Commentary

Addressing epidemiological and public health analytic challenges in outcome and impact research: a commentary on 'Prechewing Infant Food, Consumption of Sweets and Dairy and Not Breastfeeding are Associated with Increased Diarrhea Risk of Ten Month Old Infants'

Jean-Pierre Habicht and Gretel H. Peltó

Division of Nutritional Sciences, Cornell University, Ithaca, New York, USA

Abstract

Based on a paper by Conkle *et al* 2016, in which the authors use a descriptive epidemiological design to examine the relationship of premastication and other dietary behavioral variables to childhood diarrhea in the US, we address larger issues of "plausible causality" and the challenges involved in moving from epidemiological studies to public health policy. Drawing on examples from breastfeeding research and water, sanitation and hygiene (WASH) research, we discuss the following propositions: 1. Effective outcome analyses require simultaneous investigation of different, even contradictory, pathways; 2. Outcome versus impact assessments require different analytic procedures including context analysis; 3. Impact analysis requires understanding the trade-offs between detrimental and beneficial outcomes in relation to potential interventions; 4. No estimates exist for the likely detrimental and beneficial impacts of banning premastication, much less for their trade-offs.

Correspondence: Gretel H. Peltó, Division of Nutritional Sciences, Cornell University, Ithaca, NY 14850, USA. E-mail: gp32@cornell.edu

The publication of the paper by Conkle *et al.* (2016) in this issue of MCN is an important and timely contribution to the public health literature on infant feeding practices and diarrhoeal disease. It is important because it calls attention to diarrhoeal disease and nutrition, which continues to be a significant, but an often-neglected issue in wealthy countries. It is timely because it appears at a time when the relationships between nutrition, and water, sanitation and hygiene (Humphrey *et al.* 2015), and its consequences, are once again taking centre stage in research on infants' and young children's well-being. This study focuses attention on the role of behavioural-dietary aspects of diarrhoea–nutrition relationships.

Although the emphasis in Conkle *et al.* (2016) is on 'the association between prechewing and diarrhoea', the paper also shows statistically significant positive associations of diarrhoea with ingesting 'sweets' and

'dairy foods' and a negative association with 'not breastfeeding'. It is possible, even probable, that premastication is not a direct cause of diarrhoea, any more than is 'sweets' or 'dairy'. However, whether the associations identified in the paper are causal or not, one can safely conclude that they reflect a behavioural-dietary complex that is associated with higher risk of diarrhoea. Better understanding of this complex of dietary behaviours and diarrhoea, as well as their determinants and consequences, is the first step to identifying interventions to address the reasons for the associations. For example, if mothers who premasticate live in more contaminated environments, premastication is a good marker of such environments. From a policy perspective, moving to stamp out premastication, without affecting the environment, would not reduce diarrhoea in infants. We (Peltó *et al.* 2010) obviously have a stake in the argument on

premastication. If, as we suspect, it plays a valuable role in the development of infant immune systems, protects infants from illness, reduces the severity of the infections they are exposed to in their households and plays an important nutritional role in environments where caregivers do not have adequate access to appropriate complementary foods, the move to 'stamp it out' is ill considered and dangerous.

In the following paragraphs, we take up three issues that are raised by the paper by Conkle *et al.* (2016):

- i What is driving the present public health focus on premastication as a negative practice that causes child disease?
- ii How should outcome analysis of cross-sectional data be approached?
- iii Why is it important to address the differences between 'outcomes' and 'impact' in research that is undertaken to inform social or public health policy and actions?

(i) What is driving the present public health focus on premastication as a negative practice that causes child disease?

Previous to the agricultural revolution, premastication was the only way that children could be fed before they had a full set of teeth and after breast milk was no longer adequate (Pelto *et al.* 2010). This practice has continued in many societies and was even found among urban elites in China until commercial baby foods became available recently. It is also continued in the United States, a recent survey found that 19% of mothers with high school education or less

premasticated, and in HIV clinics, 36% of caretakers reported premastication at least weekly (Rakhmanina *et al.* 2011).

