Original Paper

A Social Media Mining and Questionnaire Survey Hybrid Approach to Characterizing Caregiver Knowledge, Attitudes, and Practices on House Dust Mite Control: Cross-Sectional Study

Peixin Yang[1], MMeD; Ruoling Xian[2], MMeD; Juan Li[1,3], MD, PhD; Jing Zhao[1], MMeD; Hangjiang Xie[1], MMeD; Zhenghong Jiang[1,3], MSc; Yan Hu[4], MD, PhD; Ximing Xu[5], PhD; Na Zang[3], MD, PhD; Yan Zhao[2*], BD; Heping Fang[1,3*], MD, PhD; Enmei Liu[1*], MD, PhD

[1]Department of Respiratory Medicine, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Children and Adolescents' Health and Diseases, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Municipal Health Commission Key Laboratory of Children's Vital Organ Development and Diseases, Chongqing, China

[2]Department of Pediatrics, Chongqing Health Center for Women and Children, Women and Children's Hospital of Chongqing Medical University, Chongqing, China

[3]Pediatric Research Institute, Children's Hospital of Chongqing Medical University, Chongqing, China

[4]Children's Health Care Center, Beijing Children's Hospital, Capital Medical University, Beijing, China

[5]Big Data Engineering Center, Children's Hospital of Chongqing Medical University, Chongqing, China

*these authors contributed equally

Corresponding Author:

Heping Fang, MD, PhD
Department of Respiratory Medicine, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Children and Adolescents' Health and Diseases
Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Municipal Health Commission Key Laboratory of Children's Vital Organ Development and Diseases
No. 20 Jinyu Avenue, Liangjiang New Area
Chongqing
China
Phone: 86 02348370223
Email: heping.f@outlook.com

Abstract

Background: House dust mite (HDM) sensitization commonly begins in early life and contributes to persistent allergic airway inflammation and asthma chronicity. Primary prevention via early-life environmental control is a potential pathway to reduce HDM sensitization and asthma risk.

Objective: This study aims to characterize child caregivers' knowledge, attitudes, and practices (KAP) regarding pediatric HDM control using a hybrid literature/expert-driven and social media–driven approach and to examine associations between KAP levels, child age, and caregiver HDM-related online engagement.

Methods: This cross-sectional study comprised two interconnected components: (1) mining of content published between August 2023 and July 2025 from 5 major Chinese social media platforms, analyzed via latent Dirichlet allocation (LDA), and (2) a social media–enhanced web-based KAP survey administered in November 2025 to child caregivers in Chongqing, China. In total, 132,341 social media documents and 2275 caregivers of children younger than 18 years were included in the analysis. The main outcomes included social media discourse patterns and domain-specific KAP levels across 5 dimensions: foundational knowledge (K1), recommended control knowledge (K2), attitude toward social media topics (A1), attitude toward recommended methods (A2), and control practices (P). Stratified analysis was conducted by 2 exposure variables: child age (≤3 years vs >3 years) and caregiver HDM-related online engagement (engaged vs nonengaged).

Results: LDA topic modeling identified 5 distinct topic clusters in the social media content. Commercial, emotional, and misleading content collectively dominated the information landscape, accounting for 83.3% (n=110,254) of the included documents (65.8% of the tokens), with commercial content often systematically conflating the concepts of "disinfection" and

"mite elimination." Only 16.7% (n=22,087) of the documents (34.1% of the tokens) were classified as health educational content focusing on HDM allergy prevention, reflecting that this valid educational content is substantially longer and more information-dense than other social media posts. The average KAP levels of K1, K2, A1, A2, and P domains were 62.9%, 84.7%, 57.0%, 37.8%, and 25.8%, respectively. Social media emerged as the primary knowledge source (80.7%, n=1836), with methodological knowledge gaps (47.5%) being the top implementation barrier. Caregivers of children ≤3 years had significantly lower self-rated knowledge (23.5% vs 28.3%; *P*=.01), stronger endorsement of recommended methods, but also greater information overload (odds ratio [OR] 1.39, 95% CI 1.15-1.67; *P*<.001) and efficacy uncertainty (OR 1.23, 95% CI 1.01-1.52; *P*=.04). Caregivers with HDM-related online engagement showed better performance across multiple items in 5 domains, but also increased traditional practice with uncertain evidence (ultraviolet irradiation: OR 1.85, 95% CI 1.35-2.53; *P*<.001) and misconception acceptance (allergy impact exaggeration: OR 1.39, 95% CI 1.04-1.87; *P*=.03).

Conclusions: Complex and suboptimal KAP levels exist, particularly among caregivers of young children (≤3 years). HDM-related online engagement is associated with both enhanced implementation of control practices and elevated misconception endorsement. These findings reveal critical educational gaps and the necessity of social media intervention.

J Med Internet Res 2026;28:e100446; doi: 10.2196/100446

Keywords: allergy; allergen avoidance; house dust mite; HDM; knowledge, attitudes, and practices; KAP; survey; social media; caregiver; children; health education; methodology

Introduction

Allergic diseases affect approximately 10% to 30% of the global population, peaking in childhood and adolescence [1]. House dust mites (HDMs) rank among the most prevalent indoor allergens. Immunoglobulin E (IgE) responses follow distinct age-specific patterns, with HDM sensitization typically emerging around the age of 3 years and persisting into adulthood [2-4]. Moreover, the persistent allergic airway inflammation triggered by early sensitization underpins the chronic trajectory of childhood asthma [5], suggesting a critical window for early intervention. Targeted interventions can effectively reduce HDM exposure in controlled settings [6-8] and are supported for primary prevention settings [9,10]. However, despite established recommendations for HDM control [11], real-world implementation remains inadequate among sensitized individuals and scarce in the general population [12-15]. Therefore, evaluating HDM control practices in households with children, particularly during the crucial early-life sensitization window, is essential.

Knowledge, attitudes, and practices (KAP) surveys provide a systematic approach to assess caregiver behaviors. However, current KAP methodology often relies on literature reviews and expert consultations, framing content primarily from a researcher-centered perspective [16]. This "top-down" approach risks information asymmetry, potentially overlooking participants' genuine concerns, especially in self-administered surveys [17-19]. Misalignment between researcher priorities and participant realities can diminish the clinical relevance of findings and reduce engagement due to survey fatigue [20]. While face-to-face methods might address this through interaction, their high resource demands limit sample size and efficiency, presenting a practical challenge [21,22]. Consequently, methodological innovations using digital solutions are needed to better capture authentic participant perspectives while maintaining representativeness and efficiency.

Emerging research on phenotyping via social media supports this need. Studies reveal significant discrepancies between models based on biomedical databases and those derived from social media [23], positioning social media as a novel determinant of population health [24,25]. Some studies have even used these platforms to characterize KAP levels [26,27]. This offers a participant-centered pathway to overcome current KAP limitations. However, social media data are inherently heterogeneous, mixing authentic user narratives with commercial content, algorithmic biases, and intentional discourse [28]. This complexity necessitates rigorous curation and confirms the ongoing need for structured KAP surveys. Crucially, social media data can serve as a foundational resource for a "bottom-up" approach to questionnaire development, simulating participatory design and mitigating the constraints of researcher-centered frameworks [27]. Building on this methodological imperative, our study clinically focuses on HDM control in children and innovatively addresses KAP methodological limitations through a hybrid "top-down" and "bottom-up" approach to assess caregiver KAP levels.

Methods

Study Design

This cross-sectional study comprised 2 parts. First, we crawled data on HDM control measures from social media platforms and analyzed the content using latent Dirichlet allocation (LDA), an unsupervised topic modeling technique. Second, we developed a KAP questionnaire by integrating insights from social media analysis to supplement optimization. This questionnaire was then administered online to caregivers of children in Chongqing, China. Chongqing's warm (annual mean temperature: 18.4 °C) and humid (annual mean relative humidity: 79.0%) climate provides an environment highly conducive to HDM proliferation, establishing it as a representative region for studying HDM control in subtropical humid climates [29,30]. We followed the ChecKAP (Checklist for Reporting Items for Knowledge, Attitude, and Practice) and CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklists for reporting (Checklist 1) [31,32].

