranks have been given to the significant (p-value ≤ 5 %) independent variable (household feature) categories based on the magnitude of their total t-values within each category. This way, comparing the impact of different feature categories can be done. The loadings of the principal component variables are not discussed in detail, apart from those of the waste management feature category (Table 3). Details on the other principal components can be found in the supplementary material.

In terms of coefficient of determination (R^2) values, household features are overall the best predictors of severe diarrhea in the "advanced" cluster, while they have the smallest predictive power in the "basic" cluster. The "mid-level" cluster falls in between. The quality of housing in a cluster reflects the overall wealth and well-being of its municipalities (see illiteracy and extreme poverty maps in the supplementary material). Therefore, the comparably good explanatory power in the "advanced" cluster could indicate that housing is increasingly a better indicator of severe diarrhea when overall well-being in an area is high. Note that at the same time, the more advanced the household features in a cluster, the lower the diarrhea prevalence (see Fig. 5). Conversely, in the opposite case, housing factors may explain less of the diarrhea variance in less well-off areas due to lower levels of education, income, and healthcare being stronger predictors in such municipalities (see Section 3.7).

Overall, the severe diarrhea outcome best predicted (highest R^2) with household features is diarrhea deaths of under-5 children. The models have the least predictive power for diarrhea deaths in the over-5 population. Young children are more susceptible to severe diarrhea. They also have poorer hygiene behavior than the over-5 population (Strina et al., 2003), making them more vulnerable to unsanitary household conditions. Over-5 population's diarrhea deaths seem to be a more complex phenomenon, explained almost entirely by factors outside the household (significant coefficients, but low R^2). This indicates that improvements in the household might not be effective in combating diarrhea deaths in the over-5 population. The low explanatory power may be due to the cause of death being obscured by comorbidity or old age. The over-5 population's diarrhea hospitalizations are also explained well by housing factors, as are children's diarrhea hospitalizations in the advanced cluster. This may point to people being more willing and/or able to hospitalize their children due to diarrhea in more well-off areas than in poorer areas. The findings of Coube et al. (2023) are in line with this: wealthier Brazilians are more likely to make use of hospital services.

3.3. Significance of sanitation, point-of-use water treatment, and water supply (WASH)

The field consensus is that WASH is the main way to prevent diarrhea and related deaths (e.g., Fewtrell et al., 2005; WHO, 2017; UNICEF, 2022). However, surprisingly, the types of sanitation, water supply, and household drinking water treatment (components of WASH) did not seem to be the most important diarrhea determinants in the regression analyses of this study (Table 2). Nevertheless, household drinking water treatment did rank first or second in most models. The importance of sanitation varied from being the most impactful determinant in modeling diarrhea deaths of under-5 children in the "basic" cluster to being ranked last in some other models. Interestingly, the water supply determinants were the least important overall. This finding could

A. Juvakoski et al.

Table 2

Principal component regression results. The "rank" column shows how impactful each household feature is in affecting diarrhea prevalence in comparison to the other features (1 = most impactful – bolded text; 6 = least impactful; – = not significant). The rank is determined by the combined magnitude of the *t*-values of principal components of each household feature category (e.g., PC1_sanitation and PC2_sanitation together determine the total impact of sanitation). sgn = significance (*p*-value) in stars (* ≤ 5 %, ** ≤ 1 %, *** ≤ 0.1 %). If all columns have "–", the variable was not included in the model based on the AIC. If only the "sgn" column is marked "–", the variable was selected for the model, but it is not significant. Household drinking water treatment is presented as "pou." If a *t*-value has a - sign, it means the independent variable is negatively correlated with the dependent one and vice versa.

