Selective citation in the literature on the hygiene hypothesis

Supplement

Text S1. Protocol deviations

Text S2. Data extraction

Text S3. References of included literature on Strachan's hygiene hypothesis

Table S1. All characteristics of the publications in the hygiene hypothesis network.

Table S2. Top 6 of articles and authors within network.

Table S3. Odds ratios for the chance of *empirical* publications to be cited within full network.

Table S4. Odds ratios for the chance of *empirical* publications to be cited by *synthesis* publications.

Table S5. Sensitivity analyses on odds ratios for the chance of being cited, without the four most cited publications.

Table S6. Sensitivity analyses on odds ratios for the chance of being cited, with a 1-year time lag between cited and citing publication.

Table S7. Sensitivity analyses on odds ratios for the chance of being cited, without citing publications with less than 10 potential citation paths.

Text S1. Protocol deviations

Search strategy. We ran some checks of our original search strategy (with '*hygiene hypothesis'* in combination with the pre-specified health outcomes). These checks indicated that many relevant publications in the period 1990 – 2000 were missed. Going through these missed publications we found out that the name for this hygiene hypothesis was not yet accepted or used. Also, we realised that almost all relevant publications within this network referred to Strachan's original study from 1989. We decided to change the search strategy, into *all publications referring to Strachan's original article.* Additionally, we limited the output to those publications that mentioned 'rhinitis' or a similar term in the title, abstract or keywords (see inclusion criteria).

Inclusion criteria (regarding the health outcome and types of exposure). Originally, all publications with either health outcome *rhinitis* or *asthma* were to be included. This yielded too many publications in our network, hence we decided to include only publications on rhinitis (or hay fever, or rhinoconjunctivitis). Inclusion solely based on asthma would also have yielded a network that was too large. Also, rhinitis is the original outcome as studied by Strachan. Similarly, we included only publications with exposures *number of siblings* and *infection history.* These are the two most important types of exposure related to the hygiene hypothesis as originally stated by Strachan. Number of siblings was originally studied by Strachan, and infections during childhood (or during pregnancy) was his explanation for the relationship between number of siblings and hay fever.

Types of health outcome. We intended to score both *asthma* and *rhinitis* as health outcomes. However, as we included only publications that studied the relationship with rhinitis, and excluded publications that were solely on asthma, we decided to focus on rhinitis. Thus, for the

empirical publications, we scored the relationships Siblings – Rhinitis and Infections – Rhinitis. (We also scored asthma, dermatitis and atopic sensitisation, but only for the sensitivity analyses.) Synthesis publications often did not differentiate between the different allergies in their general conclusion, so for the synthesis publications we scored statements on Siblings – Allergy and Infection – Allergy.

Types of exposure. We focused on two exposures: *number of siblings* (or household size or sibling order) and *history of infections* (as assessed by parental questionnaire, serology or medical records). We used this variable also in our analyses; it was scored as a) number of siblings only, b) only infection history, and c) both number of siblings and infection history.

Study outcome scoring strategy. There were many empirical publications with contradictory results, especially in the case of Infections where the results seemed to depend on the type of infection. In order to deal with this, we decided to use the *authors' conclusion* on Siblings – Rhinitis and Infections – Rhinitis as leading. We used 5 categories: 0. not measured or reported; 1. effect in line with hygiene hypothesis (inverse relationship); 2. no relationship; 3. effect contrary to hygiene hypothesis (positive relationship); 4. mixed or unclear results. Synthesis publications were scored in a similar way, but then on Siblings – Allergy and Infections – Allergy.

If no clear authors' conclusion was stated in the empirical publications, we used the data that were presented in the tables or the text and scored as follows: 1. statistically significant inverse relationship; 2. no statistically significant relationship; 3. statistically significant positive relationship; 4. mixed or unclear results. If both adjusted analyses and crude analyses were presented we preferred the adjusted ones. There is one exception: adjustment for Infections in

the relationship Siblings – Rhinitis; after all, the hygiene hypothesis states that infection is the mediator between siblings and rhinitis.

In the analyses we used one combined measure for study outcome, with three levels: 1) supportive; 2) mixed results / unclear; 3) non-supportive. Publications were scored as supportive if the exposure or exposures showed an inverse association with rhinitis. Publications were scored as non-supportive if there was no association or a positive association of the exposure with rhinitis. Publications were scored as mixed if the exposure or exposures showed mixed or unclear results, or if two exposures were investigated, and one showed an inverse association and the other showed no or a positive association.

General conclusion about the hygiene hypothesis. While studying the literature on the hygiene hypothesis we realised that many related hypotheses reside under this name, all evolved from one another. *The* hygiene hypothesis clearly does not exist. Support for one version of the hypothesis often implied the refutation of another (older) version. As authors assumed different versions, their general conclusion on the hygiene hypothesis would not be compatible. We decided to not score this general conclusion.

Specificity. We used 3 outcome categories for publication's specificity instead of 5.

Study design. Ecological studies were excluded from the network. Cohort studies were further classified as retrospective and prospective cohort studies. This latter step was data-driven because we realised during the analysis that prospective cohort studies were cited less often. During exploration of the data, we noticed a big difference in the citation behaviour of retrospective and prospective cohorts studies. We therefore decided to amend our preregistered data analysis plan (http://hdl.handle.net/10411/ZKGGOG). We differentiated between these

research designs (by assigning them a different categorical value), and take this difference into account in our analyses with study design as determinant or as covariate.

Publisher. In addition to the protocol we also scored the publisher of the journal, based on the information in Journal Citation Reports and in Web of Science. This could be interesting for publisher self-citation concordance analysis.

Explained variance. In addition to the original analysis plan in the protocol, we also calculated the explained variance of the adjusted models, so that these models are easier to compare. For this purpose we calculated McFadden's R^2 by the following formula: $R^2 = 1 - {LL_M}/{LL_0}$ in which LL_M stands for the log likelihood of the current regression model and LL_0 stands for the log likelihood of the empty random-regression model. Both the current and the empty model (without predictors) were nested under the citing publication. Because of missing values for certain determinants (such as sample size), some models could be tested only on a sub-selection of citation paths. If this was the case, then LL_0 was calculated on the same sub-selection of citation paths.

Text S2. Data extraction

Most variables are described in the main document. Here follows some additional information for some of the variables.

Study outcome was scored as follows: 1. supportive of the hygiene hypothesis; 2. mixed or unclear results; 3. non-supportive of the hygiene hypothesis. An inverse relationship between past exposure and rhinitis is considered to be supportive for the hygiene hypothesis, while a neutral or positive relationship was scored as non-supportive. The scoring was based on the authors' interpretation of the results, as it was stated in the text of the publication. If the authors' interpretation was unclear, we scored study outcome based on the direction and statistical significance of the data. Non-empirical publications seldom distinguished between allergy subtypes, so we used the stated conclusion on general allergy as outcome measure.

Exposure could be either number of siblings (or order of siblings or household size) or history of bacterial or viral infection (as assessed by parental questionnaire, serology, or medical records). If the impact of number of both siblings and infection history was assessed and they were contradicting each other (with one exposure showing inverse association, the other a neutral or positive association), then study outcome was scored as mixed.

Gender of the corresponding author was assessed by first name, with help of www.genderchecker.com; if first name was not given, other articles of the same author were searched, and the profile of the author at the university or at www.researchgate.com was checked.

