Journal of Clinical Epidemiology 177 (2025) 111611

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Directed acyclic graph helps to understand the causality of malnutrition in under-5 children born small for gestational age

Soumya Tiwari, Viswas Chhapola*, Nisha Chaudhary, Lokesh Sharma

Department of Pediatrics, Lady Hardinge Medical College, New Delhi, India

Accepted 11 November 2024; Published online 15 November 2024

Abstract

Objectives: Small-for-gestational age (SGA) is a causal factor for malnutrition (undernutrition). The available evidence on this causal relationship is based on observational studies and suffers from confounding and collider biases. This study aimed to construct a theoretical causal model to estimate the effect of SGA on malnutrition in children aged less than 5 years.

Methods: For the causal model, we designated term-SGA status as the exposure variable and malnutrition at 6 months to 5 years of age (diagnosed by World Health Organization criteria) as the outcome variable. Causal estimands were formulated for three stakeholders. A ‘‘rapid narrative review’’ methodology was adopted for literature synthesis. Studies (observational and randomized) listing the causal factors of malnutrition in children aged less than 5 years from the Indian subcontinent were eligible. Four databases (PubMed, Scopus, Web of Science, and ProQuest) were searched and restricted to the last 10 years (search date: December 15, 2023). Information about the causal factors (covariates) of malnutrition and study characteristics was extracted from the article abstracts. Next, a causal model in the form of a directed acyclic graph (DAG) (DAGitty software) was constructed by connecting exposure, outcome, and covariate nodes using the sequential causal criteria of temporality, face validity, recourse to theory, and counterfactual thought experiments.

Results: The search yielded 4818 records, of which 342 abstracts were included. Most of the studies were conducted in India (39%) and Bangladesh (27%). The literature synthesis identified 81 factors that were grouped into 17 nodes, referring to 5 domains: socioeconomic, parental, child-related, environmental, and political. The DAG identified 12 different minimal sufficient adjustment sets (conditioning sets for regression analysis) to estimate the total effect of SGA on malnutrition.

Conclusion: We offer an evidence-based causal diagram that will minimize bias due to improper selection of factors in studies focusing on malnutrition in term-SGA infants. The DAG and adjustment sets will facilitate the design and data analysis of future studies.

Keywords: Directed acyclic graph; DAG; Causal diagram; Small for gestational age; Malnutrition; Causality; Confounding; Collider bias

1. Introduction

Small-for-gestational age (SGA) is defined as being born with birth weight below 10th percentile for gestational age [1]. Approximately 16% of all births are SGA, with rates ranging from 7% in high-income countries to 41.5% in South Asia [2−6]. Although a diverse array of maternal, environmental, placental, and fetal factors have been implicated, a lack of nutrient supply to the fetus is believed to be the main cause of reduced fetal growth in SGA infants [7−10]. SGA has its own set of consequences, including a higher risk of mortality than its appropriate-for-gestational age counterparts. Similarly, problems related to growth are a matter of concern. Literature reports both extremes of growth abnormalities in SGA infants, that is, undernutrition as well as obesity and metabolic syndrome [11−16]. Typically, SGA infants display catch-up growth (maximally between 3 and 6 months) in postneonatal life, which is more pronounced in weight than in length [17]. Most are expected to approach their genetic growth trajectory by the age of 2 years and demonstrate normal growth afterward with appropriate care and nutrition, barring those with genetic or syndromic illnesses [18,19].

Impaired intrauterine growth in SGA infants and subsequent postnatal undernutrition are two manifestations of the same phenomenon. The factors that influence growth in the intrauterine period often influence growth in the postnatal

* Corresponding author. Department of Pediatrics, Lady Hardinge Medical College, Bangla Sahib Road, New Delhi 110001, India.

E-mail address: viswaschhapola@gmail.com (V. Chhapola).

https://doi.org/10.1016/j.jclinepi.2024.111611

What is new?

Key findings

- We identified key causal factors of childhood malnutrition in the Indian subcontinent through a systematic review.
- We applied the principles of causal theory to construct a directed acyclic graph (DAG) to depict the causal pathways leading to malnutrition in term small-for-gestational age infants.

What this adds to what is known?

- This study is the first to integrate causal theory recommendations to develop an evidence-based DAG for malnutrition in term small-for-gestational age infants.
- This DAG provides a transparent template with explicit assumptions that are open to scrutiny by other experts in malnutrition research.

What is the implication and what should change now?

- Using the estimands, a causal model, and minimum adjustment sets at their disposal, researchers can now design better observational studies to uncover causality in this population.
- The set of identified factors can be used to construct additional DAGs to investigate the relationship between different exposures and malnutrition, using their own set of assumptions.

period also because the family environment, social factors, and resources generally remain the same. For instance, the infant of a malnourished mother (owing to household food scarcity) may be deprived of transplacental nutrient supply during the antenatal period and may not receive optimal postnatal complementary feeding. Therefore, we hypothesized that SGA infants who are nurtured in an environment of care and adequate feeding are unlikely to be malnourished. Although the literature consistently associates SGA/low birth weight with childhood malnutrition, there is no clarity on whether this association is causal [12,20–23]. Health agencies such as the World Health Organization (WHO) and United Nations International Children's Emergency Fund (UNICEF) list malnutrition as one of the consequences of SGA, but causality is implied rather than proven [24,25]. Nevertheless, it is crucial to determine causality, because it guides the choice of intervention. If SGA causes childhood malnutrition, these infants would require customized postnatal monitoring and feeding plans. By contrast, if SGA is not causal, it would ignite a search for other sociodemographic causal factors of malnutrition and targeted interventions to modify them. As it is common practice to use the terms undernutrition and malnutrition interchangeably, we have referred to undernutrition as malnutrition in the remainder of the manuscript.