Variable		``Advanced'' cluster				``Mid-level'' cluster				``Basic'' cluster			
Dependent	Independent	t-value	<i>p</i> -value	sgn	rank	t-value	<i>p</i> -value	sgn	rank	t-value	<i>p</i> -value	sgn	rank
Under-5 diarrhea deaths	pc1_waste	-35.39	0.00	***	1	-5.05	0.00	***	2	8.86	0.00	***	2
	pc2_waste	12.05	0.00	***	1	-13.06	0.00	***	2	19.86	0.00	***	2
	pc1 sanitation	2.39	0.02	*	4	-9.48	0.00	***	5	16.40	0.00	***	1
	pc2 sanitation	9.15	0.00	***	4	-2.73	0.01	**	5	-21.74	0.00	***	1
	pc1 water	-2.91	0.00	**	6	_	_	_	_	9.33	0.00	***	6
	pc2 water	_	_	_	_	3.24	0.00	**	6	-1.55	0.12	_	6
	pc1 pou	-8.83	0.00	***	2	-11.58	0.00	***	1	-9.65	0.00	***	5
	pc2 pou	-12.77	0.00	***	2	-21.10	0.00	***	1	_	_	_	_
	pc3 pou	7.55	0.00	***	2	11.44	0.00	***	1	-6.42	0.00	***	5
	pc1 walls	-3.10	0.00	**	3	-5.52	0.00	***	3	-12.38	0.00	***	3
	pc2 walls	-10.96	0.00	***	3	-10.49	0.00	***	3	9.48	0.00	***	3
	Electrification	_	_	_	_	_	_	_	_	_	_	_	_
	(Constant)	22.55	0.00	***	_	23.39	0.00	***	_	18.26	0.00	***	_
		Adi, R-sou	ared: 0.220	7		Adi, R-squared: 0.2007		Adi, R-sau	Adj. R-squared: 0.0969				
Under-5 diarrhea hospitalizations	pc1 waste	-21.17	0.00	***	1	2.48	0.01	*	1	10.54	0.00	***	1
*	pc2 waste	21.17	0.00	***	1	-12.53	0.00	***	1	14.31	0.00	***	1
	pc1 sanitation	_	_	_	_	-3.78	0.00	***	5	3.82	0.00	***	6
	pc2 sanitation	6.48	0.00	***	4	-2.17	0.03	*	5	-3.81	0.00	***	6
	pc1 water	-1.53	0.13	_	3	-1.84	0.07	_	6	3.84	0.00	***	5
	pc2 water	12.94	0.00	***	3	-1.50	0.13	_	6	-7.25	0.00	***	5
	pc1 pou	-6.66	0.00	***	2	-4.71	0.00	***	3	15.31	0.00	***	2
	pc2_pou	-8.37	0.00	***	2	1.46	0.14	_	3	2.84	0.00	**	2
	pc3_pou	_	_	_	_	1.63	0.10	_	3	-3.40	0.00	***	2
	pc1_walls	_	_	_	_	-7.75	0.00	***	2	-10.20	0.00	***	3
	pc2_walls	-6.38	0.00	***	5	-6.29	0.00	***	2	-4.20	0.00	***	3
	Electrification	-	-	_	_	-	-	_	_	-	-	_	_
	(Constant)	46.43	0.00	***	-	10.47	0.00	***	_	5.72	0.00	***	_
		Adj. R-squared: 0.1602 Adj. R-squared: 0.0701						Adj. R-squared: 0.0628					
Over-5 diarrhea deaths	pc1_waste	-16.71	0.00	***	2				9.29 0.00 *** 1				
	pc2_waste	7.39	0.00	***	2	_	-	_	_	6.15	0.00	***	1
	pc1_sanitation	-	_	_	-	_	-	_	-	4.33	0.00	***	6
	pc2_sanitation	9.09	0.00	***	3	-2.90	0.00	**	3	-	_	-	-
	pc1_water	3.05	0.00	**	5	2.64	0.01	**	1	10.51	0.00	***	3
	pc2_water	4.67	0.00	***	5	-1.43	0.15	-	1	-2.81	0.00	**	3
	pc1_pou	-15.16	0.00	***	1	3.42	0.00	***	2	8.54	0.00	***	2
	pc2_pou	-5.40	0.00	***	1	_	-	-	-	-4.50	0.00	***	2
	pc3_pou	12.49	0.00	***	1	_	-	-	-	2.38	0.02	*	2
	pc1_walls	3.36	0.00	***	4	2.31	0.02	*	4	-	-	-	-
	pc2_walls	-5.38	0.00	***	4	-	-	-	-	-8.10	0.00	***	5
	Electrification	_	-	-	-	-	-	-	-	-	-	-	-
	(Constant)	9.14	0.00	***	-	56.61	0.00	***	-	11.10	0.00	***	-
		Adj. R-squ	ared: 0.069	8		Adj. R-squared: 0.0027		Adj. R-squared: 0.0255					
Over-5 diarrhea hospitalizations	pc1_waste	-32.37	0.00	***	2	-21.07	0.00	***	1	23.39	0.00	***	1
	pc2_waste	5.39	0.00	***	2	-5.59	0.00	***	1	17.08	0.00	***	1
	pc1_sanitation	9.79	0.00	***	3	-14.85	0.00	***	4	11.41	0.00	***	2
	pc2_sanitation	4.85	0.00	***	3	-	-	-	-	-20.79	0.00	***	2
	pc1_water	-5.44	0.00	***	4	2.63	0.01	**	5	7.97	0.00	***	5
	pc2_water	5.72	0.00	***	4	4.72	0.00	***	5	-5.08	0.00	***	5
	pc1_pou	-15.06	0.00	***	1	-	-	-	-	-15.87	0.00	***	3
	pc2_pou	-9.54	0.00	***	1	-2.73	0.01	**	3	-9.29	0.00	***	3
	pc3_pou	16.29	0.00	***	1	12.99	0.00	***	3	-	-	-	-
	pc1_walls	-4.35	0.00	***	5	-4.83	0.00	***	2	-	-	-	-
	pc2_walls	-6.35	0.00	***	5	-14.21	0.00	***	2	-	-	-	-
	Electrification	-	-	-	-	-	-	-	-	-	-	-	-
	(Constant)	15.62	0.00	***	-	99.11	0.00	***	-	8.48	0.00	***	-
		Adj. R-squ	Adj. R-squared: 0.1740			Adj. R-squared: 0.1689				Adj. R-squared: 0.1196			

indicate that the importance of water supply may be overestimated in diarrhea studies that do not include several other household features in their models. Nevertheless, the well-known importance of WASH and related behavioral aspects are discussed further in this section.