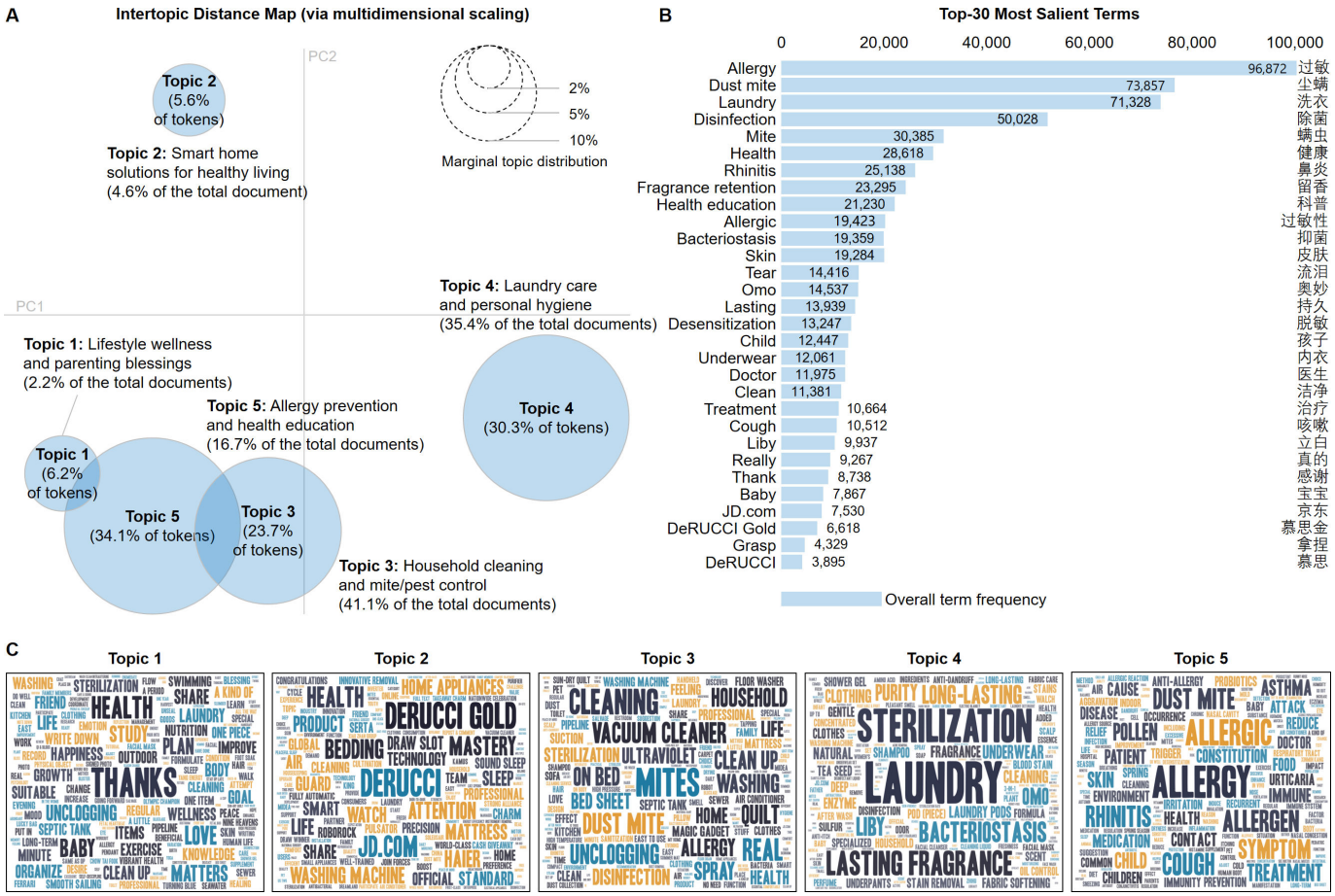
Ethical Considerations

This study received dual component-specific approvals from the Ethics Committee of the Children's Hospital of Chongqing Medical University (number 2025-402 for social media mining and number 2025-443 for the questionnaire survey). All social media data collection strictly complied with each platform's terms of service, drawing exclusively from fully public content. WeChat-derived materials were limited to public official account posts, with no semiprivate personal content collected. Data were fully deidentified at collection, no user identifiers retained, and individual informed consent for content creators was formally waived by the ethics committee for this noninteractive analysis. For the survey, participation was fully voluntary with no linked benefits or obligations. Proceeding to complete the survey after reviewing the landing-page ethics statement constituted implied informed consent; written informed consent was formally waived by the ethics committee for this fully anonymous, minimal-risk survey.

Social Media Mining and LDA Analysis

Raw data were collected from 5 major Chinese social media platforms (TikTok, Weibo, Rednote, WeChat, and KuaiShou) between August 1, 2023, and July 31, 2025 (2 years), using web crawlers [26,33]. Searching the keywords "dust mite" and "mite removal" yielded approximately 1 million initial entries. We identified substantial volumes of commercial material and semantically null text. To optimize semantic quality, we applied a domain-specific stopword lexicon to filter high-frequency terms (occurrence >1000) lacking substantive meaning, such as promotional terms ("red envelope," "repost," "subsidy," etc), while preserving brand-associated terms actively contributing to substantive discussions. Following deduplication, noise removal, and exclusion of irrelevant/incomplete content, 132,341 documents were retained for modeling (Multimedia Appendix 1). Unsupervised LDA topic modeling was performed [34,35], achieving optimal clustering at K=5 topics based on perplexity and coherence metrics, with a C_v coherence score of 0.41 falling within acceptable ranges for social media short-text contexts [36]. Interactive visualization using LDAvis confirmed strong topic distinctiveness, evidenced by minimal intertopic overlap and discrete spatial distribution in the 2D semantic space (Figure 1A), with "allergy," "dust mite," "laundry," and "disinfection" ranking the highest among the most salient terms (Figure 1B) [37,38]. Following this, we extracted the top-100 topic-word distributions for each topic, as well as the top 5 most salient terms for each topic under the λ=0 setting from LDAvis outputs, inputting these combined materials into DeepSeek-V3.2 to generate preliminary topic titles. Preliminary results were manually verified by reviewing the highest-contribution original documents assigned to each topic to ensure consistency between raw text content and the inducted topic theme, yielding the finalized set of topic labels [39,40].

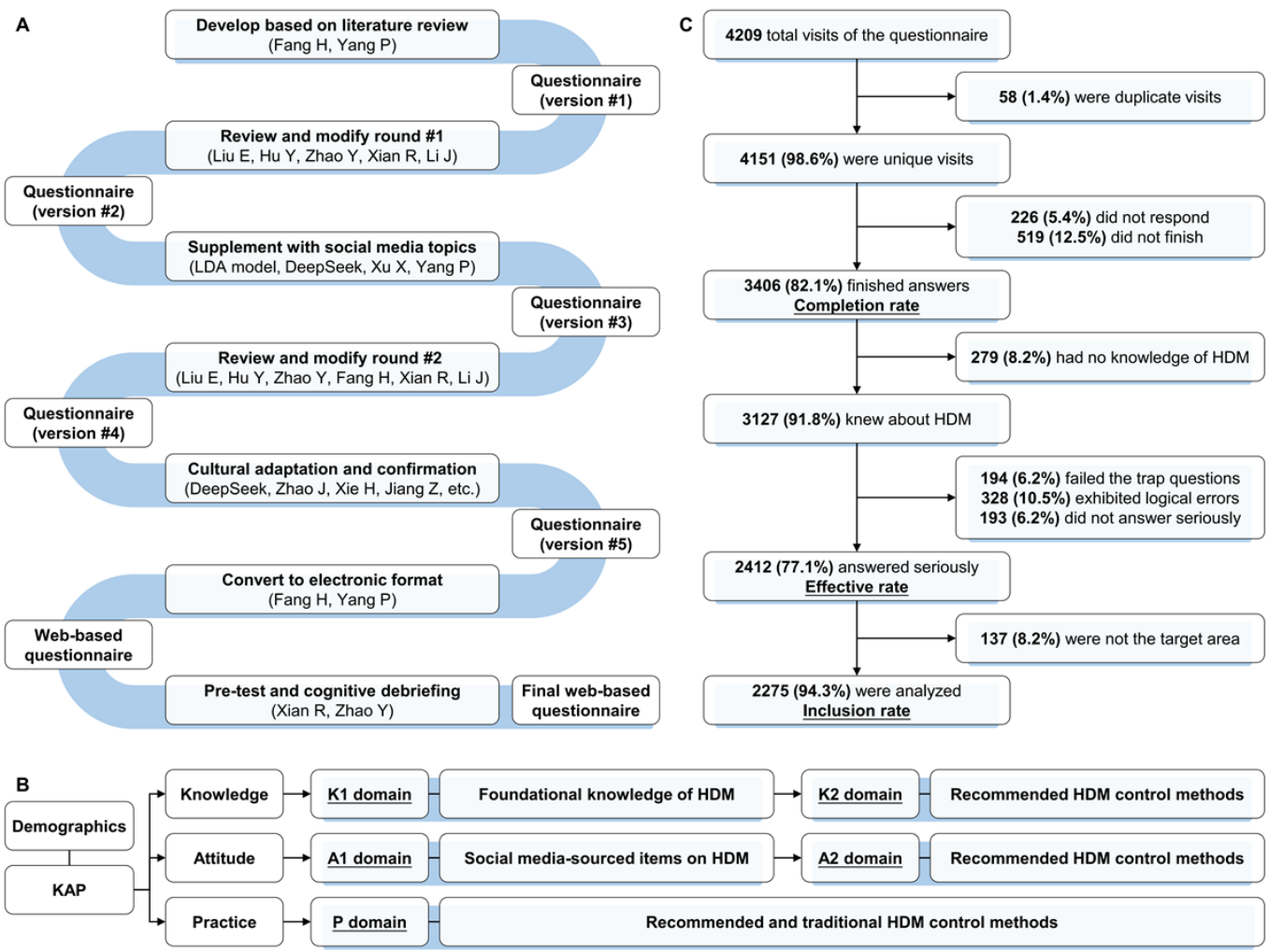
Figure 1. Social media mining and latent Dirichlet allocation (LDA) analysis of house dust mite (HDM) control discourse on 5 major Chinese social media platforms (August 1, 2023, to July 31, 2025). (A) Intertopic distance map generated via multidimensional scaling, showing minimal spatial overlap between the 5 identified topic clusters; (B) Top 30 most salient terms ranked by overall document frequency across all 132,345 included social media posts; (C) Top 200 salient terms ranked by term contribution for each of the 5 inducted topic clusters.



Development of the KAP Questionnaire

Questionnaire development followed a hybrid "top-down" (literature/expert-driven) and "bottom-up" (social media data-driven) approach (Figure 2A). The initial draft of the questionnaire was developed based on existing literature and clinical experience, forming the starting point of the traditional top-down development component. The draft then underwent its first round of revision by pediatricians from diverse subspecialties to refine the content framework. To incorporate public insights derived from social media discourse (the bottom-up component of our hybrid approach), we used the top 100 most salient terms per LDA topic as input for DeepSeek-V3.2, which generated a draft item pool for social media–related content (Multimedia Appendix 2) [39-41]. After integrating these newly generated items into the existing draft, a second round of revision was completed by the pediatrician expert team. Notably, a total of 16 items were excluded, including 14 risk assessment questions concerning HDM breeding areas, and 2 knowledge items regarding the optimal temperature and humidity for HDM proliferation. Following cultural adaptation, the finalized instrument was converted into a web-based format via the Lediaocha survey platform [42]. Finally, pilot testing among 20 participants and cognitive debriefing completed the preliminary validation process.

