RESEARCH ARTICLE

Theory and health disparities: Explaining the relevance of theory in public health research in India

JAGRITI GANGOPADHYAY, MALAVIKA SUBRAMANYAM

Abstract

This paper aims to address the concern that quantitative public health studies do not integrate theoretical considerations. This qualitative study uses content analysis to explore the application of theory in Indian public health articles listed on PubMed. Social determinants such as poverty, income, social class, education, gender, caste, socioeconomic position, socioeconomic status, immigrant status, and wealth were the keywords used to identify the articles analysed in this study. From a selection of 91 public health articles, we identified potential theoretical frameworks based on the pathways/recommendations/explanations mentioned. Additionally, using the case of tuberculosis in India, we highlight how theoretical perspectives play a critical role in providing a holistic view of major health challenges. Finally, by emphasising the need to adopt a theoretical perspective in empirical quantitative research on public health in India, we hope to motivate scholars to include a theory or theoretical paradigm in their future research.

Keywords: theory, social determinants, qualitative, content analysis, India

Introduction

Justification and motivation

Health-related social disparities have been receiving greater attention in the public health discourse in India over the last few years. By “social disparities”, we mean unjust inequalities in health outcomes in relation to tuberculosis, malaria, anaemia, under-nutrition, infant and child mortality, and so on across various social determinants such as poverty, income, gender, wealth, education, caste, and socioeconomic status [1-4]. Notably, social disparities in health in India have been documented across a wide range of health outcomes including mortality, communicable and non-communicable diseases, nutrition, mental health, risk factors of non-communicable diseases, and injuries [3]. Given that societal determinants of health are likely to contribute to the causation and distribution of several diseases, such a focus on social disparities is timely [5].

Need for theoretical conceptualisation

Given the multidisciplinary nature of public health, as researchers, we believe that theory building is an essential component that will enhance the academic rigour of public health studies. Theoretical frameworks in disciplines as diverse as public health provide a scientific approach for scholars [6]. Additionally, theory adds meaning to empirical observations and findings [7]. It also helps generate new knowledge and advances public health as a discipline [6]. Commenting on the ethics of using theory in public health research, scholars have highlighted that while studying health inequalities across races, genders, and classes, it is important to draw on theoretical frameworks of health that have an effect on race, gender and class to avoid unfounded interpretations based on empirical data [6, 8, 9]. We argue that not explicitly acknowledging the theoretical perspective adopted in a paper reduces the transparency required by readers to evaluate its claims. As such, it would be ethical for authors to clearly state the theories or models that inform their research questions as well as their interpretation of the results.

The importance of theory has been stressed both in quantitative and qualitative research [10, 11, 12]. Scholars have noted that with the help of theories, researchers can explain any phenomenon logically and meaningfully [13]. In particular, theory-driven analysis aids researchers to make assumptions transparent, improve accountability for the data, and assists in cumulating and improving the quality of research [14]. For instance, application of theory in social epidemiological quantitative research might follow a path such as identifying a problem, reviewing how various relevant theories help researchers understand the problem.
developing a theory-informed conceptual model that represents the assumptions made by the researchers, developing measurable indicators that represent those assumptions, designing a study that measures such indicators, and so on. Maintaining a balance between data and theory will also provide greater clarity to researchers in their pursuit of knowledge [12]. As discussed by these scholars, theory plays a key role in developing good quality research. In this context, we demonstrate how the use of theory will contribute to producing better quality public health and social epidemiological research in India.

Several theoretical perspectives are recognised by social epidemiology — an integral component of public health — which seeks to identify how societal characteristics affect the pattern of disease and health distribution in a society as well as their pathways of influence [15]. Social epidemiology essentially addresses the impact of both societal factors and of pathophysiology on disease [16]. In particular, theory aids in understanding and explaining the intersections between society and biology in social epidemiological research. Engaging with "notions of causation", in turn, raises not only complex philosophical issues but also, in the case of social epidemiology, issues of accountability and agency — simply invoking abstract notions of "society" and disembodied "genes" will not suffice [16]. Specifically, theory assists scholars of social epidemiology to critically and systematically examine, assimilate, and intimate connections between human social and biological existence. The theoretical perspective employed by researchers plays an important role in determining the kinds of questions that are investigated and the kinds of policies that are formulated on the basis of such research.