We performed a validity check of the gender assessment on a random sample of 20 publications. We checked our original assessment (described above) against the results from another gender assessment tool: Gender-API.com. This tool takes into account the person's country while assessing gender based on his or her first name Additionally, it gives an accuracy

score for each assessment. The results can be found in the Table below. Our reference assessment reached the same results as the original one, with a 100% accuracy.

Table. Validity check for gender assessment on random sample (N=20).

ID [1]	First Name of Corresponding Author	Country of Corresponding Author	Genderchecker Assessment	Our Data-extraction Assessment	Gender-API Assessment [2] (validity check)	Accuracy of Gender-API Assessment [2]
1	David	UK	male	male	male	99%
5	Barbara	UK	female	female	female	98%
13	Nick [3]	UK	male	male	male	98%
14	Sarah [3]	UK	female	female	female	98%
18	Juha	Finland	unisex	male [4]	male	100%
24	Anthony	UK	male	male	male	99%
25	Erika	Germany	female	female	female	98%
26	Mustafa	Turkey	male	male	male	100%
31	Johannes	Germany	male	male	male	99%
35	Paolo	Italy	male	male	male	99%
73	Anne-Louise [5]	Australia	no match	female	female	100%
	Anne		unisex			
	Louise		female			
75	Keiko	Japan	female	female	female	99%
78	Aarif	Turkey	male	male	male	75%
79	Sharad	India	male	male	male	100%
87	Woei Kang [5]	Singapore	no match	male		
	Woei		male		male	67%
	Kang		unisex		male	88%
94	Jonathan	USA	male	male	male	99%
95	Ahmet	Turkey	male	male	male	100%
97	Chun-Yuh [5]	Taiwan	no match	male [4]		
	Chun		unisex		male	53%
	Yuh		no match		male	60%
103	David	UK	male	male	male	99%
109	Katherine	USA	female	female	female	99%

Notes. 1. See Text S3 for the references. 2. Based on combination of first name and country. 3. First name was not stated in publication, but retrieved via ResearchGate.net; match between profile and correpsonding author based on surname, initials, affiliation and research topic. 4. Web search revealed a man with same name and afflilation. 5. Composite names that could not be assessed as a whole were assessed by its composites.

Time to citation was the number of years between the publication date of the cited publication and the submission date of the citing publication. This variable was not used as determinant of citation, but to determine the dataset of potential citation paths: only citation paths with a positive value for *time to citation* were considered a potential citation, and only potential citations were included in our dataset.

As *publication date* we used either the online publication date or the paper publication date, whichever was first. The average duration from submission to publication was nine months in this network. For 57 publications the *submission date* was not stated. In these cases, it was estimated by subtracting nine months from the publication date.

Text S3. References of included literature on Strachan's hygiene hypothesis

1. Strachan DP. Epidemiology of hay-fever - towards a community diagnosis. Clin Exp Allergy. 1995;25(4):296-303.

2. Strachan D. Socioeconomic factors and the development of allergy. Toxicol Lett. 1996;86(2-3):199-203.

3. Strachan DP, Taylor EM, Carpenter RG. Family structure, neonatal infection, and hay fever in adolescence. Arch Dis Child. 1996;74(5):422-6.

4. Burr ML, Merrett TG, Dunstan FDJ, Maguire MJ. The development of allergy in high-risk children. Clin Exp Allergy. 1997;27(11):1247-53.

5. Butland BK, Strachan DP, Lewis S, Bynner J, Butler N, Britton J. Investigation into the increase in hay fever and eczema at age 16 observed between the 1958 and 1970 British birth cohorts. Br Med J. 1997;315(7110):717-21.

6. Jarvis D, Chinn S, Luczynska C, Burney P. The association of family size with atopy and atopic disease. Clin Exp Allergy. 1997;27(3):240-5.

7. Matricardi PM, Rosmini F, Ferrigno L, Nisini R, Rapicetta M, Chionne P, et al. Cross sectional retrospective study of prevalence of atopy among Italian military students with antibodies against hepatitis A virus. Br Med J. 1997;314(7086):999-1003.

8. Rasanen M, Laitinen T, Kaprio J, Koskenvuo M, Laitinen LA. Hay fever, asthma and number of older siblings - a twin study. Clin Exp Allergy. 1997;27(5):515-8.

9. Serafini U. Do infections protect against asthma and atopy? Allergy. 1997;52(9):955-7.

10. Strachan DP. Allergy and family size: A riddle worth solving. Clin Exp Allergy. 1997;27(3):235-6.

11. Bodner C, Godden D, Seaton A, Aberdeen WG. Family size, childhood infections and atopic diseases. Thorax. 1998;53(1):28-32.

12. Farooqi IS, Hopkin JM. Early childhood infection and atopic disorder. Thorax. 1998;53(11):927-32.

13. Jones NS, Carney AS, Davis A. The prevalence of allergic rhinosinusitis: A review. J Laryngol Otol. 1998;112(11):1019-30.

14. Lewis SA, Britton JR. Consistent effects of high socioeconomic status and low birth order, and the modifying effect of maternal smoking on the risk of allergic disease during childhood. Respir Med. 1998;92(10):1237-44.

15. Ponsonby AL, Couper D, Dwyer T, Carmichael A. Cross sectional study of the relation between sibling number and asthma, hay fever, and eczema. Arch Dis Child. 1998;79(4):328-33.

16. von Mutius E. The influence of birth order on the expression of atopy in families: A gene-environment interaction? Clin Exp Allergy. 1998;28(12):1454-6.

17. Palmer LJ, Valinsky IJ, Pikora T, Zubrick SR, Landau LI. Environmental factors and asthma and allergy in schoolchildren from Western Australia. Eur Resp J. 1999;14(6):1351-7.

18. Pekkanen J, Remes S, Kajosaari M, Husman T, Soininen L. Infections in early childhood and risk of atopic disease. Acta Paediatr. 1999;88(7):710-4.

19. Ponsonby AL, Couper D, Dwyer T, Carmichael A, Kemp A. Relationship between early life respiratory illness, family size over time, and the development of asthma and hay fever: A seven year follow up study. Thorax. 1999;54(8):664-9.

20. Strachan DP. Lifestyle and atopy. Lancet. 1999;353(9163):1457-8.

21. Xu B, Jarvelin MR, Pekkanen J. Prenatal factors and occurrence of rhinitis and eczema among offspring. Allergy. 1999;54(8):829-36.

22. Matricardi PM, Rosmini F, Riondino S, Fortini M, Ferrigno L, Rapicetta M, et al. Exposure to foodborne and orofecal microbes versus airborne viruses in relation to atopy and allergic asthma: Epidemiological study. Br Med J. 2000;320(7232):412-7.

23. Paunio M, Heinonen OP, Virtanen M, Leinikki P, Patja A, Peltola H. Measles history and atopic diseases - a population-based cross-sectional study. JAMA-J Am Med Assoc. 2000;283(3):343-6.

24. Seaton A, Devereux G. Diet, infection and wheezy illness: Lessons from adults. Pediatr Allergy Immunol. 2000;11:37-40.

25. von Mutius E. The environmental predictors of allergic disease. J Allergy Clin Immunol. 2000;105(1):9-19.

26. Yilmaz M, Bingol G, Altintas D, Kendirli SG. Correlation between atopic diseases and tuberculin responses. Allergy. 2000;55(7):664-7.

27. Doull IJM. Does pregnancy prevent atopy? Clin Exp Allergy. 2001;31(9):1335-7.

28. Johnston SL, Openshaw PJM. The protective effect of childhood infections - the next challenge is to mimic safely this protection against allergy and asthma. Br Med J. 2001;322(7283):376-7.