Figure 2. Hybrid house dust mite (HDM) control knowledge, attitudes, and practices (KAP) questionnaire development and data processing workflow. (A) Questionnaire design procedure integrating literature/expert-driven top-down and social media–driven bottom-up approaches. (B) Domain breakdown of the finalized questionnaire covering foundational knowledge, recommended methods awareness, social media attitude, recommended method attitude, and practice. (C) Steps from survey distribution to valid sample inclusion with a final valid sample size of 2275 after cleaning attention-check and logical consistency filters.



The finalized questionnaire comprised 2 sections with a total of 55 items, covering participants' demographic characteristics and HDM-related KAP, respectively (Figure 2B; Multimedia Appendix 3). KAP items were divided into 5 domains: foundational HDM knowledge (K1), awareness of recommended control methods (K2), attitude toward social media–sourced items (A1), attitude toward recommended methods (A2), and actual control practices (P). The recommended HDM control methods included in the K2 and A2 domains were sourced from a recent evidence summary (Multimedia Appendix 4) [11]. Additionally, to avoid excluding commonly adopted traditional practices not formally endorsed by the evidence summary, we incorporated several practices from the A1 domain into the P domain, which were classified as "traditional practice with uncertain evidence." Standardized phrases with explanatory definitions were established for KAP items to ensure terminological consistency throughout results reporting (Multimedia Appendix 5).

Survey Participants, Recruitment, and Data Processing

The survey was conducted in November 2025, targeting caregivers of children and adolescents (<18 years) residing in Chongqing. All participants were recruited via nonprobability convenience sampling. Recruitment was conducted at over 10 district-level child health care facilities across Chongqing, where triage nurses distributed leaflets with a study introduction and a QR code linking to the web-based questionnaire to caregivers during routine visits. The landing page of the questionnaire also presented a brief study introduction and an explicit statement of anonymous, voluntary participation to minimize perceived mandatory participation pressure and associated response bias. During survey implementation, all response initiation events were automatically counted to enable response rate calculation, and device IDs were recorded to prevent duplicate submissions. Upon survey closure, response data were downloaded by preauthorized research staff from the institutional cloud server, and all cloud-stored raw files were permanently deleted after download to ensure data security. All submitted

responses were screened against three sequential predefined exclusion criteria: (1) participants who reported no prior knowledge of HDM; (2) responses that failed attention-check (trap) questions, logical consistency validation, or self-reported reliability checks, per established methodology [43-45]; and (3) respondents not currently residing in Chongqing.

Sample Size Calculation

Sample size estimation followed the accepted 10 events per variable rule for logistic regression [46], which yielded a minimum required valid sample of 550 for 55 questionnaire items. To ensure sufficient statistical power for the 2 pre-specified subgroup comparisons (children aged ≤3 y vs >3 y; HDM-related online engagement: engaged vs nonengaged) and retain enough stratified sample to meet the events per variable requirements for multivariable adjustment, we applied a conservative inflation factor of 4 to avoid insufficient cell size in stratified analysis, bringing the target number of valid responses to 2200. To account for the approximate 50% invalid response rate commonly reported for public online surveys [47], we set a target of 4400 initiated survey responses. Ultimately, we obtained 4209 initiated survey responses, a small shortfall from the 4400 recruitment target that did not compromise analytical power. After data cleaning, 2275 valid questionnaires were retained for final analysis, exceeding the 2200 prespecified valid sample target (Figure 2C). A post hoc power calculation (2-sided α=.05, 80% power) for 2 independent proportions confirmed adequate sample size. The minimum detectable absolute difference in KAP implementation rate was 5.0% for the child age subgroup and 9.9% for the online engagement subgroup; most observed primary outcome differences exceeded these thresholds.

Statistical Analysis

Data analysis was performed using R (version 4.5) and Python (version 3.9) with the Gensim library (version 4.3). Qualitative data are presented as frequencies (percentages), while quantitative data are presented as medians (IQRs) after normality testing. KAP levels were reported as knowledge awareness, attitude endorsement, and practice implementation rates, with group means calculated for each domain. Intergroup comparisons were conducted using chi-square tests and Mann-Whitney *U* tests, stratified by 2 factors: child age (≤3 years vs >3 years) and caregiver HDM-related online engagement status (nonengaged vs engaged). Binary logistic regression assessed associations between child age, caregiver HDM-related online engagement, and KAP items, using the >3-year age group and nonengaged group as references. Adjusted odds ratios (ORs) with 95% CIs were derived from models adjusting for covariates showing significant intergroup differences. Statistical significance was set at $P<.05$.

Results

Social Media Discussions on HDM Control

The LDA model identified 5 topic clusters. Commercially driven and emotional content (topics 1-4, 83.3% of the documents) collectively dominated the information landscape, while health educational content (topic 5, 16.7% of the documents) constituted a comparatively limited proportion (Figure 1A). The five topics and their document proportions were as follows: (1) lifestyle wellness and parenting blessings (2.2%), characterized by narrative patterns dominated by emotional expression (6.2% of the tokens); (2) smart home solutions for healthy living (4.6%), marked by commercially driven content propagation (5.6% of the tokens); (3) household cleaning and mite/pest control (41.1%), showing prevalent dissemination of specific HDM control practices (23.7% of the tokens); (4) laundry care and personal hygiene (35.4%), demonstrating systematic conceptual conflation of "disinfection" and "mite elimination" in commercial soft articles (30.3% of the tokens); and (5) allergy prevention and health education (16.7%), distinguished by longer text spans, greater information density in science communication, and the highest token proportion (34.1% of the tokens; Figure 1C, Multimedia Appendix 1).

General Characteristics of the Survey Participants

Among the analyzed 2275 questionnaires (Table 1), children had a median age of 4.4 (IQR 2.6-6.0) years, with 42.7% (n=972) aged ≤3 years and 57.3% (n=1303) aged >3 years. Caregivers had a median age of 34 (IQR 30-38) years, were predominantly mothers (n=1730, 76.0%), and 77.2% held college or higher degrees. Reported allergy prevalence among children and their first-degree relatives was 43.9%, with HDM allergy specifically at 31.7%. Regarding self-rated HDM knowledge, 73.8% reported limited/basic understanding, while 26.3% considered themselves generally/very knowledgeable. Notably, only 11.2% reported engagement in social media discussions about HDM.

Table 1. General characteristics of the survey participants (children and caregivers) and intergroup comparisons.

		Child age			HDM[a]-related online engagement		
General characteristics	Overall	≤3 years	>3 years	*P* value	Nonengaged	Engaged	*P* value
Child age (years), median (IQR)	4.4 (2.6-6.0)	2.1 (1.1-3.3)	5.8 (5.0-7.6)	<.001	4.4 (2.6-6.0)	4.5 (2.1-6.7)	.51
Child age group, n (%)				—[b]			.78
≤3 years	972 (42.7)	—	—		861 (88.6)	111 (11.4)	
>3 years	1303 (57.3)	—	—		1159 (88.9)	144 (11.1)	
Child sex, n (%)				.24			.42

General characteristics	Overall	Child age			HDM[a]-related online engagement		
		≤3 years	>3 years	*P* value	Nonengaged	Engaged	*P* value
Male	1123 (49.4)	466 (47.9)	657 (50.4)		991 (49.1)	132 (51.8)	
Female	1152 (50.6)	506 (52.1)	646 (49.6)		1029 (50.9)	123 (48.2)	
Caregiver age (years), median (IQR)	34 (30-38)	32 (30-36)	35 (32-39)	<.001	34 (30-38)	34 (30-39)	.65
Caregiver role, n (%)				.01			.84
Mother	1730 (76.0)	719 (74.0)	1011 (77.6)		1540 (76.2)	190 (74.5)	
Father	331 (14.5)	140 (14.4)	191 (14.7)		293 (14.5)	38 (14.9)	
Other relatives	141 (6.2)	71 (7.3)	70 (5.4)		122 (6.0)	19 (7.5)	
Grandparents	73 (3.2)	42 (4.3)	31 (2.4)		65 (3.2)	8 (3.1)	
Caregiver education level, n (%)				<.001			.04
Junior high school or below	166 (7.6)	43 (4.4)	123 (9.4)		150 (7.4)	16 (6.3)	
High school	352 (15.5)	108 (11.1)	244 (18.7)		318 (15.7)	34 (13.3)	
College	1574 (69.2)	722 (74.3)	852 (65.4)		1380 (68.3)	194 (76.1)	
Postgraduate or above	183 (8.0)	99 (10.2)	84 (6.4)		172 (8.5)	11 (4.3)	
Family member with allergic diseases, n (%)				.15			.48
Yes	998 (43.9)	413 (42.5)	585 (44.9)		890 (44.1)	108 (42.4)	
No	1169 (51.4)	504 (51.9)	665 (51.0)		1031 (51.0)	138 (54.1)	
Not sure	108 (4.7%)	55 (5.7)	53 (4.1)		99 (4.9)	9 (3.5)	
Family member with HDM allergy, n (%)				.09			.01
Yes	722 (31.7)	290 (29.8)	432 (33.2)		622 (30.8)	100 (39.2)	
No	1298 (57.1)	560 (57.6)	738 (56.6)		1163 (57.6)	135 (52.9)	
Not sure	255 (11.2)	122 (12.6)	133 (10.2)		235 (11.6)	20 (7.8)	
Self-rated HDM knowledge, n (%)				.03			<.001
Limited understanding	930 (40.9)	403 (41.5)	527 (40.4)		852 (42.2)	78 (30.6)	
Basic understanding	748 (32.9)	340 (35.0)	408 (31.3)		675 (33.4)	73 (28.6)	
Generally knowledgeable	500 (22.0)	185 (19.0)	315 (24.2)		417 (20.6)	83 (32.5)	
Very knowledgeable	97 (4.3)	44 (4.5)	53 (4.1)		76 (3.8)	21 (8.2)	
HDM-related online engagement, n (%)				.78			—
Nonengaged	2020 (88.8)	861 (88.6)	1159 (88.9)		—	—	
Engaged	255 (11.2)	111 (11.4)	144 (11.1)		—	—	

[a]HDM: house dust mite.
[b]— means not available.