Several theories and models have been applied in the discipline of social epidemiology and they can be classified into two broad categories based on their characteristics. The first category includes theories that emphasise the attributes of individuals as causal factors while investigating health disparities in a population. For instance, the epidemiological triad — the biomedical model, the lifestyle model, and the web of causation [17]. The other category comprises theories that underscore the importance of the social environment and other macro-level determinants of health. One such theory is the psychosocial approach, which highlights the impact of the social environment and human interaction on disease distribution, primarily by invoking the stress pathway that links social disadvantage with worse health outcomes, while highlighting the buffering effect of social networks. Another theoretical framework in this category is the social production of disease, which explicitly identifies economic and political institutions as the causal drivers of health inequalities. This theory argues that the unequal distribution of power and resources in society and the resultant societal structure determine exposure to health threats and access to health-protective resources.

The perils of not acknowledging or critically reviewing the theoretical perspectives guiding public health research have implications for public health programmes and policies as well. Let us consider the history of pulmonary tuberculosis (henceforth referred to as TB) control in India. The first organised public health approach to TB control in independent India viewed TB control — a public health issue — as identical to the diagnosis and treatment of TB in an individual patient [18]. The state attempted to control the disease by applying a biomedical lens to the community and treating as many individual cases as possible, and considered the population a sum of all its constituent individuals. Such an approach targets causal factors at the individual level and ignores those that may be operating at other levels. For instance, factors such as poor housing conditions lead to repeated infections and sapping immunity while poor distribution of agricultural produce lead to undernutrition among individuals (which increases their risk of developing pulmonary TB). Despite decades of programming that was frequently reviewed and revised, it was only after 2006 that the social context of TB received attention in the national TB control efforts. This was evident in the resulting shift in perception, where individuals were viewed as active participants in the care of their own health [19] and not just as passive cases that need to be cured. Arguably, the emphasis was on more proximal social factors such as an individual's awareness, intention, and self-efficacy, and not on upstream factors such as poverty, housing, or distribution of food. Moreover, there was a reliance on information, education, and communication (IEC) interventions, which could have led to a widening of social inequalities [20]. However, it still suggests a broadening in the conceptual model driving TB control efforts in India.

All of this raises questions such as: Would the national TB control programme have been more successful if its theoretical model had been acknowledged and critically reviewed? Would such a review have led to a discussion on potential approaches that do not rely solely on the biomedical model? Could this have led to greater attention and funding being allocated to IEC activities and the earlier inclusion of the private sector? Did efforts targeting causal factors at multiple levels and across multiple sectors lead to better TB control outcomes? These questions need to be asked in order to design effective programmes in the future.

Although we intend to expand the scope of public health as a discipline, we focus on social epidemiological theories, as these explain how health inequalities, which are at the intersection of society and biology, continue to persist across India. In the next section, through a few examples, we demonstrate how we have analysed the articles discussed in the Supplementary Table [References 21-89 have been cited in the Table, available online only].

**Illustration of a few instances**

Let us take the example of an article examining the impact of socioeconomic status (SES) and living conditions on latent
tuberculosis infection (LTBI) in a tribal population in Melghat, Maharashtra [37]. The authors argue that poor SES and living conditions (LC) are “the most important risk factors reported, since they are invariably associated with poverty, malnutrition, hygiene, and illiteracy all of which have confounding effect on outcome of both active and LTBI in high TB endemic regions” (p 373). They also posit that “Low SES status has direct impact on LC of the tribal population. Both factors ultimately may dispose other risk factors which include illiteracy, poor hygiene practices, and poor diet” (p 378). On the basis of such reasoning, they interpret the results of their study — that low SES and poor living conditions were associated with greater risk of a diagnosis of LTBI, suggesting that “.....if policy makers extend their comprehensive and integrated approach of disease control by targeting at least household level factors, like SES and LC, the prevalence of LTBI in such isolated regions of the country would be much under control” (p 379). One aim of this study was to investigate the impact of low SES and poor living conditions on LTBI and the other was to understand the relationship between those two factors and a certain diagnostic test for TB. The study’s emphasis on laboratory investigations and early diagnosis of TB, and the argument linking low SES and poor living conditions with LTBI, through social (poverty and illiteracy), nutritional (poor diet, malnutrition), and behavioural (hygiene) pathways without theorising why/how such pathways might operate, all suggest the influence of biomedical and lifestyle models [16]. The influence of the biomedical and lifestyle models can also be observed in how the authors describe SES and living conditions and the consequent factors that link them with LTBI — which they describe as having a “confounding effect” on the incidence of TB. We may recall that both the biomedical and lifestyle models emphasise the causal role of individual-level factors while treating potential societal determinants of health as nuisance factors or confounders. Such an individual-focused perspective is also evident in the recommendation that policymakers “target” “household level factors, like “SES and LC” to reduce the prevalence of LTBI, without discussing how such factors should be targeted. The authors also fail to suggest what focus, if any, policymakers should have on the factors that serve as pathways connecting low SES and poor living conditions with LTBI.