29. Martinez FD. The coming-of-age of the hygiene hypothesis. Respir Res. 2001;2(3):129-32.

30. McKeever TM, Lewis SA, Smith C, Collins J, Heatlie H, Frischer M, et al. Siblings, multiple births, and the incidence of allergic disease: A birth cohort study using the West Midlands General Practice Research Database. Thorax. 2001;56(10):758-62.

31. Ring J, Kramer U, Schafer T, Behrendt H. Why are allergies increasing? Curr Opin Immunol. 2001;13(6):701-8.

32. von Mutius E. Infection: Friend or foe in the development of atopy and asthma? The epidemiological evidence. Eur Resp J. 2001;18(5):872-81.

33. Crane J. Asthma and allergic diseases: Is there a downside to cleanliness and can we exploit it? Eur J Clin Nutr. 2002;56:S39-S43.

34. Karmaus W, Botezan C. Does a higher number of siblings protect against the development of allergy and asthma? A review. J Epidemiol Community Health. 2002;56(3):209-17.

35. Matricardi PM, Rosmini F, Panetta V, Ferrigno L, Bonini S. Hay fever and asthma in relation to markers of infection in the United States. J Allergy Clin Immunol. 2002;110(3):381-7.

36. McKeever TM, Lewis SA, Smith C, Collins J, Heatlie H, Frischer M, et al. Early exposure to infections and antibiotics and the incidence of allergic disease: A birth cohort study with the West Midlands General Practice Research Database. J Allergy Clin Immunol. 2002;109(1):43-50.

37. McKeever TM, Lewis SA, Smith C, Hubbard R. The importance of prenatal exposures on the development of allergic disease - a birth cohort study using the West Midlands General Practice Database. Am J Respir Crit Care Med. 2002;166(6):827-32.

38. Svanes C, Jarvis D, Chinn S, Omenaas E, Gulsvik A, Burney P, et al. Early exposure to children in family and day care as related to adult asthma and hay fever: Results from the European Community Respiratory Health Survey. Thorax. 2002;57(11):945-50.

39. Cremonini F, Gasbarrini A. Atopy, Helicobacter pylori and the hygiene hypothesis. Eur J Gastroenterol Hepatol. 2003;15(6):635-6.

40. McCune A, Lane A, Murray L, Harvey I, Nair P, Donovan J, et al. Reduced risk of atopic disorders in adults with Helicobacter pylori infection. Eur J Gastroenterol Hepatol. 2003;15(6):637-40.

41. Prescott SL. Allergy: The price we pay for cleaner living? Ann Allergy Asthma Immunol. 2003;90(6):64-70.

42. Gustafsson D, Andersson K. Effect of indoor environmental factors on development of atopic symptoms in children followed up to 4 years of age. Paediatr Perinat Epidemiol. 2004;18(1):17-25.

43. Jones N. Allergic rhinitis: Aetiology, predisposing and risk factors. Rhinology. 2004;42(2):49-56.

44. Resch A, Schlipkoter U, Crispin A, Behrendt H, Heinrich J, Wichmann HE, et al. Atopic disease and its determinants - a focus on the potential role of childhood infection. Clin Exp Allergy. 2004;34(8):1184-91.

45. von Hertzen LC, Haahtela T. Asthma and atopy - the price of affluence? Allergy. 2004;59(2):124-37.

46. Borchers AT, Keen CL, Gershwin ME. Hope for the hygiene hypothesis: When the dirt hits the fan. J Asthma. 2005;42(4):225-47.

47. Bresciani M, Parisi C, Manghi G, Bonini S. The hygiene hypothesis: Does it function worldwide? Curr Opin Allergy Clin Immunol. 2005;5(2):147-51.

48. Bufford JD, Gem JE. The hygiene hypothesis revisited. Immunol Allerg Clin North Am. 2005;25(2):247-+.

49. de Meer G, Janssen NAH, Brunekreef B. Early childhood environment related to microbial exposure and the occurrence of atopic disease at school age. Allergy. 2005;60(5):619-25.

50. Karmaus W, Johnson CC. Invited commentary: Sibship effects and a call for a comparative disease approach. Am J Epidemiol. 2005;162(2):133-8.

51. Kopp MV, Semmler S, Horst G, Berner R, Forster J. Hospital admission with neonatal sepsis and development of atopic disease: Is there a link? Pediatr Allergy Immunol. 2005;16(8):630-6.

52. Nafstad P, Brunekreef B, Skrondal A, Nystad W. Early respiratory infections, asthma, and allergy: 10-year follow-up of the Oslo Birth Cohort. Pediatrics. 2005;116(2):E255-E62.

53. Westergaard T, Rostgaard K, Wohlfahrt J, Andersen PK, Aaby P, Melbye M. Sibship characteristics and risk of allergic rhinitis and asthma. Am J Epidemiol. 2005;162(2):125-32.

54. Arbes SJ, Sever ML, Vaughn B, Cohen EA, Zeldin DC. Oral pathogens and allergic disease: Results from the third national health and nutrition examination survey. J Allergy Clin Immunol. 2006;118(5):1169-75.

55. Balemans WAF, Rovers MM, Schilder AGM, Sanders EAM, Kimpen JLL, Zielhuis GA, et al. Recurrent childhood upper respiratory tract infections do not reduce the risk of adult atopic disease. Clin Exp Allergy. 2006;36(2):198-203.

56. Bernsen RMD, van der Wouden JC, Nagelkerke NJD, de Jongste JC. Early life circumstances and atopic disorders in childhood. Clin Exp Allergy. 2006;36(7):858-65.

57. Biagini JM, LeMasters GK, Ryan PH, Levin L, Reponen T, Bernstein DI, et al. Environmental risk factors of rhinitis in early infancy. Pediatr Allergy Immunol. 2006;17(4):278-84.

58. Friedrich N, Volzke H, Schwahn C, Kramer A, Junger M, Schafer T, et al. Inverse association between periodontitis and respiratory allergies. Clin Exp Allergy. 2006;36(4):495-502.

59. Kinra S, Smith GD, Jeffreys M, Gunnell D, Galobardes B, McCarron P. Association between sibship size and allergic diseases in the Glasgow Alumni Study. Thorax. 2006;61(1):48-53.

60. Kocabas E, Yapicioglu H, Yildizdas D, Kendirli SG, Burgut R. The prevalence of atopy in children with antibodies against hepatitis A virus and hepatitis B virus. Turk J Pediatr. 2006;48(3):189-96.

61. Paunio M, Peltola H, Virtanen M, Leinikki P, Makela A, Heinonen OP. Acute infections, infection pressure, and atopy. Clin Exp Allergy. 2006;36(5):634-9.

62. Zutavern A, von Klot S, Gehring U, Krauss-Etschmann S, Heinrich J. Pre-natal and post-natal exposure to respiratory infection and atopic diseases development: A historical cohort study. Respir Res. 2006;7:8.

63. Cakir M, Karakas T, Orhan F, Okten A, Gedik Y. Atopy in children with chronic hepatitis B virus infection. Acta Paediatr. 2007;96(9):1343-6.

64. Cetinkaya F, Uslu HS, Nuhoglu A. Effect of neonatal sepsis on the development of allergies and asthma in later childhood. Int Arch Allergy Immunol. 2007;142(2):145-50.