KAP Levels of HDM Control Among the Participants

The average levels in the 5 KAP domains are presented in Figure 3 (Multimedia Appendix 6). In foundational knowledge domain (K1), the mean awareness rate was 62.9% (SD 28.4%; range 19.7%-91.0%). Awareness was suboptimal for "HDM-Demodex Distinction" (52.8%) and "HEPA Filtration Necessity" (19.7%). For recommended control methods (K2 domain), the mean awareness rate was 84.7% (SD 7.2%; range 67.1%-93.1%), with "Plush Toy Avoidance" showing the highest awareness (93.1%). Social media emerged as the primary knowledge source for HDM control (80.7%), followed by health care professionals (34.8%), with traditional media being the least reported (29.7%; Table 2).

Figure 3. Average knowledge, attitudes, and practices (KAP) domain levels of caregivers for house dust mite (HDM) control in Chongqing, China (November 2025).

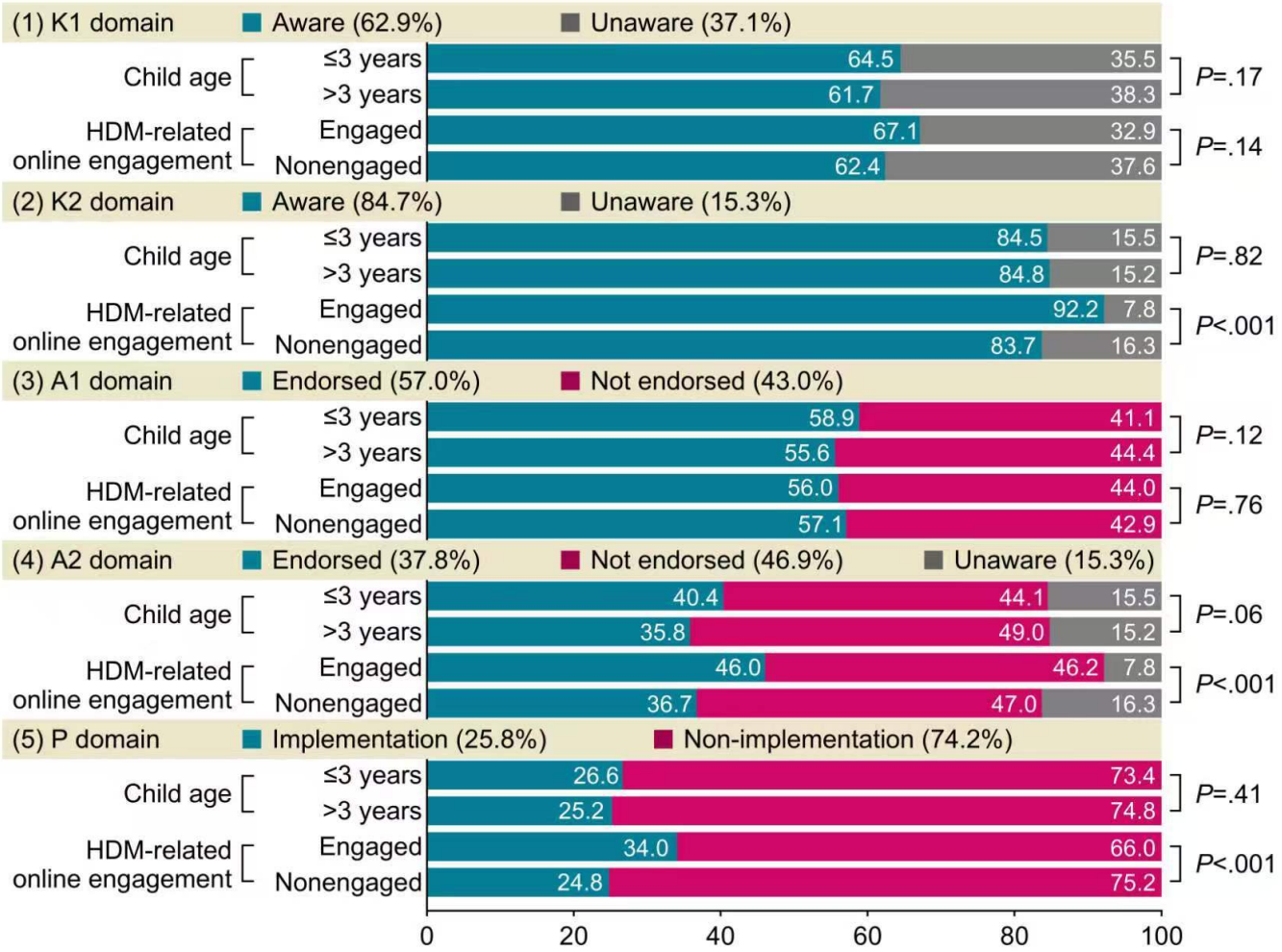


Table 2. Knowledge sources and implementation barriers for house dust mite (HDM) control and intergroup comparisons.

		Child age			HDM-related online engagement		
Domains and items of KAP[a]	Overall	≤3 years	>3 years	*P* value	Nonengaged	Engaged	*P* value
Knowledge sources, n (%)							
Social media	1835 (80.7)	794 (81.7)	1041 (79.9)	.28	1599 (79.2)	236 (92.5)	<.001
Health care professionals	791 (34.8)	334 (34.4)	457 (35.1)	.73	674 (33.4)	117 (45.9)	<.001
Traditional media	676 (29.7)	278 (28.6)	398 (30.5)	.32	575 (28.5)	101 (39.6)	<.001
Implementation barriers, n (%)							
Knowledge gaps on methods	1080 (47.5)	480 (49.4)	600 (46.0)	.12	971 (48.1)	109 (42.7)	.11
Time constraints	942 (41.4)	395 (40.6)	547 (42.0)	.52	829 (41.0)	113 (44.3)	.32
Side effect concerns	786 (34.5)	328 (33.7)	458 (35.1)	.49	689 (34.1)	97 (38.0)	.21
Price issue	693 (30.5)	284 (29.2)	409 (31.4)	.27	598 (29.6)	95 (37.3)	.01
Living condition restrictions	462 (20.3)	197 (20.3)	265 (20.3)	.97	407 (20.1)	55 (21.6)	.60
Perceived ineffectiveness	439 (19.3)	194 (20.0)	245 (18.8)	.49	386 (19.1)	53 (20.8)	.52

[a]KAP: knowledge, attitudes, and practices.

Attitudes toward social media–sourced items (A1 domain) revealed a complex picture (mean endorsement: 57.0%, SD 17.4%; range 23.6%-78.9%). Caregivers showed positive inclinations ("Prevention-Priority": 75.7%) and educational needs ("Authoritative Education Need": 77.8%). However, challenges were evident: "Allergy Information Overload" was reported by 58.5% and "Efficacy Uncertainty" by 57.4%, and safety concerns were high ("Chemical Spray Harm Concerns": 70.6%; "Disinfectant Overuse Risks": 65.7%). Misconceptions persisted, with "Antimicrobial Detergent Belief" held by 44.0% and "Ultraviolet Efficacy Belief" by 39.4%. For attitudes toward recommended methods (A2 domain), mean endorsement was lower at 37.8% (SD 14.6%; range 16.5%-64.4%). Physical interventions

such as "Endorsed Carpet Maintenance" (64.4%) and "Endorsed Hard Flooring" (56.6%) were more accepted, while device-based methods ("Endorsed HEPA Purification": 33.3%; "Endorsed Dry Environment": 20.8%) and chemical methods ("Endorsed Chemical Sprays": 21.7%) ranked lower (Multimedia Appendix 6).

Control practices (P domain) revealed implementation gaps and behavioral missteps (Multimedia Appendix 6). The mean implementation rate across surveyed methods was low (25.8%, SD 17.4%; range: 1.8%-69.5%), with 12.7% implementing none of the listed methods. Notably, traditional practice with uncertain evidence, "Sun-Drying Bedding" demonstrated the highest implementation (69.5%) and "Using Antimicrobial Detergents" (34.9%) the second highest, while recommended method "Using Mite-Proof Encasements" (highest evidence grade: 1b) ranked lowest (13.7%). Alarmingly, 15.3% reported another traditional practice with uncertain evidence ("Using Ultraviolet Devices"), among whom 26.6% were unaware of mandatory "Ultraviolet Room Evacuation" requirements. Among the key reported barriers, knowledge gaps on methods (47.5%) ranked the highest, followed by time constraints (41.4%), with perceived ineffectiveness (19.3%) being the least common (Table 2).

Association Between Child Age and Caregiver HDM-Related Online Engagement and KAP Items

Caregivers of younger children (≤3 years) reported lower self-rated HDM knowledge (generally/very knowledgeable: 23.5% vs 28.3%; *P*=.01). Logistic regression showed they had significantly higher awareness of "HDM Habitats" (OR 1.43, 95% CI 1.11-1.82; *P*=.005), while exhibiting lower awareness of "Plush Toy Avoidance" (OR 0.68, 95% CI 0.48-0.95; *P*=.02) and "Freeze-Dry Plush Protocol" (OR 0.83, 95% CI 0.68-0.99; *P*=.04). They demonstrated higher agreement with "Allergy Information Overload" (OR 1.39, 95% CI 1.15-1.67; *P*<.001) and "Efficacy Uncertainty" (OR 1.23, 95% CI 1.01-1.52; *P*=.04), along with higher endorsement of "Child-Specific Interventions" (OR 1.22, 95% CI 1.00-1.47; *P*=.045) and "Ultraviolet Room Evacuation" (OR 1.45, 95% CI 1.18-1.82; *P*<.001) in A1 domain. They also showed higher endorsement of multiple items in the A2 domain (Table 3, Multimedia Appendix 7). Behaviorally, they reported higher implementation of "Using Air Purifiers" (OR 1.28, 95% CI 1.04-1.59; *P*=.02).

Table 3. Analysis of the association of child age and caregiver house dust mite (HDM)–related online engagement with knowledge, attitudes, and practices (KAP) items[a].