In contrast, if we apply the social production of disease theory to interpret the same results, we might conclude that the societal structure — that is, the unequal distribution of resources and opportunities — increases the risk that the socially disadvantaged live in poverty and face a lack of access to a healthful diet (causing malnutrition) and of educational opportunities (leading to low access to health-promoting information and lower confidence in navigating the health system), which increase their risk of contracting TB. The theory suggests that such disadvantages due to the societal structure and prevalent distribution of power and resources increase the risk of exposure for persons with low SES and poor living conditions to the TB bacterium and further increase the risk of such exposure leading to LTBI. Furthermore, applying the theory of social production of disease would encourage authors to consider the larger societal context — the social, political, and economic forces — shaping the lives of the tribal population in Melghat. Such arguments, premised on the theory of social production of disease, advocate that policymakers target the unequal distribution of power, resources, and opportunities in this population to prevent LTBI. This is in direct contrast with the recommendation of Kashyap et al [37], wherein policymakers were encouraged to view low SES and poor living conditions as markers indicating a greater risk of incidence of LTBI, and thus, to focus their efforts on such sub-groups to lower the prevalence of LTBI. In fact, we argue that the influence of different theories is evident even in the articulation of the research question. We submit that authors applying the social production of disease theory would not refer to low SES and poor living conditions as “risk factors” or as having a “confounding effect” nor would they miss the larger context. Instead, we argue that they would articulate the question as “Do lower SES and poorer living conditions increase the risk of being diagnosed with LTBI in an underprivileged tribal population?” Thus, we observe how the theoretical perspective adopted while conceptualising a research question can lead to diverse interpretations of the same association (low SES and LTBI) in a given population (tribal population of Melghat), thus affecting the conclusions and recommendations from such studies. In an application-oriented field such as public health, the differing interpretations of the same association can lead to drastically different public health programmes and policies.

Similarly, we could argue that the application of theory would have led to more robust interpretations of the findings in Safraj et al [83] of a positive association between socioeconomic position and the prevalence of diabetes in Kerala. It is disappointing that they do not offer any explanation of their findings, although they methodically rule out the influence of other determinants in explaining this result, such as the extent of physical activity, family history of diabetes, and the differential rates of diagnosis of diabetes across varying socioeconomic positions. The only sentence in the entire article that hints at the theoretical perspective adopted by the authors in interpreting the result of their study is this: “The state has witnessed dramatic improvements in the quality of life of people, consequent to urbanization. Diets are getting richer and physical activity is declining as the middle class is burgeoning. The burden of diabetes is likely to be much higher in the immediate future, calling for effective preventive strategies” (p 6). Unfortunately, these sentences do not directly engage with the result, nor do they explicitly clarify the theoretical perspective the authors draw from. On the other hand, applying the social–ecological model to interpret their results would lead us to argue that changes in the social environment, including greater exposure to conspicuous consumption, leading consequentially to the greater prevalence of a sedentary lifestyle, more consumption of
carbohydrate-rich and nutrient-poor foods, and a culture of snacking and consuming sugary beverages instead of water, by persons of a higher versus lower socioeconomic position, likely explains the results obtained by Safraj et al [83].

In particular, applying theory in social epidemiology could prompt researchers to examine the impact of social determinants on health and bridge the gap between an individual- and population-level focus on disease prevention [90]. Including theory in research will aid social epidemiologists in constructing research designs that cover both micro- and macro-level health outcomes, since theoretical frameworks cover a range of approaches that cover factors from genetics to the environment [90].

Similarly, research on social disparities in health in India has been criticised for not paying sufficient attention to theory [91, 92]. However, the extent to which theoretical perspectives are applied or even explicitly/implicitly acknowledged in such work has not been empirically investigated. Therefore, the goal of this study is to highlight the significance of theory in social epidemiological studies in India; through this exercise, we urge future scholars of social epidemiology and public health conducting quantitative studies in India to explicitly mention the theoretical framework used in their research. In the remaining sections, we describe our review of the use of theory in articles examining the impact of social determinants on health in India and our conclusions.

Methodology
A simple count of published articles that use a theoretical framework would provide a comprehensive analysis; however, going beyond a simple count and using qualitative methods to analyse how and to what extent theory has been applied is more informative. Such a qualitative analysis may shed light on the extent of the current use of theory and initiate a conversation about the implications of theory-based research for policy and programmes.