The literature is replete with observational studies that list factors related to childhood malnutrition in low- and middle-income countries [9]. However, these studies have many limitations that render causal conclusions untenable. First, no attention has been paid to differentiating whether the relationships among the factors are associational or causal [26]. Furthermore, the aims of modeling are unclear, that is, to predict or understand causality [27]. Without these conceptual distinctions, the obtained effect estimates are neither causal nor reliable. Another common error is making regression-based adjustments for all available covariates without distinguishing among confounders, mediators, or colliders, which can introduce bias [28–30]. Finally, the list of putative factors is rarely comprehensive, and the factors for regression analysis are generally chosen according to available data [31]. Therefore, observational studies are rarely able to provide true estimates of the causal effects of exposure on outcomes. Directed acyclic graphs (DAGs) can address this problem by applying causal theory to observational data to provide a model for estimating unbiased effect sizes [32–35]. Despite the growing use of DAGs in epidemiological research, their application in studies of childhood malnutrition remains limited.

This study aimed to construct a causal model to estimate the effect of term-SGA on malnutrition (diagnosed according to WHO criteria) in children aged less than 5 years belonging to the Indian subcontinent [36]. We hypothesized that term-SGA infants who do not have genetic or endocrine abnormalities and receive good care and nutrition would not develop malnutrition.

2. Materials and methods

We followed the "Evidence Synthesis for Constructing Directed Acyclic Graphs" approach proposed by Ferguson et al [33]. A systematic review (Agency for Healthcare Research and Quality Evidence-based Practice Centers [AHRQ-EPCs] "rapid narrative review" methodology) was initially conducted to identify the causal factors of malnutrition, and then a causal diagram (DAG) was constructed using knowledge from the published literature and expert opinion [37]. We used the *"Estimand → Estimator → Estimate"* framework for study flow [32]. We describe the estimands and the estimator in the following section. The derivation of estimate was beyond the scope of this study. The review approach was narrowed down using the web-based ''Right review knowledge translation tool'' (https://rightreview.knowledgetranslation.net/) [38]. At present, there are no EQUATOR (https://www.equator-network.org/) network reporting guidelines for rapid narrative review. Therefore, we have reported this research following the WHO guidance document on rapid reviews (Supplementary Table 1) [39].

2.1. Estimands

We formulated three causal estimands aligned with our study objectives, targeting the perspectives of parents, clinicians/nutrition counselors, and public health specialists/policymakers [40,41] (Table 1).

2.2. Estimator

Causal (DAG) and statistical (regression models for covariate conditioning) models constitute the estimators. For this purpose, a two-step procedure was adopted. First, we identified the causal factors of malnutrition through a systematic review, and then constructed a causal model (Evidence Synthesis for Constructing Directed Acyclic Graphs approach) [33].

2.2.1. Step 1: systematic review

Study eligibility criteria: We included original research (observational studies or randomized controlled trials [RCTs]) that examined the factors associated with malnutrition in children aged 6 months to 5 years, belonging to the Indian subcontinent (India, Pakistan, Nepal, Bangladesh, Bhutan, Sri Lanka, and Maldives). We included studies on human subjects published in English over the last 10 years (January 2014 to December 2023).

Literature search: We searched 4 databases (PubMed, Scopus, Web of Science, and ProQuest) for all publications written in English from January 1, 2014 to December 15, 2023, using search strategies customized to each database. A literature search was conducted on December 15, 2023. The details of the search strategy are provided in Supplementary Table 2. Search terms were entered into the databases, and references and abstracts were imported into the Rayyan QCRI software [42]. Duplicate studies were excluded from the analysis. Two reviewers (S. T. and N. C.) independently screened titles and abstracts based on the eligibility criteria. Any unclear decisions were discussed with a third reviewer (V. C.) and studies that did not meet the inclusion criteria were excluded.

2.2.1.1Data extraction. During the initial scoping of the literature, we noticed that the evidence base on our research objective was very large and there were too many articles listing risk factors. Therefore, we adopted a pragmatic approach, whereby only the abstracts were reviewed to identify risk factors significantly associated with malnutrition. We did not pool the results of the studies or assess their quality. With this strategy, we hoped for a comprehensive inclusion of published literature to achieve saturation of the most frequently reported causal factors. Two authors (S. T. and N. C.) independently completed the data extraction in duplicate. Any disagreements were resolved by consensus and discussion with another reviewer (V. C.). A summary table was created for the study characteristics from the included abstracts. Summary details included the year of publication, country, research design, sample size, anthropometric indices, and factors associated with malnutrition (Supplementary Table 3).