One reason the water supply variables often ranked last or were nonsignificant in the models may be that many high- or middle-income Brazilians, especially in cities, drink bottled water (Statista, 2022). Hence, they might not consume water from the water source recorded for them in the used data. Conversely, the household drinking water treatment variables often ranked near the top. Their interpretation is, however, not completely clear, as the used data are not linked in a way that would reveal which drinking water source households are using. The benefits of filtering, chlorination, boiling, and not treating water at home vary of course, depending on whether the water source is a tap, a well, a bottle, or surface water. These are some of the main limitations of the data.

Hand washing and other hygiene practices have also been shown to significantly reduce diarrhea prevalence even in rudimentary household conditions (e.g., Melo et al., 2008; Cairncross et al., 2010; Alebel et al., 2018). Data on this could not be found, but Brazilians are known to be very attentive to personal hygiene (Corona and Mulas, 2022), which may be another contributing factor in keeping diarrhea levels low even in impoverished areas. In this study, for less well-off areas, sanitation ranked as the most important household feature in affecting under-5 diarrhea deaths in the "basic" cluster. Otherwise, its rank varied from second to last. Although this study and their study are not directly comparable, this finding agrees with that of Nandi et al. (2017). They found piped water and improved sanitation to have the strongest diarrhea-preventing effects in poorer households. Many studies have similarly found that WASH factors have varying effects based on other circumstances (Sahiledengle et al., 2021; Carlton et al., 2014; Yaya et al., 2018). For example, sex and age of children, education of parents, and household wealth may increase or decrease the effects of WASH features on diarrhea (e.g., Yaya et al., 2018).

3.4. Significance of electrification and house walls

In this study, electrification was the only household feature that was not selected for any of the models constructed based on the AIC. This might point to it being a less impactful determinant of severe diarrhea than the other household features. Nevertheless, some scholars have found a few ways in which electrification may affect diarrhea prevalence. For example, Crawford (2009) noticed that consuming meat improperly stored due to lack of electrification led to a high prevalence of diarrhea in Jamaica. Samad and Zhang (2018) noticed a similar trend in Pakistan but assessed that the connection is due to the lack of health knowledge that would be gained through electronic media. Yaya et al. (2018) found that Nigerian households without electricity had 27 % higher odds (highest odds in their multivariable study) of experiencing under-5 diarrhea, but the authors do not directly discuss why this is. Randremanana et al. (2016) found the same relationship but likewise do not discuss the mechanism behind the finding.

As with electrification, the effect of house wall materials (=house type) on diarrhea has not been widely studied. In the results of this study, its rank among the other household features varied from second to last, and sometimes it was non-significant. Yaya et al. (2018) found that lacking concrete roofing and walls increased the odds of children experiencing diarrhea by 14–16 %, compared to households with those features. Pattanayak and Wendland (2007) conversely found no connection between poor building quality and diarrhea. Some studies have looked at the floor or roof type instead of the wall type. For example, Ndikubwimana and Ngendahimana (2020) and Melese et al. (2019) found that having an unimproved floor (mud, earth, or otherwise rudimentary) was a significant diarrhea risk factor. Housing materials and house types as diarrhea determinants may require more research.

3.5. Significance of waste management

Waste management was notably the most impactful factor ranking first (in seven out of twelve models) or second (in four out of twelve models) (Table 2). Furthermore, the significance level of its principal components was ≤ 0.1 %, with only one exception. These results suggest waste management practices urgently need more attention as a preventative/predisposing feature across all severe diarrhea types and different types of municipalities. Also, some other authors have discovered a strong relationship between waste management and diarrhea. As in this study, some studies have even found the effect to be stronger than the WASH features. Randremanana et al. (2016) found the likelihood of severe diarrhea in children to be three times greater (OR = 3.2) in households with garbage on the premises than in garbage-free

households. Such a clear relationship with WASH factors was not found in their study. Recent research has moreover identified an association between living near open dumpsites and respiratory and intestinal infections (Mberu et al., 2022), as well as an association between dumping of waste into rivers and burying waste with waterborne diseases including diarrhea (Rahman et al., 2021).

A few studies have specifically explored the garbage and diarrhea issue in Brazil. Bühler et al. (2014) conducted a study on infant (age <1) diarrhea mortality and hospitalizations in Brazilian microregions in 2010. They looked at some of the most understood determinants of diarrhea (including WASH) and found only the level of garbage collection to be a significant determinant in both diarrhea deaths and hospitalizations. Moreover, Rego et al. (2005) conducted a cross-sectional study on under-2 children's diarrhea in an impoverished neighborhood close to a city garbage dump in Salvador, Brazil. They noticed that out of many variables (including WASH, mother's education, breastfeeding, and unemployed head of the family), exposure to garbage in the environment increased the risk for diarrhea the most (AOR: 3.98). Also, a few recent studies looking at the disease burden of Brazilian waste pickers noticed that along with other infectious diseases, they constantly suffer from diarrhea (Cruvinel et al., 2019, 2020).

Dumping solid waste containing pathogens on household premises introduces another source of fecal pathogens in the household through several different mechanisms. An increase in especially rotavirus diarrhea hospitalizations has recently been associated with improper waste management (Ahmed et al., 2020). When children spend time in the yard, the risk of them getting exposed to pathogens in the waste and developing diarrhea goes up. Additionally, flies or other vectors might transport fecal pathogens from waste to foods (Brown et al., 2013). The connection between dumping sites and bacterial contamination of groundwater is also well-known (e.g., Wakida and Lerner, 2005), and dumping waste on household premises could have a similar effect on soil, runoff, and groundwater on a smaller scale. Dumping solid waste on household premises might be especially risky if the waste contains feces. A few studies point to unsafe child stool disposal significantly increasing the risk of children getting diarrhea (Bawankule et al., 2017; Sinmegn Mihrete et al., 2014; Yaya et al., 2018; Soboksa, 2021). Nevertheless, no data on children's stool disposal were available for this study.