	Child age (≤3 years)		HDMs-related online engagement (engaged)	
Domains and items of KAP	OR (95% CI)	*P* value	OR (95% CI)	*P* value
K1 domain: foundational knowledge of HDM				
"HDM Habitats"	1.43 (1.11-1.82)	.005	1.79 (1.14-2.82)	.01
"HDM-Demodex Distinction"	—[b]	—	0.76 (0.58-0.99)	.04
"HEPA[c] Filtration Necessity"	—	—	1.74 (1.28-2.35)	<.001
K2 domain: recommended HDM control knowledge (evidence grade)				
"Dry Environment Maintenance" (2b)	—	—	2.12 (1.29-3.49)	.003
"HEPA Air Purification" (5a)	—	—	2.34 (1.47-3.74)	<.001
"Leather/Wooden Furniture" (1b)	—	—	2.36 (1.37-4.06)	.002
"Blind Use or Curtain Washing" (5a)	—	—	1.93 (1.08-3.45)	.03
"Hard Flooring Installation" (1b)	—	—	2.14 (1.07-4.27)	.03
"Plush Toy Avoidance" (5b)	0.68 (0.48-0.95)	.02	—	—
"Mite-Proof Encasements" (1b)	—	—	2.75 (1.60-4.71)	<.001
"Down Pillow Selection" (3b)	—	—	2.12 (1.40-3.21)	<.001
"Thermal Control" (5a)	—	—	2.07 (1.24-3.46)	.005
"Chemical Acaricide Sprays" (1b)	—	—	1.66 (1.14-2.42)	.009
"HEPA Vacuuming" (5a)	—	—	2.11 (1.45-3.06)	<.001
"≥60 °C Bedding Washing" (5b)	—	—	2.35 (1.41-3.91)	.001
"Tumble-Drying Textiles" (5b)	—	—	3.46 (1.68-7.13)	.001
"Freeze-Dry Plush Protocol" (5b)	0.83 (0.68-0.99)	.04	2.12 (1.53-2.94)	<0.001
A1 domain: social media-sourced items on HDM				
"Allergy Information Overload"	1.39 (1.15-1.67)	<.001	—	—
"Parental Care Efficacy"	—	—	0.73 (0.54-0.98)	.04
"Thorough Cleaning Necessity"	—	—	0.65 (0.48-0.88)	.005
"Efficacy Uncertainty"	1.23 (1.01-1.52)	.04	—	—
"Ultraviolet Efficacy Belief"	—	—	1.44 (1.11-1.89)	.007
"Ultraviolet Room Evacuation"	1.45 (1.18-1.82)	<.001	—	—

	Child age (≤3 years)		HDMs-related online engagement (engaged)	
Domains and items of KAP	OR (95% CI)	*P* value	OR (95% CI)	*P* value
"Chemical Spray Harm Concerns"	—	—	0.69 (0.52-0.90)	.007
"Prevention-Priority"	—	—	0.74 (0.55-1.00)	.048
"Child-Specific Interventions"	1.22 (1.00-1.47)	.045	—	—
"Allergy Impact Exaggeration"	—	—	1.39 (1.04-1.87)	.03
A2 domain: recommended HDM control methods (evidence grade)				
"Endorsed Dry Environment" (2b)	—	—	1.40 (1.04-1.90)	.03
"Endorsed HEPA Purification" (1c)	1.43 (1.19-1.75)	<.001	—	—
"Endorsed Blind/Curtain Care" (5a)	1.23 (1.02-1.47)	.03	—	—
"Endorsed Hard Flooring" (1b)	1.25 (1.03-1.49)	.02	—	—
"Endorsed Mite-Proof Encasements" (1b)	—	—	1.50 (1.13-1.98)	.005
"Endorsed Down Pillows" (3b)	1.25 (1.02-1.54)	.03	1.50 (1.12-2.01)	.006
"Endorsed Chemical Sprays" (1b)		—	1.56 (1.16-2.11)	.004
"Endorsed HEPA Vacuuming" (5b)	1.37 (1.11-1.67)	.002	—	—
"Endorsed ≥60 °C Washing" (5b)	1.20 (1.00-1.45)	.049	—	—
"Endorsed Tumble-Drying" (5b)	1.33 (1.11-1.61)	.002	1.51 (1.14-2.02)	.005
"Endorsed Freeze-Dry Protocol" (5b)	—	—	1.68 (1.22-2.32)	.001
"Endorsed Carpet Maintenance" (5b)	1.32 (1.09-1.61)	.005	—	—
P domain: actual practices (implementation vs nonimplementation)				
"Using Dehumidifiers"	—	—	1.93 (1.45-2.57)	<.001
"Using Ultraviolet Devices"	—	—	1.85 (1.35-2.53)	<.001
"Washing Bedding ≥60 °C"	—	—	1.82 (1.39-2.39)	<.001
"Tumble-Drying Textiles"	—	—	1.62 (1.23-2.14)	.001
"Using Mite-Proof Encasements"	—	—	1.92 (1.39-2.65)	<.001
"Using Air Purifiers"	1.28 (1.04-1.59)	.02	1.47 (1.08-1.98)	.01
"Using Chemical Sprays"	—	—	1.52 (1.10-2.10)	.01
"No Control Measures"	—	—	0.60 (0.37-0.98)	.04

[a]All analyses compared children aged ≤3 years (n=972) to the reference group >3 years (n=1303) and compared caregivers with house dust mite (HDM)–related online engagement (n=255) to the reference nonengaged group (n=2020). The model of child age group was adjusted for caregiver role (*P*=.01), caregiver education level (*P*<.001), and self-rated HDM knowledge (*P*=.01), while caregiver age was not included due to its collinearity with child age. The model of HDM-related online engagement group was adjusted for caregiver education level (*P*=.04), self-rated HDM knowledge (*P*<.001), and family member with HDM allergy (*P*=.01). Only items with a statistically significant association (P_{adj}<.05) in the adjusted models are presented in this table, and full results are provided in Multimedia Appendix 7. Evidence grade [11]: (1b) systematic reviews of multiple randomized controlled trials and other interventional studies; (1c) single randomized controlled trial; (2b) systematic reviews of multiple quasi-experimental studies and other low-quality interventional studies; (3b) systematic reviews of multiple cohort studies and other low-quality observational studies; (5a) systematic review of expert opinions; and (5b) expert consensus.

[b]— means not statistically significant in the adjusted model.

[c]HEPA: high-efficiency particulate air filter.

[d]Not available.

Caregivers with HDM-related online engagement demonstrated consistently better performance on multiple knowledge (K1 and K2 domains), attitude (A2 domain), and practice (P domain) items (Table 3; Multimedia Appendix 7). Paradoxically, within the A1 domain, this group showed greater agreement with "Allergy Impact Exaggeration" (OR 1.39, 95% CI 1.04-1.87; *P*=.03), while displaying lower endorsement of "Thorough Cleaning Necessity" (OR 0.65, 95% CI 0.48-0.88; *P*=.005), "Prevention-Priority" (OR 0.74, 95% CI 0.55-1.00; *P*=.048), "Parental Care Efficacy" (OR 0.73, 95% CI 0.54-0.98; *P*=.04), and "Chemical Spray Harm Concerns" (OR 0.69, 95% CI 0.52-0.90; *P*=.007). Additionally, they exhibited higher endorsement of "Ultraviolet Efficacy Belief" (OR 1.44, 95% CI 1.11-1.89; *P*=.007) and greater implementation of "Using Ultraviolet Devices" (OR 1.85, 95% CI 1.35-2.53; *P*<.001).

Discussion

Principal Results

This study provides, to our knowledge, among the first comprehensive integration of social media mining and questionnaire surveys to characterize caregiver KAP regarding HDM control. We identified pervasive commercialization and low-value content in social media data that may be associated with suboptimal caregiver KAP. High awareness of foundational knowledge coexisted with low

endorsement and implementation of recommended HDM control methods, accompanied by widespread implementation of traditional practices with uncertain evidence. Caregiver attitudes reflected proactive prevention intentions and educational needs but were challenged by significant barriers including information overload, decision-making difficulties, safety concerns, and entrenched belief in traditional practices with uncertain evidence. Caregivers of younger children demonstrated lower self-rated knowledge alongside higher endorsement of control measures, while exhibiting greater vulnerability to information overload and decision-making difficulties. Caregivers with HDM-related online engagement showed enhanced KAP but also showed patterns consistent with risk desensitization and reinforced misconceptions.

Methodological Innovation

Our study innovates by embedding social media discourse into KAP questionnaire development, forging a hybrid "top-down" and "bottom-up" framework. This approach aligns with the core purpose of KAP research: understanding the target population's needs. Their voices must actively shape the questions, not be an afterthought [17,18]. This participant-centered design proved crucial, enabling us to capture the complex and often contradictory attitudes in the social media–sourced A1 domain, as well as the emergence and risks associated with practices such as ultraviolet irradiation. While social media offers unparalleled access to public discourse, its potential for enriching health research remains underutilized [23,24]. Notably, the A1 domain achieved significantly higher endorsement rates (57.0%) than the literature-sourced A2 domain (37.8%), demonstrating closer alignment with caregivers' expressed priorities. Collectively, these findings demonstrate the unique value of leveraging social media data to ground KAP research in authentic participant perspectives and enhance ecological validity.

Unmet Needs of HDM Control in Childhood

The chronic trajectory of childhood asthma links to persistent allergic airway inflammation established early in life [5]. Age-specific HDM sensitization patterns reinforce this link, reflecting the natural course in the absence of targeted intervention [2-4]. Critically, environmental control measures have been shown in selected trials to reduce the incidence of new HDM sensitization [48,49]. Collectively, this evidence points to a potential prevention strategy: implementing targeted environmental control early in life to prevent HDM sensitization and thereby mitigate both asthma incidence and its long-term chronicity (primary prevention). Our study provides KAP evidence supporting this strategy. Currently, both caregivers' endorsement of recommended HDM control methods (37.8%) and their actual implementation rates (25.8%) remain suboptimal relative to the local 31.7% HDM allergy prevalence, with knowledge gaps related to correct operational methods identified as the predominant barrier (47.5%). Compounding this gap, caregivers of younger children exhibited 3 coexisting states: unmet knowledge needs (lower self-rated knowledge), high motivation (stronger control endorsement), and helplessness (information overload/decision difficulties). This triad creates an imperative for reinforced health education.