2.2.2. Step 2: causal modeling using directed acyclic graphs

DAGs are diagrammatic representations of the assumed causal pathways leading from a defined exposure to the outcome. DAGs are composed of nodes interconnected by unidirectional edges (arrows) that do not loop back to the same node; hence, they are termed as acyclic [32–35]. The nodes represent variables, such as exposure, outcome, and confounders. The edges depict the direction of the hypothesized relationships between the variables (nodes). The edges do not indicate whether the relationship is positive/negative, linear/nonlinear, or small/large; therefore, DAGs are nonparametric [32]. A causal path is a collection of 1 or more connected nodes, with edges flowing in the same direction. A confounder is a variable that influences both the exposure and outcome. The mediator variable falls on the causal path from exposure to the outcome. A competing exposure is the ancestor of the outcome but is unrelated to primary exposure [32]. A collider is a variable caused by at least two other variables. A direct causal effect is the effect of exposure on the outcome through paths that do not involve mediators, whereas a total causal effect is the combined effect transmitted through all causal paths. To estimate the true effect, we must close all confounding (backdoor) paths from the exposure to the outcome, which is called deconfounding [35]. Deconfounding is achieved by conditioning using regression modeling. In contrast, conditioning on the collider is not recommended because it opens a backdoor path [43]. DAG provides multiple ''minimal sufficient adjustment sets'' of variables for an unbiased estimate of the causal effect of a specific exposure on a specified outcome [44,45].

Ferguson et al recommend the construction of DAG in three sequential steps: *Mapping* → *Translation* → *Integration*.

Mapping of the theory from each study produces implied graphs (IGs). Each relationship in the individual IGs is then assessed under sequential causal criteria and a

Table 1. Estimands for research objectives

Design feature	Scenario 1	Scenario 2	Scenario 3
Stakeholders	Public health specialists/ policymakers as stakeholders	Parents	Clinicians/nutrition counselors
Expectations of stakeholders	Public health specialists/ policymakers will be interested in determining the proportion of SGA infants who have malnutrition in the under-5 age group. This information could be used to develop policies for better antenatal care and targeted postnatal intervention.	Parents will be concerned about both the risk of mortality in the neonatal period and the risk of malnutrition later.	The clinician/nutrition counselors will be interested in knowing whether dietary intervention can prevent malnutrition in SGA babies—a purist viewpoint.
Population of interest	Term newborn infants from Indian subcontinent countries	Same as Scenario 1	Same as Scenario 1
Primary exposure	**Inclusion:** **Cases:** SGA infants, defined as a weight below the 10th percentile for the gestational age. **Control group:** Neonates with appropriate weight for age. **Exclusion:** Term infants with genetic and syndromic illnesses.	Same as Scenario 1	Same as Scenario 1
Variable (end-point) at the patient level	Children who develop acute or chronic malnutrition between 6 mo and 5 y of age, as diagnosed by the WHO criteria.	Children who die (any time after birth until the age of 1 mo) or develop acute or chronic malnutrition between 6 mo and 5 y of age, as diagnosed by the WHO criteria.	Same as Scenario 1
Population-level summary measure	Odds ratio	Same as Scenario 1	Same as Scenario 1
Intercurrent events	Death at any time up to 5 y of age	Death in the neonatal period.	Death in the neonatal period.
Handling of intercurrent event	**Treatment policy strategy:** All included infants will be analyzed for the outcome (malnutrition) regardless of the intercurrent event.	**Composite strategy:** Death in the neonatal period will be added to malnutrition as a composite outcome. The likelihood of death is higher in SGA infants during the neonatal period. We assume that the number of deaths from causes other than malnutrition would be similar in the postneonatal age group in both groups.	**Principal stratum strategy:** Both cases and controls will be included in the analysis starting from 1 mo of age. By excluding neonates who died before 1 mo of postnatal age, we can isolate the effect of SGA on malnutrition. We assume that the number of deaths from causes other than malnutrition would be similar in both groups after this exclusion.

SGA, small-for-gestational age; WHO, World Health Organization.

counterfactual thought experiment, which gives rise to translational DAGs. All the translational DAGs are then combined to form one integrated DAG. Our approach deviated from this method. As we expected a large number of eligible studies, we adopted a pragmatic approach, whereby we first constructed a single IG by considering all the factors identified in our review. We designated "malnutrition" as the outcome and "SGA" as the exposure variable of interest. Two directed edges were drawn from the exposure, passing through two mediators ('poor nutritional intake' and 'persistent catabolic state') and terminating at the outcome [22]. The remaining causal factors were designated as the unassigned covariates. Directed edges were drawn from each covariate to the exposure and outcome (assuming that all were confounders). IG was saturated by drawing undirected edges between all confounders.

The hypothesized relationship between nodes was assessed in both the forward and reverse directions using Bradford Hill's criteria of temporality, face validity, and recourse to theory [46]. Finally, a counterfactual thought experiment was applied to each relationship [32,33,35].

These criteria were applied sequentially. The criterion of "temporality" requires that the postulated cause always precedes the effect (if 'yes,' proceed to the next criterion; If not, assess the reverse relationship). The next step was to assess "face validity" and whether the postulated relationship was plausible (if 'yes,' proceed to the next criterion; if not, assess the reverse relationship). ‘‘Recourse to theory’’ criterion stipulates that the postulated relationship must be supported by theory (If yes, always proceed to the counterfactual thought experiment) [46].

The final step was the ‘‘counterfactual thought experiment’’. This is a systematic thought experiment in which potential outcomes are compared in a scenario in which the entire study sample is exposed vs one in which there is no exposure. The counterfactual thought experiment entails retrospective reasoning, in which outcomes in the real world are compared with outcomes in an alternative world where causal factors are absent [32,33,35].

We prepared a final DAG by removing edges where a potential relationship could not be conceptualized. The entire process of constructing the DAG was undertaken by two authors (V. C. and S. T.) sitting together, and going by consensus. Any disagreements were resolved in a meeting involving all 4 authors. DAGitty version 3.1 was used for the construction of DAG [44].