Another potentially diarrhea-causing waste type in households involved in animal husbandry is animal feces. Close contact with food animals is associated with increased diarrhea prevalence in the review article by Zambrano et al. (2014). The authors explain that living with livestock or poultry is especially risky for young children because they are more susceptible to the fecal-oral transmission that can occur when animal feces contaminate the soil. In another review article, Penakalapati et al. (2017) recommend that future research should concentrate on thoroughly assessing and mitigating the health risks involved with animal feces.

3.6. Zooming in on waste management

Our finding of the strong association between severe diarrhea and waste management is surprising, and it may have important policy implications. For those reasons, waste management principal components and their regression outcomes are briefly discussed in this section to increase understanding of what interventions within waste management might be prioritized. Literature on the topic is discussed in the previous section.

The absolute magnitudes of loadings (+/-) of each waste management feature within PC1s and PC2s follow similar patterns (Table 3). This suggests these PC1s and PC2s might roughly have the same meanings across clusters in this case. Based on the loadings of the waste

Table 3

Loadings of waste management principal components. Only the principal components that explain over 10 % of the variance within the feature category are displayed here. These are the waste management principal components that were used as independent variables in the regression analysis. Each principal component labeled "PC1" or "PC2" in this table equals "PC1_waste" and "PC2_waste" in Table 2, respectively. The numbering of waste management features refers to the numbering of Table 1.

	Feature loadings in "Advanced"	clusters	"Mid-level"		"Basic"		
	PC1	PC2	PC1	PC2	PC1	PC2	
3.1 Collected by a public or private company	0.66	0.08	0.67	0.09	0.73	0.08	
3.2 Burned/buried	-0.57	-0.60	-0.59	-0.53	-0.31	-0.85	
3.3 Dumped/not disposed	-0.49	0.80	-0.44	0.85	-0.61	0.53	
Explained proportion of variance	0.77	0.23	0.73	0.27	0.62	0.38	

management principal components and signs (+/-) of regression *t*-values (Table 2), some conclusions can be drawn. For example, in the case of under-5 children's deaths and hospitalizations in the "basic" and "mid-level" clusters, the *t*-values of PC2s are notably larger (absolute values) than those of PC1s. This indicates that the associations between children's severe diarrhea and PC2s are especially strong in these clusters, even if the PC1s explain much more of the variance of the waste management feature category (Table 3, bottom row). In PC2s, the loading of waste collection is essentially 0, leaving almost all the gravity to dumping trash and burning/burying trash. These factors together might point to the improvement from dumped trash to burning/burying trash being an important way to combat severe diarrhea in young children. Be that as it may, as has been demonstrated in this study as well as other works described in Section 3.5, solid waste management practices require more attention and research.

3.7. Ways forward

Especially in a middle-income country like Brazil, diarrhea hospitalizations and deaths are a complex phenomenon with a myriad of possible explanations, and household conditions are only one of them. Regarding this, in a summarizing multilevel study of low-income countries, Pinzón-Rondón et al. (2015) identified the most important factors increasing the risk of high diarrhea prevalence in under-5 children on the country level. These factors were low education of the mother, a working mother, not being fully vaccinated, and country-level income inequality and poverty. They also found that a nuclear family structure, advanced household sanitation, and household wealth decreased the risk of diarrhea. Many more renowned works in the field have found similar trends (e.g., Thapar and Sanderson, 2004; Sinmegn Mihrete et al., 2014).

The effects of healthcare on diarrhea have also been specifically studied in Brazil. The low prevalence of severe diarrhea is diminished by free and high-quality healthcare provided by the national publicly funded healthcare system (SUS). For example, Rasella et al. (2010) found that the SUS system's Family Health Program (FHP) significantly reduced under-5 diarrhea deaths between 2000 and 2005. Many studies have also noted that after the start of the rotavirus vaccination campaign (2006) for 2- and 4-month-old infants, diarrhea hospitalizations (17–48%) and diarrhea deaths (22–54%) were reduced by the respective percentages in the below-5 age groups (Gurgel et al., 2011; do Carmo et al., 2011; Linhares and Justino, 2014).

Accordingly, household features are not the only factors contributing to the alleviation of the overall diarrheal burden of any country. Nevertheless, the authors argue that by comparing their impacts to each other, this study can help assess which ones should be prioritized to gain maximal health outcomes in different types of municipalities. Clustering could be utilized more often in analyses similar to this one, where localities of interest are distinctly different from each other and may therefore have dissimilar relationships with modeled diseases. Additionally, collecting household and lifestyle data on patients who are hospitalized or dead due to diarrhea (or some other disease) would make researching their determinants much more exact. In a recent article considering governmental planning in Brazilian municipalities, Lima et al. (2020) point out that municipalities are bound by law to prepare municipal plans for sanitation and solid waste management to receive funds from the union for implementing related policies. They also point out that even if Brazilian municipalities have the main responsibility for carrying out related policies, it is not known if they have the capacity to plan and implement them. As a preliminary conclusion, Lima et al. found that municipalities are very heterogeneous in terms of realizing related planning processes. Insight from this study can hopefully be helpful for Brazilian municipalities in drawing up and implementing those plans in the near future.