While our findings highlight substantial unmet KAP needs among child caregivers for standardized HDM control, it must be noted that the clinical benefit of routine allergen avoidance measures remains a widely debated topic; Global Initiative for Asthma (GINA) and Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines offer conditional or against recommendations for single-measure approaches [9,10]. For primary prevention scenarios, several cohort studies have provided supporting evidence that multifaceted interventions implemented in early life may reduce subsequent childhood asthma risk [50-54]. However, it remains unclear whether the observed protective signal can be definitively attributed to allergen control specifically, rather than other parallel co-intervention components. In secondary and tertiary prevention settings, the consistent clinical benefit delivered by HDM control interventions appears limited and far less demonstrable for symptomatic children with confirmed asthma or allergic rhinitis [55,56]. Collectively, further targeted research is needed to clarify HDM control's independent causal role in childhood asthma prevention.

Beyond these open evidentiary questions, encouraging widespread HDM control also requires careful consideration of potential downsides [57]. A primary concern is ensuring that caregivers adopt measures that effectively reduce HDM levels without inadvertently fostering excessive cleanliness or disinfection habits [58,59]. Our results further highlight this issue, indicating substantial endorsement of disinfection methods (antimicrobial detergents: 44.0%; ultraviolet irradiation: 39.4%), which may be reinforced by social media narratives conflating "disinfection" and "mite elimination" (topic 4). Maintaining an environment supporting biodiversity is crucial for fostering immune balance [60,61], underscoring the critical need for nuanced health education that emphasizes this balance. On the other hand, achieving such shifts in understanding and behavior is challenging; even highly successful programs such as the Finnish Allergy Programme (2008-2018) demonstrated that changes in public and patient perspectives occur slowly [62]. For this reason, establishing proactive public monitoring systems to detect and correct caregiver cognitive biases early in their formation is a high-priority unmet need. Social media offers a promising platform for such monitoring efforts [24,63], as shown in our findings.

The Double-Edged Sword of Social Media

Social media platforms are rife with health misinformation, negative content, and exaggerated claims [64-66]. Our LDA analysis confirmed this environment, wherein commercial marketing predominated in thematic volume while professional health education constituted a limited proportion, which correlates with the observed unscientific beliefs and practices such as ultraviolet irradiation. Moreover, a subset of content featuring maternal-infant emotional lexis including terms such as "healthy," "safe," "happy," and "growth" elicits

caregiver empathic resonance that may impair information quality discernment; this mechanism operates through social media contexts diverting cognitive resources toward emotional engagement rather than reliability evaluation [67]. Concurrently, these affective narratives establish psychologically vulnerable scenarios wherein commercial entities may influence such sentiments to propagate unsubstantiated product claims and advertisements [68-70]. This structural imbalance in information representation may be further amplified by a distinct textual feature: health educational posts are substantially longer and more information-dense than commercial or emotional HDM-control content, a pattern that creates an inherent disadvantage on social platforms that prioritize fast, short-form, high-engagement content.

On the other hand, while caregivers with HDM-related online engagement demonstrated greater knowledge awareness, they paradoxically exhibited more negative attitudes toward HDM control and showed no advantage in managing educational needs, information overload, or decision-making difficulties. Algorithmic reinforcement dynamics provide a key explanatory mechanism: expressions of concern online trigger increased exposure to affectively congruent content, including negative narratives offering psychological catharsis, thereby establishing a self-perpetuating cycle [71,72]. This pattern is observed in our data: caregivers with HDM-related online engagement were significantly more likely to agree that allergy impacts are exaggerated, a correlation consistent with platform-related cognitive framing effects. However, given the cross-sectional study design, future studies linking respondents' objective platform usage metrics with longitudinal follow-up assessments are required to establish temporal or causal relationships for the observed associations. Nevertheless, with social media serving as the reported knowledge source for 80.7% of caregivers and a strong demand (77.8%) for authoritative health education on these platforms, their potential as effective communication channels is underscored. Such demand necessitates proactive engagement from health care professionals and institutions to curate recommended discourse through designed interventions rather than passive observation [73,74].

Limitations

Several limitations need to be considered when interpreting this study. First, the web-based survey and social media text corpus both carry inherent selection bias, leading to underrepresentation of less educated caregivers and grandparent caregivers. Routine data cleaning also excluded some initiated responses that did not meet prespecified quality criteria, so nonresponse bias cannot be fully ruled out. Second, the nationally sourced social media corpus and geographically restricted Chongqing survey sample create a spatial mismatch, which limits direct generalizability of the survey KAP findings to broader populations, and the child age distribution is concentrated in early childhood, so the ">3 years" subgroup is effectively restricted to preschool and early primary children rather than older children or adolescents. Third, caregiver practices were measured via self-report without independent objective validation, which means recall bias may influence these estimates. Finally, no individual-level linkage exists between the social media and survey datasets, so all observed associations can only be interpreted as correlational, with no ability to fully rule out reverse causation.

Conclusions

This study delineates the complex KAP landscape of caregivers regarding HDM control, revealing critical gaps (especially in families with young children) that call for targeted educational initiatives. It establishes the value of integrating social media–derived insights to create more participant-centered KAP questionnaires. Simultaneously, it confirms social media's use as a potent health education channel, while starkly highlighting its role in disseminating misconceptions, necessitating vigilant monitoring and corrective interventions.

Acknowledgments

The authors gratefully acknowledge Dr Shuyao Zhu (Pediatric Neurology), Dr Keran Ling (Neonatology), and Dr Yuxin Lyu (Pediatric Dermatology) for their critical review and refinement of the questionnaire instrument.

Generative AI Use

The authors declare that DeepSeek-V3.2 generative AI tool was used under full human supervision for the development of preliminary topic title generation per the GAIDeT taxonomy (2025). All outputs were subject to full manual review to produce the finalized set of topic labels, and the authors take full and exclusive responsibility for the entire content and final outcomes of this manuscript, with no generative AI tool listed as an author.

Funding

This study was funded by the Noncommunicable Chronic Diseases—National Science and Technology Major Project (grant ID 2024ZD0541100) and the China Postdoctoral Science Foundation (2025MD784149). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

Conceptualization: PY, RX, YZ, HF (lead), EL
Data curation: PY (lead), RX, HF
Formal analysis: PY (lead), RX, HF
Funding acquisition: HF (lead), EL

Investigation: RX (lead), JL, JZ, HX, ZJ, YH, XX, NZ
Methodology: PY, XX, HF (lead)
Project administration: YZ, EL (lead)
Resources: YH, XX, NZ, YZ (lead), EL
Supervision: YZ, HF, EL (lead)
Validation: JL, JZ, HX, ZJ, YZ, EL (lead)
Visualization: PY, RX, HF (lead)
Writing – original draft: PY, RX, HF (lead)
Writing – review & editing: JL, JZ, HX, ZJ, YH, XX, NZ, YZ, EL (lead)
PY and RX contributed equally to this work and should be considered joint first authors. YZ (1294428871@qq.com), HF (heping.f@outlook.com), and EL (emliu186@126.com) contributed equally to this work and should be considered joint corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The data volumes and layout of latent Dirichlet allocation analysis.
[PDF File (Adobe File), 5975 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Methodological rationale for large language model–assisted item generation and top 100 salient terms per latent Dirichlet allocation topic.
[PDF File (Adobe File), 186 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Questionnaire used in this study.
[PDF File (Adobe File), 194 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Evidence summary of dust mite control in indoor environments.
[PDF File (Adobe File), 63 KB-Multimedia Appendix 4]

Multimedia Appendix 5

Standardized phrases and explanatory definitions for knowledge, attitudes, and practices items.
[PDF File (Adobe File), 90 KB-Multimedia Appendix 5]

Multimedia Appendix 6

Intergroup comparisons of knowledge, attitudes, and practices items by child age and caregiver social media activity groups.
[PDF File (Adobe File), 97 KB-Multimedia Appendix 6]

Multimedia Appendix 7

Association between child age and caregiver social media activity and knowledge, attitudes, and practices items.
[PDF File (Adobe File), 137 KB-Multimedia Appendix 7]

Checklist 1

ChecKAP checklist.
[PDF File (Adobe File), 151 KB-Checklist 1]

References

1. Shin YH, Hwang J, Kwon R, et al. Global, regional, and national burden of allergic disorders and their risk factors in 204 countries and territories, from 1990 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. Allergy. Aug 2023;78(8):2232-2254. [doi: 10.1111/all.15807] [Medline: 37431853]
2. Howard R, Belgrave D, Papastamoulis P, Simpson A, Rattray M, Custovic A. Evolution of IgE responses to multiple allergen components throughout childhood. J Allergy Clin Immunol. Oct 2018;142(4):1322-1330. [doi: 10.1016/j.jaci.2017.11.064] [Medline: 29428391]
3. Posa D, Perna S, Resch Y, et al. Evolution and predictive value of IgE responses toward a comprehensive panel of house dust mite allergens during the first 2 decades of life. J Allergy Clin Immunol. Feb 2017;139(2):541-549. [doi: 10.1016/j.jaci.2016.08.014] [Medline: 27793411]