The literature review by Conkle *et al.* (2016) samples the current state of discussions about the negative aspects of premastication. These discussions reflect a common belief that exchanging saliva, with all its microbes, from a mother to an infant, is not only disgusting but is also unhealthy. The alternative view is that premastication must be good because human infants have been evolutionarily programmed to be fed this way. This alternative perspective, and the evidence for it, is not as fully covered by the authors, compared with their attention to the negative aspects. For example, they do not describe the evidence of caries prevention from mother to child transmission of oral microflora (Aaltonen & Tenovuo 1994) or the protective association of saliva against allergy risk (Hesselmar *et al.* 2013).

Inevitably, the social environment is an important determinant of the choice of a research topic. This resonance also affects our approach to how we investigate a topic and our interpretation of the findings. Different emotional resonances bias our conclusions in opposite directions, even with the same evidence. The idea of exchanging saliva through premastication is apparently as abhorrent to many people today as was breastfeeding a few generations ago. Oddly, for many, this abhorrence apparently does not extend to adult tongue kissing, which also involves the exchange of saliva. Undoubtedly, there are some who find French kissing (called Florentine kissing in French) morally wrong and dangerous. The literature on the harms and benefit of French kissing and other transfers of saliva seems to be as rudimentary as that on premastication.

Given the role of the social environment in all research, and particularly on research on sensitive or controversial topics, it is important to employ careful

Key messages

- Effective outcome analyses require simultaneous investigation of different, even contradictory, pathways.
- Outcome versus impact assessments require different analytic procedures including context analysis.
- Impact analysis requires understanding the trade-offs between detrimental and beneficial outcomes in relation to potential interventions.
- No estimates exist for the likely detrimental and beneficial impacts of banning premastication, much less for their trade-offs.

analytic procedures both at the stage of 'outcome analysis' and at the stage of 'impact assessment'.

(ii) How should outcome analysis of cross-sectional data be approached?

The analysis presented by Conkle *et al.* (2016) presents descriptive bivariate analyses of the associations between various dietary behaviours and diarrhoea, and a multivariate analysis of the feeding variables and demographic and socio-economic variables. Only one outcome – the relationship between premastication and diarrhoea – is discussed in any detail. The authors state the following: 'this study ... does not firmly establish a causal relationship between prechewing and diarrhoea', and they mention the possibilities of confounding and of reverse causality, but without further examination. The paper then discusses the potential implications of their study for eliminating premastication.

We set the following discussion on analysis within the larger context of decisions to 'take action'. Decisions about taking actions should rest on a convincing 'causal explanation' (c.f. Popper 1963) or a 'story' of causality (Habicht *et al.* 1999), which is captured by the concept of 'plausible causality'. The idea of 'plausible causality', as a prerequisite for action, appears to be universal in human societies. It is as true when the explanatory model involves the actions of supernatural beings as it is for models that are based on biological and social processes. In reality in daily life, we rarely examine the evidence of causality for all of our dietary, health and political actions. We generally rely on authorities to ascertain causality for us.

To be plausible, authorities must be able to demonstrate processes and attributes that are regarded as legitimate by groups of believers, in this case, by scientists. Scientists believe in causality because of natural, but not supernatural, agents and laws. Another belief, however, which is becoming more wide spread, is that science cannot prove causality. This means that today's truth is tomorrow's mistaken belief (Kuhn 2012). This is an old philosophical stance, further elaborated by Popper (1963) and developed by epidemiologists (e.g. Krieger 1994; Susser 1991). Scientists have rules about how to challenge causality, even though they can never

prove it. The truthfulness of the outcome is determined by how well it has survived these challenges on causality, but truth is always provisional. This view has two consequences. The first is that we should be humble in offering recommendations for action. The second is that all recommendations, including global public health recommendations, such as those issued by the World Health Organization (e.g. Habicht 2004) should identify gaps that need to be investigated to further test the plausibility and appropriateness of the recommendations. Often, the cautions and gaps identified by scientists in the context of making recommendations are forgotten in the process of moving them into policy, and even more seriously, not followed up by research that might contradict or modify the recommendations (Habicht 2004).

In addition to establishing the evidence for plausibility, a major source of misunderstanding between scientists and others is the degree of plausibility that is required for action. All of the papers suggesting that premastication causes diarrhoea, or the suggestion that premastication is healthy, present their own stories. New evidence, such as the recent study identifying healthy consequences of exchanging saliva through pacifiers (Hesselmar *et al.* 2013), bolsters the story arguing for a positive health effect. However, the need to appeal to single studies shows how much work needs to be done, not only to elicit cause–effect evidence but also to bolster the stories for and against the health consequences of premastication.