65. Dunder T, Tapiainen T, Pokka T, Uhari M. Infections in child day care centers and later development of asthma, allergic rhinitis, and atopic dermatitis - prospective follow-up survey 12 years after controlled randomized hygiene intervention. Arch Pediatr Adolesc Med. 2007;161(10):972-7.

66. Harris JM, Mills P, White C, Moffat S, Taylor AJN, Cullinan P. Recorded infections and antibiotics in early life: Associations with allergy in UK children and their parents. Thorax. 2007;62(7):631-7.

67. von Mutius E. Allergies, infections and the hygiene hypothesis - the epidemiological evidence. Immunobiology. 2007;212(6):433-9.

68. Baccioglu A, Kalpaklioglu F, Guliter S, Yakaryilmaz F. Helicobacter pylori in allergic inflammation - fact or fiction? Allergol Immunopath. 2008;36(2):85-9.

69. Blaser MJ, Chen Y, Reibman J. Does helicobacter pylori protect against asthma and allergy? Gut. 2008;57(5):561-7.

70. Bremner SA, Carey IM, DeWilde S, Richards N, Maier WC, Hilton SR, et al. Infections presenting for clinical care in early life and later risk of hay fever in two UK birth cohorts. Allergy. 2008;63(3):274-83.

71. Chen Y, Blaser MJ. Helicobacter pylori colonization is inversely associated with childhood asthma. J Infect Dis. 2008;198(4):553-60.

72. Perzanowski MS, Canfield SM, Chew GL, Mellins RB, Hoepner LA, Jacobson JS, et al. Birth order, atopy, and symptoms of allergy and asthma among inner-city children attending head start in New York City. Clin Exp Allergy. 2008;38(6):968-76.

73. Ponsonby AL, Kemp A. Investigation of the hygiene hypothesis: Current issues and future directions. Allergy. 2008;63(5):506-8.

74. Svanes C. What has the ecrhs told us about the childhood risks of asthma, allergy and lung function? Clin Respir J. 2008;2:34-44.

75. Tanaka K, Miyake Y, Arakawa M, Sasaki S, Ohya Y. Dental caries and allergic disorders in Japanese children: The Ryukyus Child Health Study. J Asthma. 2008;45(9):795-9.

76. Torres-Borrego J, Molina-Teran AB, Montes-Mendoza C. Prevalence and associated factors of allergic rhinitis and atopic dermatitis in children. Allergol Immunopath. 2008;36(2):90-100.

77. Vargas C, Bustos P, Diaz PV, Amigo H, Rona RJ. Childhood environment and atopic conditions, with emphasis on asthma in a chilean agricultural area. J Asthma. 2008;45(1):73-8.

78. Eifan AO, Akkoc T, Ozdemir C, Bahceciler NN, Barlan IB. No association between tuberculin skin test and atopy in a bacillus Calmette-Guerin vaccinated birth cohort. Pediatr Allergy Immunol. 2009;20(6):545-50.

79. Gangal SV, Chowgule R. Infections in early life and susceptibility to allergic diseases: Relevance of hygiene hypothesis. Curr Sci. 2009;96(6):784-93.

80. Kendirli SG, Yilmaz M, Bayram I, Altintas DU, Inal A, Karakoc G. Potential association between allergic diseases and pertussis infection in schoolchildren: Results of two cross-sectional studies seven years apart. Allergol Immunopath. 2009;37(1):21-5.

81. Kramer MS, Matush L, Bogdanovich N, Dahhou M, Platt RW, Mazer B. The low prevalence of allergic disease in Eastern Europe. Clin Exp Allergy. 2009;39(5):708-16.

82. Matheson MC, Walters EH, Simpson JA, Wharton CL, Ponsonby AL, Johns DP, et al. Relevance of the hygiene hypothesis to early vs. late onset allergic rhinitis. Clin Exp Allergy. 2009;39(3):370-8.

83. Ohfuji S, Miyake Y, Arakawa M, Tanaka K, Sasaki S. Sibship size and prevalence of allergic disorders in Japan: The Ryukyus Child Health Study. Pediatr Allergy Immunol. 2009;20(4):377-84.

84. Bunyavanich S, Soto-Quiros ME, Avila L, Laskey D, Senter JM, Celedon JC. Risk factors for allergic rhinitis in Costa Rican children with asthma. Allergy. 2010;65(2):256-63.

85. Codispoti CD, Levin L, LeMasters GK, Ryan P, Reponen T, Villareal M, et al. Breast-feeding, aeroallergen sensitization, and environmental exposures during infancy are determinants of childhood allergic rhinitis. J Allergy Clin Immunol. 2010;125(5):1054-60.

86. Arbes SJ, Matsui EC. Can oral pathogens influence allergic disease? J Allergy Clin Immunol. 2011;127(5):1119-27.

87. Liew WK, Lim CWT, Tan TH, Wong KY, Tai BC, Quek SC, et al. The effect of Kawasaki disease on childhood allergies - a sibling control study. Pediatr Allergy Immunol. 2011;22(5):488-93.

88. Matheson MC, Dharmage SC, Abramson MJ, Walters EH, Sunyer J, de Marco R, et al. Early-life risk factors and incidence of rhinitis: Results from the European community respiratory health study-an international population-based cohort study. J Allergy Clin Immunol. 2011;128(4):816-+.

89. Flohr C, Nagel G, Weinmayr G, Kleiner A, Williams HC, Ait-Khaled N, et al. Tuberculosis, bacillus Calmette-Guerin vaccination, and allergic disease: Findings from the International Study of Asthma and Allergies in Childhood Phase Two. Pediatr Allergy Immunol. 2012;23(4):324-31.

90. Lee SY, Kwon JW, Seo JH, Song YH, Kim BJ, Yu J, et al. Prevalence of atopy and allergic diseases in Korean children: Associations with a farming environment and rural lifestyle. Int Arch Allergy Immunol. 2012;158(2):168-74.

91. Lighter-Fisher J, Peng CH. Infection with mycobacterium tuberculosis is inversely associated with childhood asthma. Pediatr Allergy Immunol Pulmonol. 2012;25(2):80-5.

92. Nagel G, Weinmayr G, Flohr C, Kleiner A, Strachan DP, Grp IPTS. Association of pertussis and measles infections and immunizations with asthma and allergic sensitization in ISAAC Phase Two. Pediatr Allergy Immunol. 2012;23(8):737-46.

93. Schmitz R, Atzpodien K, Schlaud M. Prevalence and risk factors of atopic diseases in German children and adolescents - a nationwide health report. Pediatr Allergy Immunol. 2012;23(8):716-23.

94. Silverberg JI, Kleiman E, Silverberg NB, Durkin HG, Joks R, Smith-Norowitz TA. Chickenpox in childhood is associated with decreased atopic disorders, IgE, allergic sensitization, and leukocyte subsets. Pediatr Allergy Immunol. 2012;23(1):50-8.

95. Duksal F, Akcay A, Becerir T, Ergin A, Becerir C, Guler N. Rising trend of allergic rhinitis prevalence among Turkish schoolchildren. Int J Pediatr Otorhinolaryngol. 2013;77(9):1434-9.

96. Kondrashova A, Seiskari T, Ilonen J, Knip M, Hyoty H. The 'hygiene hypothesis' and the sharp gradient in the incidence of autoimmune and allergic diseases between Russian Karelia and Finland. Apmis. 2013;121(6):478-93.

97. Kuo HC, Chang WC, Yang KD, Yu HR, Wang CL, Ho SC, et al. Kawasaki disease and subsequent risk of allergic diseases: A population-based matched cohort study. BMC Pediatr. 2013;13:6.