4. Fang H, Li J, Ren L, Liu E. Age-related differences in IgE between childhood and adulthood allergic asthma: analysis of NHANES 2005-2006. World Allergy Organ J. Dec 2023;16(12):100842. [doi: 10.1016/j.waojou.2023.100842] [Medline: 38213391]
5. Illi S, von Mutius E, Lau S, et al. Perennial allergen sensitisation early in life and chronic asthma in children: a birth cohort study. Lancet. Aug 26, 2006;368(9537):763-770. [doi: 10.1016/S0140-6736(06)69286-6] [Medline: 16935687]
6. Sobczak M, Kowal K, Pawliczak R. Are there effective methods to reduce exposure to house dust mite allergens? A meta-analysis of randomized clinical trials. Int Forum Allergy Rhinol. Aug 2025;15(8):792-802. [doi: 10.1002/alr.23565] [Medline: 40089899]
7. Zuiani C, Custovic A. Update on house dust mite allergen avoidance measures for asthma. Curr Allergy Asthma Rep. Jun 19, 2020;20(9):50. [doi: 10.1007/s11882-020-00948-y] [Medline: 32561997]
8. Morgan WJ, Crain EF, Gruchalla RS, et al. Results of a home-based environmental intervention among urban children with asthma. N Engl J Med. Sep 9, 2004;351(11):1068-1080. [doi: 10.1056/NEJMoa032097] [Medline: 15356304]
9. Global strategy for asthma management and prevention. Global Initiative for Asthma (GINA); 2026. URL: https://ginasthma.org/wp-content/uploads/2026/05/GINA-2026-Strategy-Report-WMS.pdf [Accessed 2026-10-09]
10. Brozek JL, Bousquet J, Baena-Cagnani CE, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) guidelines: 2010 revision. J Allergy Clin Immunol. Sep 2010;126(3):466-476. [doi: 10.1016/j.jaci.2010.06.047] [Medline: 20816182]
11. Zhou Q, Zhang H, Kang N, et al. Evidence summary of indoor environmental controls in dust mite allergic patients [Article in Chinese]. Chin J Allergy Clin Immunol. 2025;19(1):32-41. [doi: 10.3969/j.issn.1673-8705.2025.01.005]
12. Grant TL, Wood RA, Chapman MD. Indoor environmental exposures and their relationship to allergic diseases. J Allergy Clin Immunol Pract. Oct 2023;11(10):2963-2970. [doi: 10.1016/j.jaip.2023.08.034] [Medline: 37652348]
13. Bracken M, Fleming L, Hall P, et al. The importance of nurse-led home visits in the assessment of children with problematic asthma. Arch Dis Child. Oct 2009;94(10):780-784. [doi: 10.1136/adc.2008.152140] [Medline: 19546102]
14. Pingitore G, Pinter E. Environmental interventions for mite-induced asthma: a journey between systematic reviews, contrasting evidence and clinical practice. Eur Ann Allergy Clin Immunol. May 2013;45(3):74-77. [Medline: 23862395]
15. Liu S, Zhou Q, Dai B, et al. Parents' knowledge, attitudes and practices towards the prevention and treatment of dust mite allergy: a cross-sectional study in Shenyang (China). BMJ Open. Dec 20, 2024;14(12):e085905. [doi: 10.1136/bmjopen-2024-085905] [Medline: 39806689]
16. Wei X, Chen X, Chen H, Chen J. Development and validation of a knowledge, attitude, and practice questionnaire related to prone positioning among intensive care unit nurses. BMC Nurs. Oct 8, 2025;24(1):1249. [doi: 10.1186/s12912-025-03900-5] [Medline: 41063108]
17. Duburcq A, Molins M, Debremaeker S, et al. Overview of migraine care, impact of the disease, and patient experience in France through patient voices: a cross-sectional and participatory survey study. J Headache Pain. Apr 14, 2025;26(1):76. [doi: 10.1186/s10194-025-02018-y] [Medline: 40229668]
18. Schreck LD, Meyer S, Pedersen ESL, et al. "Nobody has ever spoken to me about PCD and fertility issues": fertility experiences of people with primary ciliary dyskinesia and their family caregivers. Health Expect. Jun 2025;28(3):e70316. [doi: 10.1111/hex.70316] [Medline: 40470648]
19. Patel P. KAP survey: does it really measure knowledge, attitudes and practices? Natl J Community Med. 2022;13(5):271-273. [doi: 10.55489/njcm.130520222063]
20. Christensen AI, Ekholm O, Glümer C, Juel K. Effect of survey mode on response patterns: comparison of face-to-face and self-administered modes in health surveys. Eur J Public Health. Apr 2014;24(2):327-332. [doi: 10.1093/eurpub/ckt067] [Medline: 23766339]
21. Ödling M, Lagercrantz B, Lundin S, Sandelowsky H, Janson C, Kull I. Challenges of transitional care for young allergy and asthma patients from healthcare professionals' perspectives. Clin Transl Allergy. Aug 2025;15(8):e70088. [doi: 10.1002/clt2.70088] [Medline: 40760775]
22. Zhang J, Wu Q, Wang X, Zhang Y, Helena van Velthoven M. High response rate and low cost of WeChat-based data collection of infant and young child feeding data in rural Qinghai Province, China. J Glob Health. Oct 21, 2022;12:11011. [doi: 10.7189/jogh.12.11011] [Medline: 36269216]
23. Slater K, Schofield PN, Wright J, et al. Talking about diseases: developing a model of patient- and public-prioritized disease phenotypes. NPJ Digit Med. Sep 30, 2024;7(1):263. [doi: 10.1038/s41746-024-01257-8] [Medline: 39349692]
24. Zollo F, Baronchelli A, Betsch C, Delmastro M, Quattrociocchi W. Understanding the complex links between social media and health behaviour. BMJ. Jun 20, 2024;385:e075645. [doi: 10.1136/bmj-2023-075645] [Medline: 38901860]
25. Sousa-Pinto B, Jankin S, Vieira RJ, Marques-Cruz M, Fonseca JA, Bousquet J. English tweets on allergy: content analysis and association with surveillance data. Clin Exp Allergy. Jul 2024;54(7):500-508. [doi: 10.1111/cea.14479] [Medline: 38567657]

26. Hu L, Jin X, Li Y, et al. Perspective of people with type 2 diabetes toward self-management: qualitative study based on web crawler data. J Med Internet Res. Feb 2, 2023;25:e39325. [doi: 10.2196/39325] [Medline: 36729569]
27. Thornton C, Lanyi K, Wilkins G, et al. Scoping the priorities and concerns of parents: infodemiology study of posts on Mumsnet and Reddit. J Med Internet Res. Nov 28, 2023;25:e47849. [doi: 10.2196/47849] [Medline: 38015600]
28. Mc Carthy CM, de Vries R, Mackenbach JD. The influence of unhealthy food and beverage marketing through social media and advergaming on diet-related outcomes in children—a systematic review. Obes Rev. Jun 2022;23(6):e13441. [doi: 10.1111/obr.13441] [Medline: 35301815]
29. Chongqing Municipal Bureau of Statistics and NBS Survey Office in Chongqing. Chongqing statistical yearbook 2025 [Report in Chinese]. China Statistics Press; 2026. URL: https://tjj.cq.gov.cn/zwgk_233/tjnj/tjnj.html?url=https://tjj.cq.gov.cn/cslm/tjsjzl/202603/P020260330529346889373.pdf [Accessed 2026-03-30]
30. Zeng D, Li W, Zhou J, et al. Analysis of the immunoglobulin E molecular sensitization profile in children with allergic asthma and predictive factors for the efficacy of allergy immunotherapy. Ann Transl Med. Nov 2020;8(21):1459. [doi: 10.21037/atm-20-7314] [Medline: 33313204]
31. Zarei F, Dehghani A, Ratansiri A, et al. ChecKAP: a checklist for reporting a knowledge, attitude, and practice (KAP) study. Asian Pac J Cancer Prev. Jul 1, 2024;25(7):2573-2577. [doi: 10.31557/APJCP.2024.25.7.2573] [Medline: 39068593]
32. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. Sep 29, 2004;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
33. Fu J, Yang J, Li Q, et al. What can we learn from a Chinese social media used by glaucoma patients? BMC Ophthalmol. Nov 20, 2023;23(1):470. [doi: 10.1186/s12886-023-03208-5] [Medline: 37986061]
34. Albalawi R, Yeap TH, Benyoucef M. Using topic modeling methods for short-text data: a comparative analysis. Front Artif Intell. 2020;3:42. [doi: 10.3389/frai.2020.00042] [Medline: 33733159]
35. Laureate CDP, Buntine W, Linger H. A systematic review of the use of topic models for short text social media analysis. Artif Intell Rev. May 1, 2023:1-33. [doi: 10.1007/s10462-023-10471-x] [Medline: 37362887]
36. Robert, Delir Haghighi P, Burstein F, Urquhart D, Cicuttini F. Investigating individuals' perceptions regarding the context around the low back pain experience: topic modeling analysis of Twitter data. J Med Internet Res. Dec 23, 2021;23(12):e26093. [doi: 10.2196/26093] [Medline: 36260398]
37. Sievert C, Shirley K. LDAvis: a method for visualizing and interpreting topics. Proc Workshop Interact Lang Learn Vis Interfaces. 2014:63-70. [doi: 10.3115/v1/W14-3110]
38. Chuang J, Manning CD, Heer J. Termite: visualization techniques for assessing textual topic models. In: AVI '12: Proceedings of the International Working Conference on Advanced Visual Interfaces. Association for Computing Machinery (ACM); 2012:74-77. [doi: 10.1145/2254556.2254572]
39. Ma L, Chen R, Ge W, et al. AI-powered topic modeling: comparing LDA and BERTopic in analyzing opioid-related cardiovascular risks in women. Exp Biol Med (Maywood). 2025;250:10389. [doi: 10.3389/ebm.2025.10389] [Medline: 40093658]
40. Castellanos A, Jiang H, Gomes P, Vander Meer D, Castillo A. Large language models for thematic summarization in qualitative health care research: comparative analysis of model and human performance. JMIR AI. Apr 4, 2025;4:e64447. [doi: 10.2196/64447] [Medline: 40611510]
41. Chao AFY, Wang CS, Li BY, Chen HY. From hate to harmony: leveraging large language models for safer speech in times of COVID-19 crisis. Heliyon. Aug 30, 2024;10(16):e35468. [doi: 10.1016/j.heliyon.2024.e35468] [Medline: 39220951]
42. Lediaocha [Article in Chinese]. URL: https://lediaocha.com/ [Accessed 2026-10-09]
43. Fang H, Xian R, Li J, et al. Urban-suburban differences in public perspectives on digitalizing pediatric research: cross-sectional survey study. J Med Internet Res. Jan 7, 2025;27:e60324. [doi: 10.2196/60324] [Medline: 39773676]
44. Fang H, Lv Y, Chen L, Zhang X, Hu Y. The current knowledge, attitudes, and practices of the neglected methodology of web-based questionnaires among Chinese health workers: web-based questionnaire study. J Med Internet Res. Jan 27, 2023;25:e41591. [doi: 10.2196/41591] [Medline: 36533302]
45. Fang H, Chen L, Li J, et al. A web-based instrument for infantile atopic dermatitis identification (electronic version of the modified child eczema questionnaire): development and implementation. J Med Internet Res. Jul 19, 2023;25:e44614. [doi: 10.2196/44614] [Medline: 37467020]
46. Lu Y, Fang J, Tian L, Jin H, editors. Advanced Medical Statistics. 2nd ed. World Scientific Publishing; 2015. ISBN: 9789814583299
47. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in preclinical, clinical, and laboratory studies. Biochem Med (Zagreb). Feb 15, 2021;31(1):010502. [doi: 10.11613/BM.2021.010502] [Medline: 33380887]