Preliminary work performed as part of a class assignment led us to suspect that most India-focused social epidemiological research does not explicitly state its theoretical influence. Discussions with a few professors teaching public health at different universities in India suggested that a focus on theory was rare in the Indian context when it came to public health degree programmes. Hence, we hypothesised that India-based studies that use and acknowledge theory are primarily authored by researchers based outside India. While these hypotheses are based on preliminary readings and multiple personal conversations, we clearly state them because we wanted to be transparent about the a priori assumptions we made at the time we began this study.

Furthermore, we expect that every article included in this study will still have theoretical underpinnings, whether explicitly stated or not. We aim to examine the practice of acknowledging one’s theoretical perspectives while undertaking public health research in the Indian context.

Our goal was not to accurately identify the theory that the authors have employed. Our goal was to demonstrate that statements made by the authors suggest one or more theoretical perspective(s). We have presented our best-educated guess in case the authors have not explicitly identified the theory they have used.

Inclusion criteria of the articles analysed
Quantitative studies on social disparities in health and studies published in 1990–2019 were included, to present a contemporary picture. To maintain a focus on articles that primarily examine the association of a social factor with health, we restricted our search to studies that mention the keywords in the title or abstract of the article. We also searched for articles with either disparity/disparities or inequality/inequalities in the title or abstract, in combination with “India” (in “Title OR Abstract”).

Exclusion criteria
Articles on sexually transmitted diseases and mental health outcomes were excluded as these studies have a history of drawing on social and psychological theories. We also excluded qualitative studies as the approach entails acknowledging the lens used. Finally, we did not consider studies without implicit or explicit theoretical statements (pathways/policy recommendations/explanations).

Literature search
The search for articles to be included in the sample was conducted on PubMed. Since PubMed is the largest search engine for the MEDLINE database for life sciences and biomedical research, we conducted our search on PubMed. We used several keywords denoting the social determinants of health (in “Title OR Abstract”) and the word “India” (in “Title OR Abstract”). The following social determinants were employed in the search: poverty, income, gender, wealth, education, caste, social class, religion, immigration, socioeconomic position and socioeconomic status. The selected social determinants are some of the major factors impacting the health of individuals as suggested by the World Health Organization [93]. However, we did not include race or ethnicity as these terms are not relevant when conceptualising social determinants of health in the Indian context. Ethnicity is invoked to indicate the historical and geographical connection of a person to a place in India and is not a social determinant of health. Additionally, we included “religion” as a key search word but did not find any articles that met our inclusion criteria. While there may be a few studies that compare specific religions with each other, we hesitate to include them as the religious labels might be proxies for several sociocultural implications.

Content analysis: Suitability of the approach and collation of its findings
To identify the underlying theory that might have been used in studies that did not explicitly name a theoretical
framework, we used content analysis as a qualitative methodological tool. In simple terms, content analysis is the study of documents and communication in either textual or audio–visual format [94]. Content analysis has been used widely in both quantitative and qualitative studies. Quantitative content analysis involves coding key terms and analysing the frequency of their occurrence [95] — a simple example would be the frequency with which a phrase of interest, such as “gender discrimination”, appears in a text. On the other hand, qualitative content analysis is used to interpret and analyse the multiple layers and meanings of a particular document or text [96]. In particular, studies using qualitative content analysis emphasise subjective interpretations of the text [97]. Within the qualitative content analysis method, three approaches are most widely used: conventional, directed, and summative. The conventional content analysis approach derives coding categories directly from the text data. The directed approach involves creating the initial codes based on a particular theory.

The summative content analysis begins with identifying and quantifying particular words in the text to explain the contextual use of the words or content [98]. The main aim of summative analysis is to highlight the usage of the word and understand its context. In this type of content analysis, the focus is on ascertaining the underlying meanings of the content [99]. Finally, the summative approach involves comparing keywords to interpret the text [100]. We chose summative content analysis to examine how the use of certain keywords determines the theoretical tone. We used summative content analysis as it goes beyond examining the frequency of terms and fleshes out their inner meaning or content [100]. In summative content analysis, the data analysis begins with the identification of the frequency of the keywords. Following this, the context in which the word was used is explored and interpreted.

The social determinants listed above (in the Literature Search section) acted as the keywords, and the context of these social determinants in the statements related to explanations, pathways, or policy recommendations, in the selected articles, were analysed. The next and final step involves explaining why that particular theoretical model was suggested for that study.

Since all the articles included only implicit theoretical postulations and no explicit statements, we performed summative content analysis to speculatively identify the theories most relevant for each of the articles.

To further elaborate on our methods, we have demonstrated them step-by-step with the help of a flowchart [Figure 1].

Taken together, we found support for our hypothesis. No studies explicitly identified the theory adopted, and an overwhelming majority included statements that implicitly drew on one or more theoretical approaches. However, four of these studies did have references that drew from theoretical frameworks; nonetheless, an explicit mention of the theory was missing in the article.