98. Prokopakis E, Vardouniotis A, Kawauchi H, Scadding G, Georgalas C, Hellings P, et al. The pathophysiology of the hygiene hypothesis. Int J Pediatr Otorhinolaryngol. 2013;77(7):1065-71.

99. Lionetti E, Leonardi S, Lanzafame A, Garozzo MT, Filippelli M, Tomarchio S, et al. Helicobacter pylori infection and atopic diseases: Is there a relationship? A systematic review and meta-analysis. World J Gastroenterol. 2014;20(46):17635-47.

100. Tantilipikorn P. The relationship between allergic rhinitis and viral infections. Curr Opin Otolaryngol Head Neck Surg. 2014;22(3):249-52.

101. Kimura H, Konno S, Isada A, Maeda Y, Musashi M, Nishimura M. Contrasting associations of body mass index and measles with asthma and rhinitis in young adults. Allergy Asthma Proc. 2015;36(4):293-9.

102. Lin CH, Lin WC, Wang YC, Lin IC, Kao CH. Association between neonatal urinary tract infection and risk of childhood allergic rhinitis. Medicine (Baltimore). 2015;94(38):6.

103. Strachan DP, Ait-Khaled N, Foliaki S, Mallol J, Odhiambo J, Pearce N, et al. Siblings, asthma, rhinoconjunctivitis and eczema: A worldwide perspective from the international study of asthma and allergies in childhood. Clin Exp Allergy. 2015;45(1):126-36.

104. Han CS. A specific hygiene hypothesis. Med Hypotheses. 2016;93:146-9.

105. Indinnimeo L, Porta D, Forastiere F, De Vittori V, De Castro G, Zicari AM, et al. Prevalence and risk factors for atopic disease in a population of preschool children in Rome: Challenges to early intervention. Int J Immunopathol Pharmacol. 2016;29(2):308-19.

106. Sastra S, Irsa L, Loebis MS, Evalina R. Number of siblings and allergic rhinitis in children. Paediatr Indones. 2016;56(1):1-7.

107. Christiansen ES, Kjaer HF, Eller E, Bindslev-Jensen C, Host A, Mortz CG, et al. Early childhood risk factors for rhinoconjunctivitis in adolescence: A prospective birth cohort study. Clin Transl Allergy. 2017;7:7.

108. Miftahussurur M, Nusi IA, Graham DY, Yamaoka Y. Helicobacter, hygiene, atopy, and asthma. Front Microbiol. 2017;8:11.

109. Wander K, Shell-Duncan B, Brindle E, O'Connor K. Hay fever, asthma, and eczema and early infectious diseases among children in Kilimanjaro, Tanzania. Am J Hum Biol. 2017;29(3):13.

110. Yildizdas HY, Ozcan A, Sertdemir Y, Yilmaz M. Effect of healthcare associated infections and broad spectrum antibiotic use in newborn period on development of asthma, allergic rhinitis and atopic dermatitis in early childhood. Cukurova Med J. 2017;42(1):132-9.

Table S1. All characteristics of the publications in the hygiene hypothesis network.

		N publications	n potential citations	n actual citations (%)
Total		110	5551	392 (7%)

Publication characteristics - content-related	**category**	**N publications**	**n potential citations**	**n actual citations (%)**
Type of Exposure	only Number of Siblings	28	1512	100 (7%)
	only Infection History	48	1946	144 (7%)
	both Siblings & Infections	34	2093	148 (7%)
Study Outcome	supportive	41	2322	198 (9%)
Exposure - Rhinitis	mixed results	35	1913	129 (7%)
	non-supportive	34	1316	65 (5%)
Publication Type / Study Design	Empirical	73	3517	337 (10%)
	cross-sectional	39	1697	179 (11%)
	case-control	4	249	36 (14%)
	cohort	29	1535	121 (8%)
	-retrospective	15	817	89 (11%)
	-prospective	14	718	32 (4%)
	intervention	1	36	1 (3%)
	Synthesis	37	2034	55 (3%)
	narrative review	27	1423	16 (1%)
	systematic review	2	80	20 (25%)
	-with meta-analysis	1	8	1 (13%)
	editorial, etc	8	531	19 (4%)
Sample Size	low (1 – 999)	24	909	56 (6%)
(cat; for empirical publications)	medium (1000 – 7999)	25	1327	143 (11%)
	high (>= 8000)	24	1281	138 (11%)
Specificity	0 (non-specific)	27	1402	25 (2%)
	1	39	1657	65 (4%)
	2 (specific)	44	2492	302 (12%)

Publication characteristics - not content-related	**category**	**N publications**	**n potential citations**	**n actual citations (%)**
Conclusive Title	not conclusive	99	5026	375 (7%)
	conclusive	11	525	17 (3%)

Funding Source	non-profit	44	2188	214 (10%)
	for-profit	1	38	1 (3%)
	both	12	559	51 (9%)
	not reported / unclear	53	2766	126 (5%)
Number of Authors	1 - 2	32	2017	89 (4%)
	3 - 5	41	2143	155 (7%)
	>= 6	37	1391	148 (11%)
Number of Affiliations	1	36	2276	111 (5%)
	2	24	1168	108 (9%)
	>= 3	50	2107	173 (8%)
Number of References	< 30	35	2307	194 (8%)
	30 – 50	49	2060	159 (8%)
	>= 50	26	1184	39 (3%)

Journal characteristics	**category**	**N publications**	**n potential citations**	**n actual citations (%)**
Journal Impact Factor (cat)	0 - 2	28	1275	27 (2%)
	2 - 4	41	2087	145 (7%)
	>= 4	32	1671	176 (11%)
Publisher	Wiley-Blackwell	41	2107	82 (4%)
	BMJ	15	1170	213 (18%)
	Elsevier	18	894	43 (5%)
	other	36	1380	54 (4%)

Author characteristics	category	N publications	n potential citations	n actual citations (%)
Gender	male	65	3368	265 (8%)
	female	42	2024	123 (6%)
	unclear	3	159	4 (3%)
Affiliation	university	88	4402	258 (6%)
	government	9	410	22 (5%)
	industry / other	13	739	112 (15%)
Country	Europe	62	3903	324 (8%)
	UK	26	1946	165 (8%)
	Germany	11	594	19 (3%)
	Finland	8	516	33 (6%)
	Italy	7	418	85 (20%)
	North-America	19	688	38 (6%)
	USA	18	662	36 (5%)
	Asia	21	484	9 (2%)
	Turkey	9	303	7 (2%)
	Japan	4	60	0 (0%)
	Australia / New Zealand	8	476	21 (4%)
	Australia	7	407	21 (5%)

Citation characteristics	category	n potential citations	n actual citations (%)
Authority	low (0-2)	2279	81 (4%)
	medium (2-10)	1326	108 (8%)
	high (>= 10)	1946	203 (10%)
Time to Citation (in years)	0 – 1	494	38 (8%)
	1 – 2	521	56 (11%)
	2 – 3	527	50 (9%)
	3 – 4	459	33 (7%)
	4 – 5	456	40 (9%)
	5 – 6	441	35 (8%)
	6 – 7	404	28 (7%)
	7 – 8	372	22 (6%)
	=> 8	1877	90 (5%)
Self-citation	no	5462	365 (7%)
	yes	89	27 (30%)

Table S2. Top 6 of articles (above) and authors (below) within network, based on the number of received citations up to 2017.