48. Arshad SH, Bojarskas J, Tsitoura S, et al. Prevention of sensitization to house dust mite by allergen avoidance in school age children: a randomized controlled study. Clin Exp Allergy. Jun 2002;32(6):843-849. [doi: 10.1046/j.1365-2222.2002.01378.x] [Medline: 12047429]
49. Klain A, Senatore AA, Licari A, et al. The prevention of house dust mite allergies in pediatric asthma. Children (Basel). Apr 15, 2024;11(4):469. [doi: 10.3390/children11040469] [Medline: 38671686]
50. Arshad SH, Bateman B, Matthews SM. Primary prevention of asthma and atopy during childhood by allergen avoidance in infancy: a randomised controlled study. Thorax. Jun 2003;58(6):489-493. [doi: 10.1136/thorax.58.6.489] [Medline: 12775858]
51. Chan-Yeung M, Ferguson A, Watson W, et al. The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. J Allergy Clin Immunol. Jul 2005;116(1):49-55. [doi: 10.1016/j.jaci.2005.03.029] [Medline: 15990772]
52. Schönberger H, Dompeling E, Knottnerus JA, et al. The PREVASC study: the clinical effect of a multifaceted educational intervention to prevent childhood asthma. Eur Respir J. Apr 2005;25(4):660-670. [doi: 10.1183/09031936.05.00067704] [Medline: 15802340]
53. van Schayck OCP, Maas T, Kaper J, Knottnerus AJA, Sheikh A. Is there any role for allergen avoidance in the primary prevention of childhood asthma? J Allergy Clin Immunol. Jun 2007;119(6):1323-1328. [doi: 10.1016/j.jaci.2007.02.024] [Medline: 17399772]
54. Scott M, Roberts G, Kurukulaaratchy RJ, Matthews S, Nove A, Arshad SH. Multifaceted allergen avoidance during infancy reduces asthma during childhood with the effect persisting until age 18 years. Thorax. Dec 2012;67(12):1046-1051. [doi: 10.1136/thoraxjnl-2012-202150] [Medline: 22858926]
55. Gøtzsche PC, Johansen HK. House dust mite control measures for asthma: systematic review. Allergy. Jun 2008;63(6):646-659. [doi: 10.1111/j.1398-9995.2008.01690.x] [Medline: 18445182]
56. Nurmatov U, van Schayck CP, Hurwitz B, Sheikh A. House dust mite avoidance measures for perennial allergic rhinitis: an updated Cochrane systematic review. Allergy. Feb 2012;67(2):158-165. [doi: 10.1111/j.1398-9995.2011.02752.x] [Medline: 22103686]
57. von Hertzen LC, Savolainen J, Hannuksela M, et al. Scientific rationale for the Finnish Allergy Programme 2008-2018: emphasis on prevention and endorsing tolerance. Allergy. May 2009;64(5):678-701. [doi: 10.1111/j.1398-9995.2009.02024.x] [Medline: 19383025]
58. Wang M, Tan G, Eljaszewicz A, et al. Laundry detergents and detergent residue after rinsing directly disrupt tight junction barrier integrity in human bronchial epithelial cells. J Allergy Clin Immunol. May 2019;143(5):1892-1903. [doi: 10.1016/j.jaci.2018.11.016] [Medline: 30500342]
59. Gharpure R, Hunter CM, Schnall AH, et al. Knowledge and practices regarding safe household cleaning and disinfection for COVID-19 prevention - United States, May 2020. MMWR Morb Mortal Wkly Rep. Jun 12, 2020;69(23):705-709. [doi: 10.15585/mmwr.mm6923e2] [Medline: 32525852]
60. Moreira A, Maesano IA, Walusiak-Skorupa J, et al. Healthy planet, healthier people: EAACI's integrated planetary health strategy for asthma, allergies and immune diseases. Allergy. Jan 2026;81(1):7-10. [doi: 10.1111/all.70164] [Medline: 41268612]
61. Erickson S, Lauring B, Cullen J, Medzhitov R. Environmentally driven immune imprinting protects against allergy. Nature. Feb 2026;650(8103):987-996. [doi: 10.1038/s41586-025-10001-5] [Medline: 41606323]
62. Haahtela T, Jantunen J, Saarinen K, et al. Managing the allergy and asthma epidemic in 2020s-lessons from the Finnish experience. Allergy. Aug 2022;77(8):2367-2380. [doi: 10.1111/all.15266] [Medline: 35202479]
63. Merchant RM, South EC, Lurie N. Public health messaging in an era of social media. JAMA. Jan 19, 2021;325(3):223-224. [doi: 10.1001/jama.2020.24514] [Medline: 33393964]
64. Chandrasekaran R, Sadiq T M, Moustakas E. Racial and demographic disparities in susceptibility to health misinformation on social media: national survey-based analysis. J Med Internet Res. Nov 6, 2024;26:e55086. [doi: 10.2196/55086] [Medline: 39504121]
65. John JN, Gorman S, Scales D, Gorman J. Online misleading information about women's reproductive health: a narrative review. J Gen Intern Med. Apr 2025;40(5):1123-1131. [doi: 10.1007/s11606-024-09118-6] [Medline: 39511120]
66. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med. Nov 2019;240:112552. [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
67. Epstein Z, Sirlin N, Arechar A, Pennycook G, Rand D. The social media context interferes with truth discernment. Sci Adv. Mar 3, 2023;9(9):eabo6169. [doi: 10.1126/sciadv.abo6169] [Medline: 36867704]
68. Pérez-Escamilla R, Tomori C, Hernández-Cordero S, et al. Breastfeeding: crucially important, but increasingly challenged in a market-driven world. Lancet. Feb 11, 2023;401(10375):472-485. [doi: 10.1016/S0140-6736(22)01932-8] [Medline: 36764313]

69. Rick JW, Morgan A, De DR, Hsiao JL, Shi VY. Allergens and marketing claims of the most popular baby moisturizers: a product analysis. J Am Acad Dermatol. Jan 2023;88(1):205-207. [doi: 10.1016/j.jaad.2022.04.034] [Medline: 35487334]
70. Cetthakrikul N, Kelly M, Banwell C, Baker P, Smith J. Regulation of baby food marketing in Thailand: a NetCode analysis. Public Health Nutr. Oct 2022;25(10):2680-2692. [doi: 10.1017/S1368980022001446] [Medline: 35733357]
71. Brady WJ, Jackson JC, Lindström B, Crockett MJ. Algorithm-mediated social learning in online social networks. Trends Cogn Sci. Oct 2023;27(10):947-960. [doi: 10.1016/j.tics.2023.06.008] [Medline: 37543440]
72. Sadlier MA. Smartphones and children: algorithm driven social media should be the focus of action. BMJ. May 15, 2025;389:r964. [doi: 10.1136/bmj.r964] [Medline: 40374289]
73. Khadija U, Vohra-Miller S, Jeimy S. Allergy misinformation across social media platforms: content analysis and public response. J Allergy Clin Immunol Pract. Apr 2026;14(4):952-958. [doi: 10.1016/j.jaip.2025.12.032] [Medline: 41661117]
74. Conway AE, Gupta E, Verdi M, et al. A media advocacy toolkit for the allergist-immunologist. J Allergy Clin Immunol Pract. Oct 2024;12(10):2678-2686. [doi: 10.1016/j.jaip.2024.07.003] [Medline: 38996838]

Abbreviations

ARIA: Allergic Rhinitis and its Impact on Asthma
ChecKAP: Checklist for Reporting Items for Knowledge, Attitude, and Practice
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
GINA: Global Initiative for Asthma
HDM: house dust mite
IgE: immunoglobulin E
KAP: knowledge, attitude, and practice
LDA: latent Dirichlet allocation
OR: odds ratio

Edited by Amaryllis Mavragani; peer-reviewed by Arjavi Patel, Ayman Allam; submitted 06.May.2026; final revised version received 17.Sep.2026; accepted 20.Sep.2026; published 16.Nov.2026

Please cite as:
Yang P, Xian R, Li J, Zhao J, Xie H, Jiang Z, Hu Y, Xu X, Zang N, Zhao Y, Fang H, Liu E
A Social Media Mining and Questionnaire Survey Hybrid Approach to Characterizing Caregiver Knowledge, Attitudes, and Practices on House Dust Mite Control: Cross-Sectional Study
J Med Internet Res 2026;28:e100446
URL: https://www.jmir.org/2026/1/e100446
doi: 10.2196/100446

© Peixin Yang, Ruoling Xian, Juan Li, Jing Zhao, Hangjiang Xie, Zhenghong Jiang, Yan Hu, Ximing Xu, Na Zang, Yan Zhao, Heping Fang, Enmei Liu. Originally published in the Journal of Medical Internet Research (https://www.jmir.org), 16.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on https://www.jmir.org/, as well as this copyright and license information must be included.