3. Results

3.1. Results of the literature search

The literature search yielded 4818 records, of which 405 were screened for detailed assessment. Ultimately, 342 abstracts were used for the data extraction (Fig. 1).

3.2. Study characteristics

The sample sizes of the included studies varied widely from 47 to 564,518 participants with a median (interquartile range) of 2035 (7557). The highest number of studies were from India (39%), followed by Bangladesh (27%). The most common study designs were cross-sectional (76%) and cluster surveys (16%). Both urban and rural populations were included in 82% of studies. The anthropometric outcome measures were length/height-for-age (72%), weight-for-age (54%), and weight-for-height/length (52%). Mid-upper arm circumference and body mass index were used as nutritional indicators in less than 5% of the studies (Supplementary Table 4).

3.3. Factor identification and handling

The literature review identified 81 factors, with the most commonly reported being maternal education, poverty, maternal malnutrition, socioeconomic status, and low birth weight (Supplementary Table 5). Upon scrutiny, we labeled the 4 factors as purely predictive and excluded them from the causal model. The remaining 77 factors were grouped into 17 single or composite nodes, referring to five domains: socioeconomic, parental, child-related, environmental, and political (Table 2).

Composite nodes (supernodes) comprised of factors that had similar causal mechanisms. The availability of home,

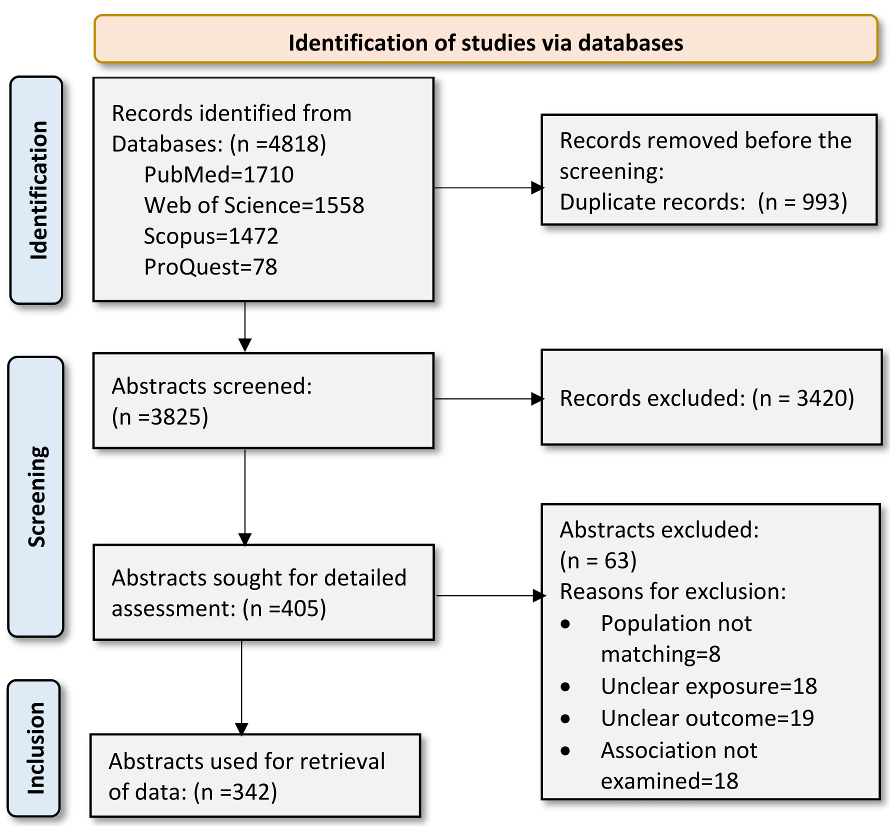

Figure 1. Flow chart showing the result of literature search and selection of abstracts for qualitative synthesis.

Table 2. Causal factors, causal domains, and potential causal mechanisms of malnutrition

S. No.	Causal factors	Causal domain	Number of studies reporting the factor	Potential causal mechanisms
1.	Poverty	Socioeconomic	211	Less availability of food and housing; poor sanitation; maternal malnutrition
2.	Mother's education	Parental	155	A less educated mother is likely to be unemployed and have high-risk fertility behavior; likely to be less aware of care during pregnancy and child-rearing practices; might lack autonomy in family decisions.
3.	High-risk fertility behavior	Parental	153	Very young or elderly mothers or those with less birth spacing can be malnourished, and all these can lead to low birth weight, lactation failure, and poor infant feeding.
4.	Maternal malnutrition	Parental	89	This can lead to pregnancy complications, low birth weight, and lactation failure.
5.	Household security and hygiene	Socioeconomic	88	Lack of housing, safe water, and sanitation will lead to social insecurity, less availability of food, and recurrent infection; similar to poverty.
6.	Awareness of mothers regarding antenatal and child care	Parental	45	Lack of awareness might lead to high-risk fertility behavior and malnutrition in mother; low birth weight infant, suboptimal feeding and immunization.
7.	Child neglect	Child-related	37	Apathy toward child can lead to poor feeding and immunization, and recurrent infection.
8.	Recurrent infections	Child-related	33	Catabolic state, poor appetite affecting nutrient intake.
9.	Immunization of child	Child-related	20	Unimmunized children can have serious/life-threatening and/or recurrent infections.
10.	Mother's employment status	Parental	16	Maternal unemployment can contribute to poverty and lack of mother's autonomy.
11.	Access to healthcare	Political	15	Lack of health facilities in the vicinity affects immunization rates and awareness of mothers regarding antenatal and child care.
12.	Domestic violence or lack of mother's autonomy	Parental	12	This can affect nutritional status of pregnant lady and her children and immunization of children. Can also lead to neglect of children in the household.
13.	Bottle feeding	Child-related	8	Delays in complementary feeding can lead to inadequate nutrition; unhygienic bottle-feeding practices can cause recurrent infections.
14.	Migration, residence in disaster-prone areas, and marginalized communities	Environmental	7	Acts through household insecurity, decreased food availability, maternal malnutrition, and poor immunization.
15.	Psychiatric illness/addiction in the family	Parental	5	It can affect nutrition of mother and child, employment of mother; can cause poverty and poor awareness regarding maternal and child care.
16.	Air pollution	Environmental	3	Low birth weight through uteroplacental insufficiency; chronic or recurrent respiratory morbidities.
17.	Genetic disorder or chronic systemic illness	Child-related	2	Through chronic catabolic state or poor feeding.