![Image](image-url)

**Figure 1: Step-by-step demonstration of the study methods**

Using social determinants of health (in “Title OR Abstract”) and the word “India” (in “Title OR Abstract”) looked for articles in PubMed.

Articles that met the inclusion criteria that were analyzed using summative content analysis (n = 91)

Next, an explanation justifying why that theoretical model was suggested for that particular study was added.

Based on these statements, the authors identified approaches likely employed in each of the articles.

Articles that included statements in the form of explanations, pathways, or policy recommendations that suggested the theoretical approach (n = 79)

The total number of articles that met the inclusion criteria was 91. Seventy-nine of these articles included statements in the form of explanations, pathways, or policy recommendations that suggested the theoretical approach applied (detailed breakdown of the findings have been indicated in Table 1). Based on these statements, the authors identified theoretical approaches for each of the articles.

<table>
<thead>
<tr>
<th>Social factor</th>
<th>Number of articles that PubMed listed for each search</th>
<th>Number of articles that met the inclusion criteria</th>
<th>Number of articles that explicitly identified a theoretical perspective</th>
<th>Number of articles with explanation/pathway/policy recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poverty</td>
<td>114</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Income</td>
<td>333</td>
<td>8</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Wealth</td>
<td>32</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Gender</td>
<td>347</td>
<td>18</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Caste</td>
<td>106</td>
<td>12</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Education</td>
<td>598</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social Class</td>
<td>24</td>
<td>16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Socioeconomic Position</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>59</td>
<td>10</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Religion</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Immigration Status</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Disparity/Disparities</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Inequality/Inequalities</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>1629</td>
<td>91</td>
<td>0</td>
<td>79</td>
</tr>
</tbody>
</table>

**Table 1: Distribution of articles on social determinants of health and their acknowledgment of theoretical perspective**

**Discussion and critical observations**

This paper began with the underlying research hypothesis that studies on social disparities in health in India do not draw from theoretical frameworks. Our summative content analysis of the published literature found support for this
hypothesis. Additionally, we had hypothesised that while the studies may not state the theoretical lens, they still advocate for a policy perspective. Most of the studies we analysed concluded with a policy recommendation. Finally, we also postulated that most of the studies would include statements that suggest potential theoretical paradigms. This hypothesis was also supported as most of the articles included such statements.

By performing summative content analysis, we found that none of the articles that met our inclusion criteria explicitly mentioned a theoretical framework. However, most of the articles on the socioeconomic determinants of health in the Indian context did have statements that allowed us to make an educated guess about the theoretical approach applied in those studies. This suggests that social epidemiological studies in India are not atheoretical and are conceptualised based on one or more theoretical approaches. However, the need to acknowledge the underlying theoretical lens has not been recognised by the authors of these studies.

Health disparities research typically focuses on questions regarding the extent, trends, and causes of disparities and ways to narrow or eliminate them. However, decisions related to how disparities are conceptualised and measured, which disparities are studied and which are not, which causes of disparities are investigated and which are not are all determined by the theoretical perspective adopted by researchers. We argue that a critical review of the underlying theory could lead, at the very least, to a more comprehensive understanding of the issue, which is imperative for designing effective and impactful programmes and policies.

We hope that our analysis sparks debate on the importance of theoretical perspectives in health disparities and public health research in the Indian context. Analysing and applying theory may help in avoiding errors, inspiring new ideas, and critically conceptualising the methods used in different studies and programmes.

**Conclusion**

We submit that theoretical frameworks enhance studies on health outcomes in India by influencing the types of research questions focused on and thus the policy/programme implications of the findings of such studies. Additionally, ignoring theory has great implications for public health practice as well. For instance, the numerous revisions of the TB prevention programme did not differ extensively in their theoretical perspectives, thus missing an opportunity to address the underlying social determinants of TB in India. Additionally, given India’s diversity and regional specificities, adopting theoretical models that acknowledge the importance of the social environment, in addition to “traditional” risk factors, is likely to lead to a better understanding of social disparities in health. Furthermore, context-appropriate theory-building might facilitate the global exchange of ideas to understand why different populations show different kinds of disease profiles [101].

Such a practice may help in lowering the chance of harming individuals and increasing the chance of formulating policies that prevent diseases, promote health equity, and improve public health.

**References**

22. Singh RB, Beegom R, Verma SP, Haque M, Singh R, Mehta AS, De AK, Kundu S, Roy S, Krishnan A, Simhadri H. Association of dietary factors and other coronary risk factors with social class in women in