These considerations show that plausibility (persuasiveness) is socially constructed. The recognition of this fundamental feature of scientific inquiry is precisely why epidemiologists are wary of its importance as a criterion of causality (Hill 1965; Porta 2008). But we cannot act responsibly without sound plausibility, whatever other criteria are fulfilled.

Epidemiological designs in order of greater likelihood of demonstrating plausible causality (c.f. Hill 1965; Susser 1991) can be described as follows: descriptive (Porta 2008), exploratory (Bender & Lange 2001), analytic (Porta 2008), confirmatory (Bender & Lange 2001) and experimental (Porta 2008). Within each of these designs, one should follow Rothman's admonition for data analysis: 'Causal inference based on conjecture and refutation fosters a highly desirable critical

scrutiny' (Rothman 2002). In other words, the art of interpreting the meaning of cross-sectional descriptive data is detective work: imagining scenarios and testing them. The Conkle paper is descriptive in that it examined many possible variables in cross-sectional data. It found some that are positively related to diarrhoea, but the authors only present those that remained statistically significant at $P < 0.05$ in the multivariate analyses. This study is also suited for exploratory analyses, in which statistical significance is less important than are leads for developing the story. For instance, the intriguing discovery that infants who eat premasticated foods have a greater-than-expected increase in diarrhoea when 'sweets' are added merits follow-up, even though it is not statistically significant.

Three different detective stories, all of which are adumbrated by Conkle *et al.* (2016), can explain the Conkle *et al.* findings. Their main story is that premastication causes diarrhoea. The literature on environmental causes of diarrhoea is relevant for this story because they show that the influences of environmental factors as 'casual factors' for diarrhoea are different in different kinds of environments (Eisenberg *et al.* 2007; Esrey & Habicht 1986). For example, previous research has shown that eliminating pathogens in water has little effect if there is not enough water in the environment. In these situations, increasing the amount of water (clean or not) has little effect when the environment is heavily polluted from faeces, which have not been disposed of sanitarily. Thus, clean water has no effect unless three conditions are present: clean water, enough water, use of the water and sequestration of faeces. One would expect similar factors would affect the impact of premastication. Analyses that fail to take account of the differential effects of environmental conditions and miss these facts concluded that interventions to improve water and sanitation are ineffective (Engell & Lim 2013). This is the opposite conclusion than is the story that takes these facts into account (Eisenberg *et al.* 2007).

Finding different effects across different circumstances, as has been found for breastfeeding, water and sanitation, would bolster or weaken the Conkle story depending on the findings. Their present story is less convincing because they did not investigate other possibilities, including, for example, their finding of

more diarrhoea than expected in children who were fed both premasticated foods and sweets. Their story is also less convincing because they do not offer a rationale for concentrating only on premastication. Is there any reason to believe that premastication causes diarrhoea, whereas eating sweets and dairy foods does not?

A second story, which is the one we propose, is that that the statistical significance of premastication and diarrhoea occurs because diarrhoea is an indicator that identifies a combination of environments and foods that are contaminated with diarrhoea pathogens. This seems likely in view of the fact that some of the subgroups in the United States who premasticate, such as those who attend HIV clinics (Rakhmanina *et al.* 2011), are disadvantaged. Thus, both the environment and the premastication of the already contaminated foods are the causes of the diarrhoea, but eliminating premastication would not decrease food pollution. Evidence to reject this possibility would bolster the Conkle *et al.* story. Another possibility is that premastication protects those who are living in unsanitary environments, as is the case with breastfeeding (Habicht *et al.* 1988), but the protection is insufficient to overcome the combined environmental and food exposures.

A third possible story relates to the analytic problem of reverse causality. In this story, premastication does not cause diarrhoea, but diarrhoea encourages mothers to premasticate because they believe it is good for children with diarrhoea. The analogy is the association that occurs between breastfeeding and poor growth in infants and toddlers because mothers breastfeed children who are not growing well longer, and not because breastfeeding stunts growth (Habicht 2000). Reverse causality for the association between breastfeeding and poor growth only became apparent from a combination of ethnographic and longitudinal data, whereby the ethnographic data have been the most convincing. Therefore, similar investigations about why mothers premasticate is essential.