safe water, and sanitation measures was labeled as 'household security and hygiene.' Inadequate nutrition and recurrent infections are likely to occur if household security and hygiene are compromised. Marginalized or migrant communities and people living in disaster-prone areas were grouped as a single node, as all of them were vulnerable because of poor access to household security and hygiene, opportunities for education, stable occupation, immunization, and access to healthcare. Domestic violence and a lack of mothers' autonomy were combined, as they could cause emotional deprivation in the child and affect their nutritional intake and healthcare needs. In this context, the 'autonomy' implies "freedom of decision-making in matters related to herself and her child" [47]. Psychiatric illness and addiction in the family were clubbed together, with the assumption that any such family circumstance could cause neglect of the child, resulting in malnutrition. Similarly, genetic disorders and chronic systemic illnesses were combined into a single node, as both could lead to malnutrition via a persistent catabolic state and poor nutrient intake. Table 2 lists the 17 factors, causal domains, and causal reasoning.

3.4. DAG and minimal sufficient adjustment sets

The final DAG constructed using Bradford Hill's criteria and a counterfactual thought experiment is shown in Figure 2. A detailed account of the decision for each possible edge is provided as a decision log in Supplementary Table 6. A summary of the retained edges is presented as a directed edge index (Supplementary Table 7). The DAG consisted of 21 nodes, including primary exposure, outcome, mediators, and covariates. There were 59 edges interconnecting the nodes. The DAG identified 12 minimal sufficient adjustment sets to estimate the total effect of SGA on malnutrition. Minimal sufficient adjustment set-1 consisted of a total of 9 variables of 17 candidate variables (access to healthcare, air pollution, antenatal/childcare awareness, domestic violence/lack of mother's autonomy, genetic disorder/chronic systemic illness, maternal malnutrition, migration/residence in disaster-prone area/marginalized communities, poverty, psychiatric illness/addiction). The other adjustment sets are listed in Supplementary Table 8. No model has been proposed for estimating the direct effect of SGA on malnutrition, because the relationship between these 2 is only indirect through mediators. The DAGitty code for the DAG is provided in Supplementary Table 9.

4. Discussion

Here, we present a case study aimed at addressing an important gap in the literature regarding the causal effect of SGA on malnutrition. Although we hypothesized that SGA does not cause malnutrition, the available literature cannot answer this question. Whether our hypothesis is true can only be answered by a well-designed observational study in which all possible confounders have been conditioned upon, all colliders identified, and there is no

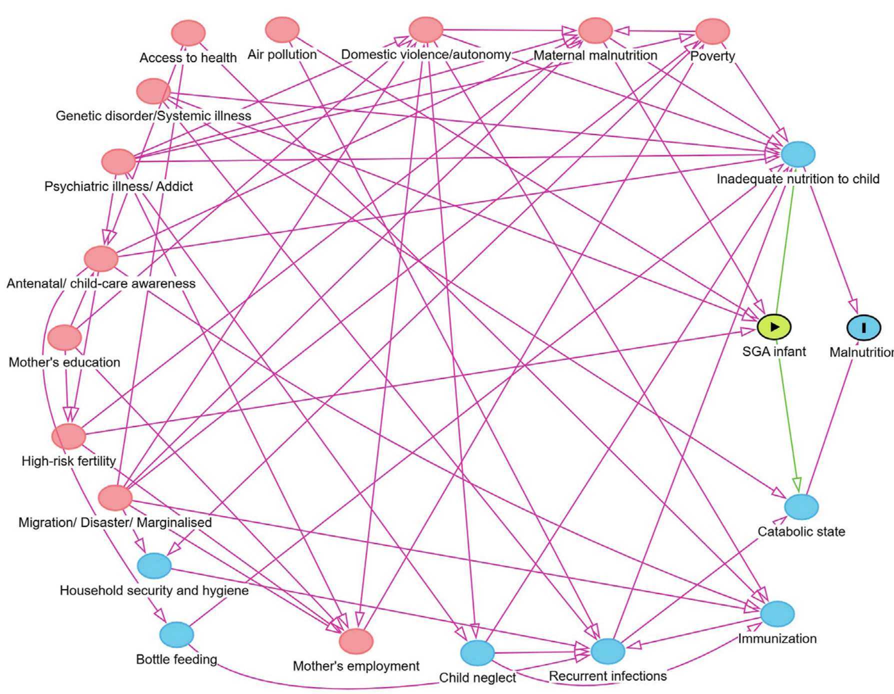

Figure 2. Directed acyclic graph (DAG) depicting the causal relationship between term-small for gestational age status (exposure) and malnutrition at 6 months to 5 years of age (outcome), incorporating causal pathways and covariates (mediators, confounders, and colliders).

measurement or selection bias [29]. Even large sample size datasets from cluster surveys on childhood nutrition cannot mitigate collider and confounding bias, because the list of covariates is often incomplete and selected without causal reasoning [31].