Perhaps the most surprising finding of this study-that waste management may be more strongly associated with diarrhea than with WASH-requires more attention and research. The body of literature pointing to poor waste management being a serious health concern has grown significantly in the 2000s (Section 3.5). In their recent review paper, Al-Dailami et al. (2022) found that poor waste management threatens health through several mechanisms, such as contributing to spread of disease vectors and contaminating the air, soil, and surface as well as groundwaters. At the moment, however, waste management is not extensively discussed, for example, in the WHO Housing and Health guidelines (2018). Many risks to human health, such as poor air quality, household crowding, and high temperature, are reviewed in the WHO guidelines, but waste management is only discussed in the context of wastewater. Nonetheless, this study points to housing and public health policymakers possibly reaching superior health results by prioritizing waste management relative to some other housing improvements, especially when reducing the prevalence of dumped waste.

3.8. Study limitations

The results of this study on the significance of household features are compared to those of other works that have used a mixture of diarrhea outcomes (mild and severe) as their dependent variables. Their works have been conducted in areas comparable to the "mid-level" or "basic" housing clusters of this study. In the comparisons, it is assumed that the same factors that increase the risk of mild diarrhea also increase the risk of severe diarrhea. To keep this paper as concise as possible, the effects of each municipality type and feature type on each diarrhea outcome are not discussed separately.

Severe diarrhea has many different origins beyond housing. These include environmental and food-related causes, chronic diseases, and so forth (see Section 3.7). Diarrhea diseases also spread from other humans and animals. This is an observational study, which can only capture the correlation between household features and diarrhea on the municipal level. Therefore, causality between tested independent household features and severe diarrhea cannot be directly proven here. The data cannot be used to identify which features specifically affected each case of severe diarrhea. Additionally, the data do not include information on which feature combinations were utilized in households. This is unfortunate, as especially the significance of household water treatment is dependent on the used water source. Furthermore, the data do not specify details about household features, e.g., whether they are functional and appropriately managed by households and stakeholders in each case. It is also likely that some errors have occurred in the data collection phase, especially as the data are so vast. Therefore, as already mentioned, some error removal was performed in clear cases. (See details in the supplementary material.)

Behavioral, economic, and practical factors may affect the ability and willingness to hospitalize patients with diarrhea, although the Brazilian health care system (SUS) is free of cost. Determining the diagnosis for deaths and hospitalizations may also be complicated by comorbidity. Some time has passed since the data were collected, so it does not accurately represent present-day Brazil. However, the development level of the data time period can be assumed to mirror that of many low- and middle-income countries and areas today. Diarrhea hospitalizations and diarrhea deaths are assumed to reflect the unreported overall burden of diarrheal diseases. Milder diarrhea cases are naturally much more prevalent. The authors argue that understanding derived from the data can thus be useful for current policy planning and also in mitigating milder forms of diarrhea.

Using PCA was efficient and even necessary due to the collinearity of household feature variables. Unfortunately, using PCA also obscures the effects of each household feature. However, the effects of each variable category are easily interpretable in the analysis of this study (Table 2). The causal relationship of each housing category with diarrhea has also been shown in other works (see the Sections 3.2–3.5). Therefore, the authors maintain that by comparing the *t*-values and *p*-values of household feature principal components to each other, they can be ranked, and thus preliminary recommendations can be given on how to prioritize the features in municipal policy planning.

4. Conclusion

In Brazil and many other countries, municipalities are the main dutybearers in providing public health services, which is why the design of related policies should also rely on data analyses that fit the intervention scale. We aim to answer the research question: *How do housing features relate to severe diarrhea prevalence in different types of municipalities?* Exceptionally many household and diarrhea variables of vast and rare municipal-level data were utilized for the task.

Based on clustering, municipalities with the most "advanced" housing are in the Southeast region of Brazil and "basic" housing municipalities in the Northeast region, while the rest of the country roughly has "mid-level" housing. Under-5 children's diarrhea deaths relative to the population were three times higher in the "basic" cluster than in the "advanced" cluster. Household features explained severe diarrhea prevalence best in the "advanced" cluster ($R^2 = 16-22$ %) and least in the "basic" cluster ($R^2 = 6-12$ %). Of the different severe diarrhea outcomes, deaths of under-5 children were best explained with household features ($R^2 = 10-22$ %), while those in the over-5 population were the least best explained ($R^2 = 0.3-7$ %).

Unlike previously thought, especially in the water field, improving WASH did not show up as the clear silver bullet for reducing diarrhea in any of the clusters (i.e., municipality types). In contrast, waste management turned out to have the strongest association with severe diarrhea prevalence overall. These findings add value to the existing literature: implications of waste management on diarrhea have not been under scrutiny with a dataset of the extent and granularity as the one employed here. The study brings more scientific backing on waste management's crucial role as a diarrhea determinant and might prove essential for driving change in practical applications, such as housing and health policies.