Article rank	Article's first author	Title	Year	Nr. of received citations (% of potential citations)
1	Matricardi	Cross sectional retrospective study of prevalence of atopy among Italian military students with antibodies against hepatitis A virus	1997	35 (35 %)
2	Bodner	Family size, childhood infections and atopic diseases	1998	32 (33 %)
3	Matricardi	Exposure to foodborne and orofecal microbes versus airborne viruses in relation to atopy and allergic asthma: epidemiological study	2000	32 (38 %)
4	Strachan	Family structure, neonatal infection, and hay fever in adolescence	1996	28 (26 %)
5	Farooqi	Early childhood infection and atopic disorder	1998	21 (23 %)
6	Karmaus	Does a higher number of siblings protect against the development of allergy and asthma? A review	2002	19 (26 %)

Author rank	Author	Affiliation	Country	Nr. of received citations (= authority)
1	P. Matricardi	Consiglio Nazionale delle Ricerche, Rome	Italy	84
2	F. Rosmini	Istituto Superiore di Sanita, Rome	Italy	84
3	L. Ferrigno	Istituto Superiore di Sanita, Rome	Italy	84
4	M. Rapicetta	Istituto Superiore di Sanita, Rome	Italy	67
5	D. Strachan	University of London, London	United Kingdom	57
6	S. Bonini	Consiglio Nazionale delle Ricerche, Rome	Italy	49

Table S3. Odds ratios (95% CIs) for the chance of *empirical* publications to be cited within full network (N = 73, n = 3517).

Publication characteristics, content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Type of Exposure (ref: both Siblings & Infections)			0.04
only Number of Siblings	0.5 (0.3 – 0.6)	0.4 (0.3 – 0.5)	
only Infection History	1.0 (0.8 – 1.3)	0.7 (0.5 – 0.9)	
Study Outcome (ref: non-supportive results) **			0.07
mixed / unclear results	1.4 (0.9 – 2.3)	1.1 (0.7 – 1.9)	
supportive results	4.8 (3.2 – 7.0)	5.1 (3.3 – 7.8)	

Publication characteristics, other content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Study Design (ref: cross-sectional)			0.02 (crude)
case control	1.4 (0.9 – 2.2)		
retrospective cohort	0.9 (0.7 – 1.2)		
prospective cohort	0.3 (0.2 – 0.5)		
Sample Size (ref: low)		***	0.02
medium	1.6 (1.2 – 2.3)	1.6 (1.0 – 2.6)	
high	1.6 (1.2 – 2.3)	1.9 (1.2 – 3.0)	
Specificity (ref: low)			0.05
medium	4.6 (2.6 – 8.2)	3.5 (1.8 – 6.5)	
high	7.4 (4.5 – 12)	6.1 (3.5 – 10)	

Publication characteristics, not content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Conclusive Title (yes vs no)	0.3 (0.2 – 0.6)	0.3 (0.1 – 0.5)	0.03
Funding Source (ref: exclusively non-profit)			0.02
profit or both profit/non-profit	0.9 (0.6 – 1.2)	0.8 (0.5 – 1.1)	
not reported	0.8 (0.6 – 1.1)	0.8 (0.6 – 1.1)	
Number of Authors (ref: 1-2)			0.03
3 - 5	1.1 (0.7 – 1.6)	1.2 (0.7 – 1.8)	
>= 6	1.7 (1.2 – 2.6)	1.8 (1.1 – 2.9)	
Number of Affiliations (ref: 1)			0.03
2	1.7 (1.2 – 2.4)	1.8 (1.2 – 2.6)	
>= 3	1.3 (0.9 – 1.7)	1.6 (1.2 – 2.4)	
Number of References (ref: <30)			0.02
30 - 50	0.9 (0.7 – 1.1)	1.0 (0.7 – 1.2)	
>= 50	0.3 (0.1 – 0.7)	0.3 (0.1 – 0.8)	

Journal characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Journal Impact Factor (ref: 0-2, n = 3266)			0.06
2 – 4	2.8 (1.7 – 4.5)	2.6 (1.6 – 4.3)	
>= 4	5.9 (3.7 – 9.5)	6.6 (4.0 – 11)	

Author characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Gender (female vs male, n = 3457)	0.6 (0.5 – 0.8)	0.7 (0.5 – 0.9)	0.03
Region (ref: Europe)			0.04
North-America	0.5 (0.3–0.97)	0.6 (0.3 – 1.1)	
Asia	0.2 (0.1 – 0.4)	0.1 (0.1 – 0.3)	
Australia / New-Zealand	0.5 (0.3 – 0.8)	0.5 (0.3 – 0.9)	
Type of Affiliation (other vs university)	2.4 (1.9 – 3.2)	2.2 (1.7 – 2.9)	0.04

Citation characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Authority (ref: low)			0.05
medium	2.0 (1.4 – 3.0)	2.1 (1.4 – 3.0)	
high	3.6 (2.6 – 5.1)	3.8 (2.7 – 5.5)	

***** adjusted for study design and log sample size. ****** both the 'crude' and adjusted analyses are (additionally) adjusted for type of exposure. ******* only adjusted for study design. **supportive:** supportive for Strachan's original hygiene hypothesis, i.e. inverse association between siblings/infections and allergy. **non-supportive:** no association or positive association between siblings/infections and allergy. N = number of potentially cited publications; n = number of potential citation paths.

Table S4. Odds ratios (95% CIs) for the chance of *empirical* publications to be cited by *synthesis* publications (N = 73, n = 1097).

Publication characteristics, content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Type of Exposure (ref: both Siblings & Infections)			0.07
only Number of Siblings	0.4 (0.2 – 0.7)	0.3 (0.2 – 0.6)	
only Infection History	1.8 (1.1 – 3.0)	1.3 (0.7 – 2.3)	
Study Outcome (ref: non-supportive results) **			0.12
mixed / unclear results	0.5 (0.2 – 1.0)	0.4 (0.2 – 0.9)	
supportive results	6.0 (3.1 – 12)	7.3 (3.5 – 15)	

Publication characteristics, other content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Study Design (ref: cross-sectional)			0.04 (crude)
case control	2.1 (1.1 – 4.2)		
retrospective cohort	1.0 (0.6 – 1.6)		
prospective cohort	0.3 (0.1 – 0.6)		
Sample Size (ref: low)		***	0.04
medium	1.6 (0.9 – 2.8)	1.9 (0.8 – 4.6)	
high	1.3 (0.7 – 2.3)	1.9 (0.8 – 4.4)	
Specificity (ref: low)			0.06
medium	6.9 (2.8 – 17)	4.3 (1.6 – 12)	
high	7.6 (3.4 – 17)	5.6 (2.3 – 13)	

Publication characteristics, not content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Conclusive Title (yes vs no)	0.4 (0.2 – 0.9)	0.2 (0.1 – 0.6)	0.05
Funding Source (ref: exclusively non-profit)			0.04
profit or both profit/non-profit	0.7 (0.4 – 1.2)	0.7 (0.4 – 1.3)	
not reported	0.6 (0.4 – 1.0)	0.5 (0.3 – 0.9)	
Number of Authors (ref: 1-2)			0.06
3 - 5	0.6 (0.3 – 1.1)	0.6 (0.3 – 1.2)	
>= 6	1.5 (0.8 – 2.8)	1.5 (0.7 – 3.2)	
Number of Affiliations (ref: 1)			0.04
2	1.3 (0.7 – 2.4)	1.6 (0.8 – 3.1)	
>= 3	1.2 (0.7 – 2.0)	2.0 (1.1 – 3.6)	
Number of References (ref: <30)			0.04
30 - 50	0.7 (0.5 – 1.1)	0.8 (0.5 – 1.2)	
>= 50	0.2 (0.04–0.9)	0.2 (0.04–1.0)	