We performed a robust systematic review of the literature to identify the causal factors unique to the Indian subcontinent. This ensured comprehensiveness in our adjustment set and minimized unmeasured confounding. We acknowledge that an observational design can never perfectly address unmeasured confounding, which is an advantage offered by the randomization procedure in an RCT [48]. Moreover, even with an optimal collection of causal factors, residual confounding cannot be eliminated because of the imperfect measurement of latent traits in the variables [31]. Despite the aforementioned limitations, observational studies are the only recourse in this scenario because it is not ethical or feasible to conduct an RCT [49].

Most previous studies on this causal query have not considered a formal causal model to inform data generation. To our knowledge, this is the first study to integrate causal theory recommendations to develop an evidence-based DAG for malnutrition among SGA infants. In a previous systematic review, Mertens et al synthesized data from longitudinal studies and identified the causes of growth faltering in children from low- and middle-income countries [9]. The authors constructed the DAGs for each exposure. However, there are several noticeable weaknesses in DAGs that cannot be compensated by the robust statistical techniques used in that review. First, there were errors in selecting mediators and other covariates. Second, there is a conflation between causal and predictive goals, although both require different methodological approaches. Ramspek et al highlighted that conflation is a significant problem in observational studies and results in biased estimates and erroneous conclusions [50]. In another systematic review, Christian et al estimated study-specific and pooled risk estimates of stunting, wasting, and underweight due to SGA and preterm birth. They found that the odds of stunting, wasting, and underweight were 2−3 times higher in term-SGA infants than in term appropriate-for-gestational age infants [12]. However, the trustworthiness of this effect size is debatable, because the adjustment set was incomplete and not based on a causal model.

We provide a simplified DAG for malnutrition in SGA infants. A notable strength is that we formulated estimands that catered to the interests of different stakeholders. Explicit estimands that account for intercurrent events add to the validity of estimates drawn from this causal model. Interestingly, the causal model halved the number of covariates required for deconfounding, thereby allowing for parsimony in variable selection. Moreover, the availability of 12 minimal sufficient adjustment sets ensures flexibility in selection of confounders. We did not review the full texts of the articles because of the large number of eligible manuscripts. Instead, we adopted a pragmatic approach to extract the required information from abstracts. Unlike a typical systematic review process, we adopted a rapid review methodology in which research quality appraisals were not performed [37]. Nonetheless, we believe that in the context of our objectives, this method provides credible information from the end-user perspective [51]. Additionally, we assume that factors that were consistently (consistency assumption of Bradford Hill criteria) reported across studies provide a reasonable idea of important causal factors [46].

Translating a causal model into data generation requires the operationalization of causal variables [52]. Most of the factors depicted in this DAG can be quantified using scales (eg, health-seeking behavior, child neglect, domestic violence, mothers' autonomy, psychiatric illness, and socioeconomic status), health cards (eg, immunization and antenatal/postnatal care), or targeted interviews (eg, history of infections, drug abuse, dietary intake, breastfeeding, and complementary feeding).

Although we aimed to understand the causal effect of SGA on malnutrition, the same set of factors can be used to construct more DAGs to examine the focal relationship between different exposures and malnutrition, with their own set of assumptions. Our DAG provides a transparent template with explicit assumptions that are open to scrutiny by other experts in malnutrition research. We did not include the structural approach of time-varying confounding in this DAG, although many covariates can have time-varying properties, such as breastfeeding, complementary feeding, poverty status, parental occupation, and availability of housing and sanitation [53]. We refrained from including the time-varying structure, as it would have made the DAG extremely complex. This qualitative evidence synthesis is not backed by quantitative analyses as that was beyond the scope of this study.

5. Conclusion

We offer an evidence-based DAG that will minimize bias due to improper selection of causal factors in studies focusing on malnutrition in term-SGA infants. Understanding the causal factors of malnutrition is important for both clinicians and policymakers. The availability of different estimands, causal models, and minimum adjustment sets can help researchers design future research and analyze their data. In addition to investigating causality, DAG can be applied to develop predictive models. Our study bridges an important gap in the malnutrition literature.

Ethics statement

This study did not involve human participants. Therefore, permission was not required from the ethics committee.

CRediT authorship contribution statement

Soumya Tiwari: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Viswas Chhapola:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Conceptualization. **Nisha Chaudhary:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Lokesh Sharma:** Writing – review & editing, Methodology, Formal analysis.

Declaration of competing interest

The authors declare no conflict of interest related to this work.

Supplementary Data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2024.111611.

Data availability

Data is available as supplementary material

References

[1] WHO Expert Committee on Physical Status: the Use and Interpretation of Anthropometry (1993 : Geneva, Switzerland) & World Health Organization. Physical status: the use of and interpretation of anthropometry, report of a WHO expert committee. World health organization; 1995. Available at: https://iris.who.int/handle/10665/37003. Accessed November 8, 2024.

[2] He H, Miao H, Liang Z, Zhang Y, Jiang W, Deng Z, et al. Prevalence of small for gestational age infants in 21 cities in China, 2014–2019. Sci Rep 2021;11:7500. https://doi.org/10.1038/s41598-021-87127-9.