In conclusion, municipal spending on housing could be tailored to each context to maximize inhabitant health, and waste management may require more attention. The findings of this extensive study may have important implications for public health and housing planning in low- and middle-income countries, where resources are typically scarce, and need to be carefully allocated to policies with best expected response.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data is publicly available (see bibliography). The code will be shared upon request.

Acknowledgments

This study was financially supported by Maa- ja vesitekniikan tuki ry. (MVTT) and Aalto University. The authors also thank Laboratorio Nacional de Ciencias de La Sostenibilidad (LANCIS), Instituto de Ecología (UNAM), for hosting Anni Juvakoski for a research stay at the time of finalizing this manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2023.120805.

References

- Adane, M., Mengistie, B., Medhin, G., Kloos, H., Mulat, W., 2017. Piped water supply interruptions and acute diarrhea among under-five children in Addis Ababa slums, Ethiopia: a matched case-control study. PLoS One 12 (7), e0181516.
- Ahmed, A.S., Halabi, Z., Antoun, J., 2020. The effect of the waste disposal crisis on the rates of hospitalization due to acute diarrheal illness in a middle-income country: retrospective chart review. Int. J. Infect. Dis. 90, 65–70.
- Al-Dailami, A., Ahmad, I., Kamyab, H., Abdullah, N., Koji, I., Ashokkumar, V., Zabara, B., 2022. Sustainable solid waste management in Yemen: environmental, social aspects, and challenges. Biomass Convers. Biorefinery 1–27.
- Alebel, A., Tesema, C., Temesgen, B., Gebrie, A., Petrucka, P., Kibret, G.D., 2018. Prevalence and determinants of diarrhea among under-five children in Ethiopia: a systematic review and meta-analysis. PLoS One 13 (6), e0199684.

Aquino, R., De Oliveira, N.F., Barreto, M.L., 2009. Impact of the family health program on infant mortality in Brazilian municipalities. Am. J. Public Health 99 (1), 87–93.

- Bawankule, R., Singh, A., Kumar, K., Pedgaonkar, S., 2017. Disposal of children's stools and its association with childhood diarrhea in India. BMC Public Health 17 (1), 1–9. Brown, J., Cairncross, S., Ensink, J.H., 2013. Water, sanitation, hygiene and enteric
- infections in children. Arch. Dis. Child. 98 (8), 629–634.
- Bühler, H.F., Ignotti, E., daSilva Neves, S.M.A., Hacon, S.S., 2014. Spatial analysis of integrated determinant indicators of mortality from acute diarrhea in children under 1 year of age in geographical regions. Ciència Saúde Coletiva 19 (10), 4131.
- Cairneross, S., Hunt, C., Boisson, S., Bostoen, K., Curtis, V., Fung, I.C., Schmidt, W.P., 2010. Water, sanitation and hygiene for the prevention of diarrhoea. Int. J. Epidemiol. 39 (suppl_1), i193–i205.
- Carlton, E.J., Eisenberg, J.N., Goldstick, J., Cevallos, W., Trostle, J., Levy, K., 2014. Heavy rainfall events and diarrhea incidence: the role of social and environmental factors. Am. J. Epidemiol. 179 (3), 344–352.
- Contreras, J.D., Eisenberg, J.N., 2020. Does basic sanitation prevent diarrhea? Contextualizing recent intervention trials through a historical lens. Int. J. Environ. Res. Public Health 17 (1), 230.
- Corona, F., Mulas, M., 2022. Conversations Over Zoom in 2020-2022.
- Coube, M., Nikoloski, Z., Mrejen, M., Mossialos, E., 2023. Persistent inequalities in health care services utilisation in Brazil (1998–2019). Int J Equity Health 22 (1), 1–15. Crawford, T.V., 2009. Alternate electrification and non-potable water: a health concern
- for Jamaicans. N. Am. J. Med. Sci. 1 (7), 338. Cruvinel, V.R.N., Zolnikov, T.R., Bashash, M., Marques, C.P., Scott, J.A., 2019.
- Waterborne diseases in waste pickers of Estrutural, Brazil, the second largest openair dumpsite in world. Waste Manag. 99, 71–78.
- Cruvinel, V.R.N., Zolnikov, T.R., Obara, M.T., de Oliveira, V.T.L., Vianna, E.N., do Santos, F.S.G., Scott, J.A., 2020. Vector-borne diseases in waste pickers in Brasilia, Brazil. Waste Manag. 105, 223–232.

A. Juvakoski et al.

Darvesh, N., Das, J.K., Vaivada, T., Gaffey, M.F., Rasanathan, K., Bhutta, Z.A., 2017. Water, sanitation and hygiene interventions for acute childhood diarrhea: a systematic review to provide estimates for the lives saved tool. BMC Public Health 17 (4), 101–111.