Journal characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Journal Impact Factor (ref: 0-2, n = 1015)			0.11
2 – 4	2.6 (1.2 – 5.6)	2.2 (1.0 – 4.9)	
>= 4	8.2 (3.8 – 18)	9.2 (4.1 – 21)	

Author characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Gender (female vs male, n = 1079)	0.5 (0.3 – 0.8)	0.6 (0.4 – 0.9)	0.04
Region (ref: Europe)			0.07
North-America	0.7 (0.2 – 2.2)	0.9 (0.3 – 2.9)	
Asia	0.2 (0.1 – 0.7)	0.1 (0.02–0.3)	
Australia / New-Zealand	0.2 (0.1 – 0.6)	0.2 (0.9 – 0.7)	
Type of Affiliation (other vs university)	3.3 (2.2 – 5.1)	3.1 (2.0 – 4.8)	0.07

Citation characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Authority (ref: low)			0.07
medium	2.6 (1.4 – 4.6)	2.7 (1.4 – 5.1)	
high	4.0 (2.3 – 7.1)	4.1 (2.2 – 7.6)	

***** adjusted for study design and log sample size. ****** both the 'crude' and adjusted analyses are (additionally) adjusted for type of exposure. ******* only adjusted for study design. **supportive:** supportive for Strachan's original hygiene hypothesis, i.e. inverse association between siblings/infections and allergy. **non-supportive:** no association or positive association between siblings/infections and allergy. N = number of potentially cited (*empirical*) publications; n = number of potential citation paths.

Table S5. Sensitivity analyses on odds ratios (95% CIs) for the chance of being cited within full network, without the four most cited publications. N = 106, n = 5164).

Publication characteristics, content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Type of Exposure (ref: both Siblings & Infections)			0.08
only Number of Siblings	1.4 (1.0 – 1.9)	0.9 (0.6 – 1.3)	
only Infection History	1.0 (0.8 – 1.4)	1.0 (0.7 – 1.4)	
Study Outcome (ref: non-supportive results) **			0.08
mixed / unclear results	0.8 (0.5 – 1.1)	1.2 (0.8 – 2.0)	
supportive results	1.0 (0.7 – 1.5)	1.5 (1.0 – 2.2)	

Publication characteristics, other content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Publication Type (empirical vs synthesis)	2.9 (2.1 – 3.9)		0.03 (crude)
Study Design (ref: cross-sectional)			0.08 (crude)
case control	0.3 (0.1 – 0.8)		
retrospective cohort	1.1 (0.8 – 1.5)		
prospective cohort	0.5 (0.3 – 0.8)		
narrative review	0.1 (0.1 – 0.2)		
systematic review	4.5 (2.5 – 8.0)		
editorial / other	0.4 (0.2 – 0.7)		
Sample Size (ref: low, n = 3130)		***	0.02
medium	2.3 (1.4 – 3.7)	1.6 (0.9 – 2.9)	
high	3.3 (2.1 – 5.3)	2.5 (1.5 – 4.2)	
Specificity (ref: low)			0.09
medium	2.4 (1.5 – 3.9)	3.3 (1.9 – 5.6)	
high	5.6 (3.6 – 8.7)	3.5 (2.1 – 5.6)	

Publication characteristics, not content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Conclusive Title (yes vs no)	0.6 (0.4 – 1.0)	0.5 (0.3 – 0.9)	0.08
Funding Source (ref: exclusively non-profit)			0.08
profit or both profit/non-profit	1.4 (1.0 – 2.0)	1.4 (0.9 – 2.0)	
not reported	0.5 (0.4 – 0.7)	1.0 (0.7 – 1.4)	
Number of Authors (ref: 1-2)			0.08
3 - 5	1.2 (0.9 – 1.7)	0.7 (0.5 – 1.1)	
>= 6	2.0 (1.4 – 2.8)	0.9 (0.6 – 1.4)	
Number of Affiliations (ref: 1)			0.08
2	1.8 (1.3 – 2.5)	1.1 (0.8 – 1.6)	
>= 3	1.0 (0.7 – 1.4)	0.6 (0.4 – 0.9)	

Number of References (ref: <30)			0.08
30 - 50	0.9 (0.7 – 1.2)	0.7 (0.6 – 1.0)	
>= 50	0.6 (0.4 – 0.8)	0.8 (0.5 – 1.3)	

Journal characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Journal Impact Factor (ref: 0-2, n = 4752)			0.09
2 – 4	2.8 (1.8 – 4.3)	2.5 (1.6 – 3.9)	
>= 4	4.0 (2.6 – 6.1)	3.5 (2.2 – 5.5)	

Author characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Gender (female vs male, n = 5005)	1.4 (1.1 – 1.8)	1.2 (0.9 – 1.6)	0.07
Region (ref: Europe)			0.08
North-America	1.2 (0.8 – 1.8)	1.0 (0.6 – 1.6)	
Asia	0.4 (0.2 – 0.8)	0.4 (0.2 – 0.8)	
Australia / New-Zealand	0.8 (0.5 – 1.2)	0.7 (0.5 – 1.2)	
Type of Affiliation (other vs university)	1.6 (1.2 – 2.2)	1.4 (1.1 – 2.0)	0.08

Citation characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Authority (ref: low)			0.08
medium	2.1 (1.5 – 2.9)	1.6 (1.1 – 2.2)	
high	2.2 (1.6 – 3.0)	1.6 (1.1 – 2.2)	

In these sensitivity analyses, the four most cited publications shown in Table S2 are excluded as *cited* publications; they are still included as *citing* publications. * adjusted for study design and log sample size. ** both the 'crude' and adjusted analyses are (additionally) adjusted for type of exposure. *** only adjusted for study design. **supportive:** supportive for Strachan's original hygiene hypothesis, i.e. inverse association between siblings/infections and allergy. **non-supportive:** no association or positive association between siblings/infections and allergy. N = number of potentially cited publications; n = number of potential citation paths.

Table S6. Sensitivity analyses on odds ratios (95% CIs) for the chance of being cited within full network, with a 1-year time lag between cited and citing publication. (N = 110, n = 5057).