[3] Hocquette A, Durox M, Wood R, Klungsøyr K, Szamotulska K, Berrut S, et al. International versus national growth charts for identifying small and large-for-gestational age newborns: a population-based study in 15 European countries. Lancet Reg Health Eur 2021;8:100167. https://doi.org/10.1016/j.lanepe.2021.100167.

[4] Lee ACC, Katz J, Blencowe H, Cousens S, Kozuki N, Vogel JP, et al. National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. Lancet Glob Health 2013;1:e26–36. https://doi.org/10.1016/S2214-109X(13)70006-8.

[5] Cho WK, Jung IA, Suh BK. Current growth status and metabolic parameters of Korean adolescents born small for gestational age: results from the Korea National Health and Nutrition Examination Surveys (KNHANES) 2010-2011. Pediatr Int 2014;56:344–8. https://doi.org/10.1111/ped.12278.

[6] Tudehope D, Vento M, Bhutta Z, Pachi P. Nutritional requirements and feeding recommendations for small for gestational age infants. J Pediatr 2013;162:S81–9. https://doi.org/10.1016/j.jpeds.2012.11.057.

[7] Pollack RN, Divon MY. Intrauterine growth retardation: definition, classification, and etiology. Clin Obstet Gynecol 1992;35:99–107. https://doi.org/10.1097/00003081-199203000-00015.

[8] Falcão IR, Ribeiro-Silva RC, de Almeida MF, Fiaccone RL, Silva NJ, Paixao ES, et al. Factors associated with small- and large-for-gestational-age in socioeconomically vulnerable individuals in the 100 Million Brazilian Cohort. Am J Clin Nutr 2021;114:109–16. https://doi.org/10.1093/ajcn/nqab033.

[9] Mertens A, Benjamin-Chung J, Colford JM, Coyle J, van der Laan MJ, Hubbard AE, et al. Causes and consequences of child growth faltering in low-resource settings. Nature 2023;621:568–76. https://doi.org/10.1038/s41586-023-06501-x.

[10] Gluckman PD, Harding JE. The physiology and pathophysiology of intrauterine growth retardation. Horm Res 1997;48(Suppl 1):11–6. https://doi.org/10.1159/000191257.

[11] Ahmad N, Sutan R, Tamil AM, Hajib N. Growth patterns and nutritional status of small for gestational age infants in Malaysia during the first year of life. Child Health Nurs Res 2021;27:317–27. https://doi.org/10.4094/chnr.2021.27.4.317.

[12] Christian P, Lee SE, Donahue Angel M, Adair LS, Arifeen SE, Ashorn P, et al. Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries. Int J Epidemiol 2013;42:1340–55. https://doi.org/10.1093/ije/dyt109.

[13] Leunissen RWJ, Kerkhof GF, Stijnen T, Hokken-Koelega A. Timing and tempo of first-year rapid growth in relation to cardiovascular and metabolic risk profile in early adulthood. JAMA 2009;301:2234–42. https://doi.org/10.1001/jama.2009.761.

[14] Leunissen RWJ, Kerkhof GF, Stijnen T, Hokken-Koelega ACS. Effect of birth size and catch-up growth on adult blood pressure and carotid intima-media thickness. Horm Res Paediatr 2012;77:394–401. https://doi.org/10.1159/000338791.

[15] Kerkhof GF, Duivenvoorden HJ, Leunissen RWJ, Hokken-Koelega ACS. Pathways leading to atherosclerosis: a structural equation modeling approach in young adults. Hypertens Dallas Tex 2011;57:255–60. https://doi.org/10.1161/HYPERTENSIONAHA.110.163600.

[16] Singhal A, Lucas A. Early origins of cardiovascular disease: is there a unifying hypothesis? Lancet Lond Engl 2004;363:1642–5. https://doi.org/10.1016/S0140-6736(04)16210-7.

[17] Huang L, Yang S, Yang F, Xiong F. A prospective study about physical growth of children from birth to 2 years old born full-term small-for-gestational-age. J Paediatr Child Health 2019;55:199–204. https://doi.org/10.1111/jpc.14162.

[18] Karlberg JP, Albertsson-Wikland K, Kwan EY, Lam BC, Low LC. The timing of early postnatal catch-up growth in normal, full-term infants born short for gestational age. Horm Res 1997;48(Suppl 1):17–24. https://doi.org/10.1159/000191279.

[19] Karlberg J, Albertsson-Wikland K, Kwan CW, Chan FYS. Early spontaneous catch-up growth. J Pediatr Endocrinol Metab JPEM 2002;15(Suppl 5):1243–55.

[20] Rahman MS, Howlader T, Masud MS, Rahman ML. Association of low-birth weight with malnutrition in children under five years in Bangladesh: do mother's education, socio-economic status, and birth interval matter? PLoS One 2016;11:e0157814. https://doi.org/10.1371/journal.pone.0157814.

[21] Abbas F, Kumar R, Mahmood T, Somrongthong R. Impact of children born with low birth weight on stunting and wasting in Sindh province of Pakistan: a propensity score matching approach. Sci Rep 2021;11:19932. https://doi.org/10.1038/s41598-021-98924-7.

[22] Ntenda PAM. Association of low birth weight with undernutrition in preschool-aged children in Malawi. Nutr J 2019;18:51. https://doi.org/10.1186/s12937-019-0477-8.