- DATASUS (Ministério da Saúde Sistema de Informações Hospitalares do SUS (SIH/ SUS), 2023a. Database of hospitalizations in Brazil. http://tabnet.datasus.gov.br/cg i/tabcgi.exe?sih/cnv/niuf.def (accessed 22.4.2020).
- DATASUS (Ministério da Saúde Sistema de Informações sobre Mortalidade (SIM), 2023b. Database of mortalities in Brazil. http://tabnet.datasus.gov.br/cgi/tabcgi. exe?sim/cnv/obt10br.def. (accessed 22.4.2020).
- Degebasa, M.Z., Weldemichael, D.Z., Marama, M.T., 2018. Diarrheal status and associated factors in under five years old children in relation to implemented and unimplemented community-led total sanitation and hygiene in Yaya Gulele in 2017. Pediatric Health, Med. Ther., 9, 109.
- ..., do Carmo, G.M.I., Yen, C., Cortes, J., Siqueira, A.A., de Oliveira, W.K., Cortez-Escalante, J.J., Patel, M., 2011. Decline in diarrhea mortality and admissions after routine childhood rotavirus immunization in Brazil: a time-series analysis. PLoS Med. 8 (4), e1001024.
- Fewtrell, L., Kaufmann, R.B., Kay, D., Enanoria, W., Haller, L., Colford Jr., J.M., 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. Lancet Infect. Dis. 5 (1), 42–52.
- Fuller, J.A., Eisenberg, J.N., 2016. Herd protection from drinking water, sanitation, and hygiene interventions. Am. J. Trop. Med. Hyg. 95 (5), 1201.
- Gurgel, R.Q., Ilozue, C., Correia, J.B., Centenari, C., Oliveira, S.M., Cuevas, L.E., 2011. Impact of rotavirus vaccination on diarrhoea mortality and hospital admissions in Brazil. Trop. Med. Int. Health 16 (9), 1180–1184.
- Jung, Y.T., Hum, R.J., Lou, W., Cheng, Y.L., 2017. Effects of neighbourhood and household sanitation conditions on diarrhea morbidity: systematic review and metaanalysis. PLoS One 12 (3), e0173808.
- Lima, L.L., Lui, L., Dias, G.V.R.S., Papi, L.P., Demarco, D.J., 2020. Government planning in Brazilian municipalities: towards a research agenda. Cadernos EBAPE.BR 18, 323–335.
- Linhares, A.C., Justino, M.C.A., 2014. Rotavirus vaccination in Brazil: effectiveness and health impact seven years post-introduction. Expert Rev. Vaccines 13 (1), 43–57.
- Machoski, E., de Araujo, J.M., 2020. Corruption in public health and its effects on the economic growth of Brazilian municipalities. Eur. J. Health Econ. 21 (5), 669–687.
 MacQueen, J., 1967. Classification and analysis of multivariate observations. In: 5th
- Berkeley Symp. Math. Statist. Probability, pp. 281–297.
 Mberu, B., Mutua, M., Kabaria, C., Amugsi, D., Muindi, K., 2022. Levels of household exposure to solid waste dumpsites and associated loss to health in urban Kenya and Senegal. Cities Health 6 (1), 168–179.
- Melese, B., Paulos, W., Astawesegn, F.H., Gelgelu, T.B., 2019. Prevalence of diarrheal diseases and associated factors among under-five children in Dale District, Sidama zone, Southern Ethiopia: a cross-sectional study. BMC Public Health 19 (1), 1–10.
- Melo, M.C.N.D., Taddei, J.A., Diniz-Santos, D.R., Vieira, C., Carneiro, N.B., Melo, R.F., Silva, L.R., 2008. Incidence of diarrhea in children living in urban slums in Salvador, Brazil. Braz. J. Infect. Dis. 12 (1), 89–93.
- Nandi, A., Megiddo, I., Ashok, A., Verma, A., Laxminarayan, R., 2017. Reduced burden of childhood diarrheal diseases through increased access to water and sanitation in India: a modeling analysis. Soc. Sci. Med. 180, 181–192.
- Ndikubwimana, J.B., Ngendahimana, F., 2020. Risk factors associated with under-five diarrhea and their effect on under-five mortality in Rwanda: secondary data analysis of 2014–2015 Rwanda demographic and health survey. Available at [SSRN 3601648].
- Neves, E.M.S.C., 2012. Environmental policy, municipalities and intergovernmental cooperation in Brazil. Estud. Avançados 26, 137–150.
- OECD iLibrary, 2021. OECD reviews of health systems: Brazil 2021. 4. Health Data Infrastructure and Information System in Brazil | OECD Reviews of Health Systems: Brazil 2021 | OECD iLibrary (oecd-ilibrary.org) (accessed 28.9.2023).
- Patrícia, S.D.A., Ribeiro Jr, H.D.C., Mendes, C.M.C., 2013. Temporal trends of overall mortality and hospital morbidity due to diarrheal disease in Brazilian children younger than 5 years from 2000 to 2010. J. Pediatr. 89 (3), 315–325.
- Pattanayak, S.K., Wendland, K.J., 2007. Nature's care: diarrhea, watershed protection, and biodiversity conservation in Flores, Indonesia. Biodivers. Conserv. 16, 2801–2819.
- Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. Philos. Mag. 2, 559–572.
- Penakalapati, G., Swarthout, J., Delahoy, M.J., McAliley, L., Wodnik, B., Levy, K., Freeman, M.C., 2017. Exposure to animal feces and human health: a systematic

review and proposed research priorities. Environ. Sci. Technol. 51 (20), 11537–11552.