Publication characteristics, content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Type of Exposure (ref: both Siblings & Infections)			0.11
only Number of Siblings	0.7 (0.5 – 1.0)	0.4 (0.3 – 0.7)	
only Infection History	1.2 (1.0 – 1.6)	0.8 (0.6 – 1.1)	
Study Outcome (ref: non-supportive results) **			0.12
mixed / unclear results	1.2 (0.8 – 1.7)	2.2 (1.4 – 3.5)	
supportive results	1.7 (1.2 – 2.4)	3.2 (2.2 – 4.6)	

Publication characteristics, other content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Publication Type (empirical vs synthesis)	4.4 (3.2 – 6.0)		0.05 (crude)
Study Design (ref: cross-sectional)			0.09 (crude)
case control	1.4 (0.9 – 2.2)		
retrospective cohort	0.9 (0.7 – 1.2)		
prospective cohort	0.3 (0.2 – 0.5)		
narrative review	0.1 (0.0 – 0.1)		
systematic review	3.4 (1.9 – 6.2)		
editorial / other	0.3 (0.1 – 0.4)		
Sample Size (ref: low, n = 3199)			0.02
medium	1.6 (1.1 – 2.2)	1.4 (0.8 – 2.4)	
high	1.4 (1.0 – 2.1)	1.6 (0.9 – 2.6)	
Specificity (ref: low)			0.12
medium	2.8 (1.6 – 4.7)	3.1 (1.7 – 5.6)	
high	10.3 (6.4 – 17)	6.1 (3.6 – 10)	

Publication characteristics, not content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Conclusive Title (yes vs no)	0.5 (0.3 – 0.8)	0.3 (0.2 – 0.5)	0.10
Funding Source (ref: exclusively non-profit)			0.10
profit or both profit/non-profit	0.8 (0.6 – 1.2)	0.8 (0.5 – 1.1)	
not reported	0.4 (0.3 – 0.5)	0.8 (0.6 – 1.1)	
Number of Authors (ref: 1-2)			0.10
3 - 5	2.0 (1.5 – 2.8)	1.1 (0.7 – 1.8)	
>= 6	3.7 (2.7 – 5.0)	1.7 (1.0 – 2.6)	
Number of Affiliations (ref: 1)			0.10
2	3.0 (2.2 – 4.1)	2.1 (1.4 – 3.0)	
>= 3	2.2 (1.7 – 2.9)	1.7 (1.2 – 2.3)	
Number of References (ref: <30)			0.09
30 - 50	1.1 (0.9 – 1.4)	1.0 (0.8 – 1.3)	
>= 50	0.5 (0.3 – 0.7)	0.7 (0.4 – 1.3)	

Journal characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Journal Impact Factor (ref: 0-2, n = 4580)			0.12
2 – 4	3.2 (2.0 – 4.9)	2.5 (1.6 – 4.0)	
>= 4	5.4 (3.5 – 8.4)	4.5 (2.9 – 7.1)	

Author characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Gender (female vs male, n = 4913)	0.8 (0.6 – 1.0)	0.7 (0.6 – 1.0)	0.09
Region (ref: Europe)			0.11
North-America	1.0 (0.7 – 1.5)	1.0 (0.6 – 1.8)	
Asia	0.3 (0.2 – 0.7)	0.2 (0.1 – 0.4)	
Australia / New-Zealand	0.5 (0.3 – 0.9)	0.5 (0.3 – 0.9)	
Type of Affiliation (other vs university)	2.3 (1.8 – 2.9)	1.9 (1.5 – 2.5)	0.10

Citation characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Authority (ref: low)			0.11
medium	2.8 (1.9 – 3.9)	1.9 (1.3 – 2.7)	
high	4.0 (2.9 – 5.5)	2.8 (2.0 – 4.0)	

***** adjusted for study design. ****** both the 'crude' and adjusted analyses are (additionally) adjusted for type of exposure. **supportive:** supportive for Strachan's original hygiene hypothesis, i.e. inverse association between siblings/infections and allergy. **non-supportive:** no association or positive association between siblings/infections and allergy. **N**: number of publications. **n**: number of potential citation paths.

Table S7. Sensitivity analyses on odds ratios (95% CIs) for the chance of being cited, without citing publications with less than 10 potential citation paths (N = 110, n = 5507).

Publication characteristics, content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Type of Exposure (ref: both Siblings & Infections)			0.10
only Number of Siblings	0.8 (0.6 – 1.1)	0.5 (0.4 – 0.7)	
only Infection History	1.2 (1.0 – 1.6)	0.8 (0.6 – 1.1)	
Study Outcome (ref: non-supportive results) **			0.12
mixed / unclear results	1.2 (0.8 – 1.7)	2.3 (1.5 – 3.6)	
supportive results	1.7 (1.3 – 2.4)	3.0 (2.1 – 4.2)	

Publication characteristics, other content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Publication Type (empirical vs synthesis)	4.2 (3.1 – 5.6)		0.04 (crude)
Study Design (ref: cross-sectional)			0.09 (crude)
case control	1.4 (1.0 – 2.2)		
retrospective cohort	0.8 (0.6 – 1.1)		
prospective cohort	0.4 (0.2 – 0.5)		
narrative review	0.1 (0.0 – 0.1)		
systematic review	3.3 (1.8 – 5.8)		
editorial / other	0.3 (0.2 – 0.4)		
Sample Size (ref: low, n = 3423)			0.02
medium	1.5 (1.1 – 2.1)	1.5 (0.9 – 2.5)	
high	1.4 (1.0 – 2.0)	1.7 (1.1 – 2.8)	
Specificity (ref: low)			0.11
medium	2.2 (1.4 – 3.6)	2.5 (1.4 – 4.2)	
high	8.6 (5.6 – 13)	4.9 (3.1 – 7.9)	

Publication characteristics, not content-related	**Crude OR**	**Adjusted OR ***	**R^2 ***
Conclusive Title (yes vs no)	0.4 (0.3 – 0.7)	0.3 (0.2 – 0.5)	0.10
Funding Source (ref: exclusively non-profit)			0.09
profit or both profit/non-profit	0.9 (0.6 – 1.3)	0.8 (0.6 – 1.2)	
not reported	0.4 (0.3 – 0.5)	0.7 (0.5 – 1.0)	
Number of Authors (ref: 1-2)			0.09
3 - 5	2.0 (1.5 – 2.7)	1.1 (0.7 – 1.7)	
>= 6	3.8 (2.8 – 5.2)	1.7 (1.1 – 2.6)	
Number of Affiliations (ref: 1)			0.09
2	2.8 (2.1 – 3.7)	1.9 (1.3 – 2.6)	
>= 3	2.2 (1.7 – 2.9)	1.6 (1.2 – 2.3)	
Number of References (ref: <30)			0.09
30 - 50	1.0 (0.8 – 1.3)	0.9 (0.7 – 1.2)	
>= 50	0.4 (0.3 – 0.6)	0.8 (0.4 – 1.3)	

Journal characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Journal Impact Factor (ref: 0-2, n = 4955)			0.11
2 – 4	3.2 (2.1 – 5.0)	2.5 (1.6 – 3.9)	
>= 4	5.7 (3.8 – 8.8)	4.6 (3.0 – 7.2)	

Author characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Gender (female vs male, n = 5350)	0.8 (0.6 – 1.0)	0.8 (0.6 – 1.0)	0.09
Region (ref: Europe)			0.11
North-America	0.9 (0.6 – 1.3)	0.9 (0.5 – 1.4)	
Asia	0.3 (0.2 – 0.6)	0.2 (0.1 – 0.4)	
Australia / New-Zealand	0.5 (0.3 – 0.8)	0.5 (0.3 – 0.8)	
Type of Affiliation (other vs university)	2.4 (1.9 – 3.0)	2.0 (1.6 – 2.6)	0.10

Citation characteristics	**Crude OR**	**Adjusted OR ***	**R^2 ***
Authority (ref: low)			0.11
medium	2.4 (1.8 – 3.3)	1.8 (1.3 – 2.5)	
high	3.8 (2.9 – 5.1)	2.9 (2.2 – 4.0)	

***** adjusted for study design. ****** both the 'crude' and adjusted analyses are (additionally) adjusted for type of exposure. **supportive:** supportive for Strachan's original hygiene hypothesis, i.e. inverse association between siblings/infections and allergy. **non-supportive:** no association or positive association between siblings/infections and allergy. **N**: number of publications. **n**: number of potential citation paths.