[23] Namiiro FB, Batte A, Rujumba J, Nabukeera-Barungi N, Kayom VO, Munabi IG, et al. Nutritional status of young children born with low birthweight in a low resource setting: an observational study. BMC Pediatr 2023;23:520. https://doi.org/10.1186/s12887-023-04356-9.

[24] Low birthweight. UNICEF DATA n.d. Available at: https://data.unicef.org/topic/nutrition/low-birthweight/. Accessed August 22, 2024.

[25] Low birth weight n.d. Available at: https://www.who.int/data/nutrition/nlis/info/low-birth-weight. Accessed August 22, 2024.

[26] Altman N, Krzywinski M. Association, correlation and causation. Nat Methods 2015;12:899−900. https://doi.org/10.1038/nmeth.3587.

[27] Shmueli G. To explain or to predict? Stat Sci 2010;25:289−310.

[28] Spector PE, Brannick MT. Methodological urban legends: the misuse of statistical control variables. Organ Res Methods 2011;14:287−305. https://doi.org/10.1177/1094428110369842.

[29] Hernan MA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol 2002;155:176−84. https://doi.org/10.1093/aje/155.2.176.

[30] Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. Am J Public Health 2018;108:616−9. https://doi.org/10.2105/AJPH.2018.304337.

[31] Danziger J, Zimolzak AJ. Residual confounding lurking in big data: a source of error. Secondary analysis of electronic health records. Cham: Springer International Publishing; 2016:71−8. https://doi.org/10.1007/978-3-319-43742-2_8.

[32] Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. Int J Epidemiol 2021;50:620−32. https://doi.org/10.1093/ije/dyaa213.

[33] Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. Int J Epidemiol 2020;49:322−9. https://doi.org/10.1093/ije/dyz150.

[34] Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999;10:37−48. https://doi.org/10.1097/00001648-199901000-00008.

[35] Pearl J, Mackenzie D. The book of why: the new science of cause and effect. New York: Basic books; 2018.

[36] Organization WH. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva: World Health Organization; 2006.

[37] Hartling L, Guise J-M, Hempel S, Featherstone R, Mitchell MD, Motu'apuaka ML, et al. EPC methods: AHRQ End-user perspectives of rapid reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2016.

[38] Amog K, Pham B, Courvoisier M, Mak M, Booth A, Godfrey C, et al. The web-based "Right Review" tool asks reviewers simple questions to suggest methods from 41 knowledge synthesis methods. J Clin Epidemiol 2022;147:42−51. https://doi.org/10.1016/j.jclinepi.2022.03.004.

[39] Tricco AC, Langlois EV, Straus SE, Research A for HP and S, Organization WH. Rapid reviews to strengthen health policy and systems: a practical guide. Geneva: World Health Organization; 2017.

[40] Chen J, Scharfstein D, Wang H, Yu B, Song Y, He W, et al. Estimands in real-world evidence studies. Stat Biopharm Res 2024;16:257−69. https://doi.org/10.1080/19466315.2023.2259829.

[41] Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICH E9(R1) addendum. BMJ 2024;384:e076316. https://doi.org/10.1136/bmj-2023-076316.

[42] Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016;5:210. https://doi.org/10.1186/s13643-016-0384-4.

[43] Suzuki E, Shinozaki T, Yamamoto E. Causal diagrams: pitfalls and tips. J Epidemiol 2020;30:153−62. https://doi.org/10.2188/jea.JE20190192.

[44] Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package "dagitty". Int J Epidemiol 2016;45:1887−94. https://doi.org/10.1093/ije/dyw341.

[45] Textor J. Drawing and Analyzing Causal DAGs with DAGitty. 2015. Available at: https://doi.org/10.48550/arXiv.1508.04633. Accessed November 28, 2024.

[46] Hill AB. The environment and disease: association or causation? Proc R Soc Med 1965;58:295−300.

[47] Paul P, Saha R. Is maternal autonomy associated with child nutritional status? Evidence from a cross-sectional study in India. PLoS One 2022;17:e0268126. https://doi.org/10.1371/journal.pone.0268126.

[48] Jager KJ, Zoccali C, Macleod A, Dekker FW. Confounding: what it is and how to deal with it. Kidney Int 2008;73:256−60. https://doi.org/10.1038/sj.ki.5002650.

[49] Bosdriesz JR, Stel VS, Van Diepen M, Meuleman Y, Dekker FW, Zoccali C, et al. Evidence-based medicine—when observational studies are better than randomized controlled trials. Nephrology 2020;25:737−43. https://doi.org/10.1111/nep.13742.

[50] Ramspek CL, Steyerberg EW, Riley RD, Rosendaal FR, Dekkers OM, Dekker FW, et al. Prediction or causality? A scoping review of their conflation within current observational research. Eur J Epidemiol 2021;36:889−98. https://doi.org/10.1007/s10654-021-00794-w.

[51] Hartling L, Guise JM, Hempel S, Featherstone R, Mitchell MD, Motu'apuaka ML, et al. Fit for purpose: perspectives on rapid reviews from end-user interviews. Syst Rev 2017;6:32. https://doi.org/10.1186/s13643-017-0425-7.

[52] Andrade C. A student's guide to the classification and operationalization of variables in the conceptualization and design of a clinical study: Part 2. Indian J Psychol Med 2021;43:265−8. https://doi.org/10.1177/0253717621996151.

[53] Fewell Z, Hernán MA, Wolfe F, Tilling K, Choi H, Sterne JAC. Controlling for time-dependent confounding using marginal structural models. Stata J Promot Commun Stat Stata 2004;4:402−20. https://doi.org/10.1177/1536867X0400400403.