None of the three stories are plausible enough in their outcomes to mobilize for action yet. However, these stories should be investigated simultaneously by ensuring that the data and analyses within any design are appropriate for investigating their stories. This simultaneous examination is an effective strategy to achieve Popperian refutability (falsifiability)

(Popper 1963) by contrasting the stories' plausibility in relation to each other.

(iii) Why is it important to address the differences between 'outcomes' and 'impact' in research that is undertaken to inform social or public health policy and actions?

When one moves from a plausible story about an outcome to implications for public health action, the analytic challenge shifts from consideration of what is required to establish outcome to what is required to establish impact. It is useful to separate the examination of impact in the population where the outcome analyses have been conducted from outcomes in other populations. The first is important when the goal is policy for a specific population, and the second is important when the goal is a global policy. For both purposes, further analytic steps are necessary.

To examine the procedures for moving from outcome to impact, we need to draw examples for which plausible outcomes have already been developed. Because the plausibility studies on premastication are not yet credible, we draw our examples from the water and sanitation literature, which is appropriate because we believe that premastication's effect on diarrhoea will be similar to those of improving or degrading water, sanitation and breastfeeding.

Population impact is the change in an outcome due to a change in the cause. The outcome analyses used to develop the causal stories do not predict population impact. However, the outcome analyses do identify characteristics of children who are more or less likely to respond to changes in the cause. Impacts can then be estimated in groups of children who are more and less likely to benefit. The data for these impact analyses may be the same as those analysed to develop the outcome story, if the data come from a definable population.

Our first example comes from the effect of breastfeeding in infants living in different water and sanitation conditions in Malaysia. Our stories, that breastfeeding saved lives (Habicht *et al.* 1986), and that this outcome was greater if water supply and sanitation

were poor (Butz *et al.* 1984), were widely accepted as plausible. We then calculated the impacts on measured mortality in groups of infants living in environments with different degrees of water and sanitation deficiencies (Habicht *et al.* 1988). The impact measures that we used were the proportion of death due to not breastfeeding (called the 'etiologic fraction') and the proportion of infants whose deaths were prevented because they breastfed (the 'preventive fraction'). The etiologic fraction estimates the beneficial impact of better breastfeeding, and the preventive fraction estimates the detrimental impact of deteriorating breastfeeding. These impact analyses showed that the impact of improving water and sanitation is greater where breastfeeding is less but that breastfeeding was beneficial in all environments, which is important for policy and programmes.

These impacts reflected the past in one place, but the same equations can be used to calculate impacts for future populations with different mortality and breastfeeding rates, and different water and sanitation environments. In other words, they can be used to localize decisions for action. In this case, policy and practice outstripped the impact calculations to support policy in other populations because the story that breastfeeding is good reinforced the beliefs of many public health professions. The outcome analyses alone carried the day without the impact analyses.

Our second example concerns the heterogeneous effects of improving water quality across different water and sanitation environments described previously. A review of the literature (Esrey & Habicht 1986) and subsequent outcome analyses suggested that appropriate faecal disposal was necessary before one saw effects from greater water supply. However, these effects were only seen when accompanied by improved hygiene behaviours and increased water usage (Esrey *et al.* 1992). Improved water quality was effective only if there was sufficient water, used properly and effective sanitary faecal disposal. This model would indicate that the first priority should be faecal disposal, then increased water supply coupled to better hygienic behaviours and lastly, improving water quality.

This story was widely accepted. It changed the emphasis on water quality to one of water supplies in developing countries, because it is easier by far to

supply any water than pure water. Improving water quality could come later. Recently, outcome parameters have been used to estimate the impact of improving water quality in the presence of high pathogen transmission rates as proxied by an index (Eisenberg *et al.* 2007). This represents a part of the model. It predicts much larger impacts than do analyses that do not take the model into account (Engell & Lim 2013). A full impact analysis depends on parameters from outcome analyses that correspond to the full model. As these are presently lacking, meta-analyses (Cairncross *et al.* 2010) have to work with parameters that are contaminated by mixing the different levels of water and sanitation. These impact analyses, which remain to be carried out, are important because the results will dictate appropriate sequences of action, and, in some cases, different actions.