- Pinzón-Rondón, Á.M., Zárate-Ardila, C., Hoyos-Martínez, A., Ruiz-Sternberg, Á.M., Vélez-van-Meerbeke, A., 2015. Country characteristics and acute diarrhea in children from developing nations: a multilevel study. BMC Public Health 15, 1–11.
- Rahman, M.M., Bohara, A.K., Vazquez, J.E., 2021. Geospatial analysis of health risks and solid waste management behaviour. J. Environ. Econ. Policy 10 (4), 400–427.
- Randremanana, R.V., Razafindratsimandresy, R., Andriatahina, T., Randriamanantena, A., Ravelomanana, L., Randrianirina, F., Richard, V., 2016. Etiologies, risk factors and impact of severe diarrhea in the under-fives in Moramanga and Antananarivo, Madagascar. PLoS One 11 (7), e0158862.
- Rasella, D., Aquino, R., Barreto, M.L., 2010. Reducing childhood mortality from diarrhea and lower respiratory tract infections in Brazil. Pediatrics 126 (3), e534–e540.
- Reeve, B., Ashe, M., Farias, R., Gostin, L., 2015. State and municipal innovations in obesity policy: why localities remain a necessary laboratory for innovation. Am. J. Public Health 105 (3), 442–450.
- Rego, R.F., Moraes, L.R.S., Dourado, I., 2005. Diarrhoea and garbage disposal in Salvador, Brazil. Trans. R. Soc. Trop. Med. Hyg. 99 (1), 48–54.
- Roquette, R., Nunes, B., Painho, M., 2018. The relevance of spatial aggregation level and of applied methods in the analysis of geographical distribution of cancer mortality in mainland Portugal (2009–2013). Popul. Health Metr. 16 (1), 1–12.
- Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.
- Sahiledengle, B., Teferu, Z., Tekalegn, Y., Zenbaba, D., Seyoum, K., Atlaw, D., Chattu, V. K., 2021. A multilevel analysis of factors associated with childhood diarrhea in Ethiopia. Environ. Health Insights 15, 11786302211009894.
- Samad, H.A., Zhang, F., 2018. Electrification and household welfare: evidence from Pakistan. World Bank policy research working paper (8582).
- Sinmegn Mihrete, T., Asres Alemie, G., Shimeka Teferra, A., 2014. Determinants of childhood diarrhea among under-five children in Benishangul Gumuz regional state, northwest Ethiopia. BMC Pediatr. 14 (1), 1–9.
- Sistema de Informação da Atenção Básica (SIAB) by DATASUS (Ministério da Saúde), 2023. Household feature data from Brazilian municipalities. http://www2.datasus. gov.br/SIAB/index.php?area=01 (accessed 2.3.2020).
- Soboksa, N.E., 2021. Associations between improved water supply and sanitation usage and childhood diarrhea in Ethiopia: an analysis of the 2016 demographic and health survey. Environ. Health Insights 15, 11786302211002552. https://doi.org/ 10.1177/11786302211002552.
- Statista, 2022. Consumer Market Insights: Non-Alcoholic Drinks: Bottled Water (Brazil). https://www.statista.com/outlook/cmo/non-alcoholic-drinks/bottled-water/brazil (accessed 24.2.2023).
- Strina, A., Cairncross, S., Barreto, M.L., Larrea, C., Prado, M.D.S., 2003. Childhood diarrhea and observed hygiene behavior in Salvador, Brazil. Am. J. Epidemiol. 157 (11), 1032–1038.
- Szwarcwald, C.L., Souza Júnior, P.R.B.D., Marques, A.P., Almeida, W.D.S.D., Montilla, D. E.R., 2016. Inequalities in healthy life expectancy by Brazilian geographic regions: findings from the National Health Survey, 2013. Int. J. Equity Health 15 (1), 1–9.
- Thapar, N., Sanderson, I.R., 2004. Diarrhoea in children: an interface between developing and developed countries. Lancet 363 (9409), 641–653.
- Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18 (4), 267–276.
- Troeger, C.E., Khalil, I.A., Blacker, B.F., Biehl, M.H., Albertson, S.B., Zimsen, S.R., Seyedmousavi, S., 2020. Quantifying risks and interventions that have affected the burden of diarrhoea among children younger than 5 years: an analysis of the global Burden of disease study 2017. Lancet Infect. Dis. 20 (1), 37–59.
- UNICEF, 2022. Diarrhoea. https://data.unicef.org/topic/child-health/diarrhoeal-di sease/ (accessed 13.2.2023).
- Wakida, F.T., Lerner, D.N., 2005. Non-agricultural sources of groundwater nitrate: a review and case study. Water Res. 39 (1), 3–16.
- WHO, 2017. Diarrheal disease. https://www.who.int/news-room/fact-sheets/detail/dia rrhoeal-disease (accessed 20.1.2020).
- WHO, 2018. WHO housing and health guidelines. Geneva, Switzerland (2018).World Bank, 2020. Gini index. https://data.worldbank.org/indicator/SLPOV.GINI (accessed 7.3.2023).
- Yaya, S., Hudani, A., Udenigwe, O., Shah, V., Ekholuenetale, M., Bishwajit, G., 2018. Improving water, sanitation and hygiene practices, and housing quality to prevent diarrhea among under-five children in Nigeria. Trop. Med. Infect. Dis. 3 (2), 41.
- Zambrano, L.D., Levy, K., Menezes, N.P., Freeman, M.C., 2014. Human diarrhea infections associated with domestic animal husbandry: a systematic review and meta-analysis. Trans. R. Soc. Trop. Med. Hyg. 108 (6), 313–325.