Proposed actions must not only take into account the impact on the targeted health outcomes but also on other outcomes, including unintended outcomes. The impact assessment should include not only the positive effects but also the negative consequences of action (Andrews *et al.* 2013). Social, cultural and political issues that lie outside the scientific challenges often become centrally important, and it is difficult to disentangle these in conclusions about impact. In this sense, we can say that the definition and interpretation of impact are 'socially constructed'.

To illustrate the problems that can arise from not considering potential negative impacts, the experience of HIV and breastfeeding is instructive. When the possibility that the HIV virus could be transmitted to uninfected infants through breast milk was initially recognized, and its impact quantified, campaigns were initiated for HIV + mothers to stop breastfeeding because even exclusive breastfeeding transmitted the virus. This ban discommodulated breastfeeding policy and made breastfeeding promotion difficult for whole populations, even where HIV was rare. The infant morbidity and mortality risks of not breastfeeding were not taken into account, although they were already known to be high in populations where HIV had higher prevalences.

Over time, evidence accumulated concerning high mortality in the non-breastfed infants, which far outweighed the benefits of reduced HIV transmission (Kuhn *et al.* 2008). World Health Organization reversed its position, and other agencies followed suit, but not

before a great deal of harm was done. The breastfeeding ban would not have been proposed, much less implemented, had the already well-known negative impacts been weighed against the positive impacts. The ill effects of this ban on policy, and therefore on programmes, remain with us today, although they are difficult to quantify.

The foregoing examples demonstrate the importance of impact analysis after plausibility has been well established. Finally, we note that only positive evaluations of interventions and programmes can reinforce the plausibility of recommendations based on the steps described previously. Negative evaluations are difficult to interpret because they could be because of either incorrect outcome stories or to inadequate application of the intervention (Habicht & Peltó 2014). Negative results are too weak to support Popperian refutation. Other approaches, such as those described previously in the context of outcome analyses, are necessary. This higher order challenge of refutation through interventions, field trials and programmes remains to be addressed.

In conclusions, we concur with Conkle *et al.* (2016) in strongly encouraging research to estimate the epidemiological and other impacts of encouraging and discouraging premastication. This assessment must take into account other factors, including other causes of diarrhoea, as well as other outcomes of premastication, both detrimental and beneficial. However, the research community concerned with premastication has only started on a productive research route. We do not yet have plausible stories of causality against or in favour of premastication, and in what population groups. The story to be told by this research promises to be exciting.

Going beyond premastication, rigorous analyses to examine 'outcomes' and assess 'impact' are critical for making public health policy decisions that are grounded in the best procedures that current developments in science provide. Undoubtedly, these will change over time as new social constructions of the scientific process emerge.

Source of funding

None.

Conflicts of interest

The authors declare that they have no conflict of interest.

Contributions

Both authors JPH and GHP participated in all aspects of this paper.

References

- Aaltonen A.S. & Tenovu J. (1994) Association between mother-infant salivary contacts and caries resistance in children: a cohort study. *Pediatric Dentistry* **16** (2), 110–116.
- Andrews J.C. *et al.* (2013) GRADE guidelines: 15. Going from evidence to recommendation – determinants of a recommendation's direction and strength. *Journal of Clinical Epidemiology* **66** (7), 726–735.
- Bender R. & Lange S. (2001) Adjusting for multiple testing – when and how? *Journal of Clinical Epidemiology* **54** (4), 343–349.
- Butz W.P., Habicht J.P. & DaVanzo J. (1984) Environmental factors in the relationship between breastfeeding and infant mortality: the role of sanitation and water in Malaysia. *American Journal of Epidemiology* **119** (4), 516–525.
- Cairncross S, Hunt C, Boisson S, Bostoen K, Curtis V., Fung ICH, *et al.* (2010) Water, sanitation and hygiene for the prevention of diarrhoea. *International Journal of Epidemiology* **39** (Suppl), 193–205.
- Conkle J., Ramakrishnan U. & Freeman M. (2016) Prechewing infant food, consumption of sweets and dairy, and not breastfeeding are associated with increased diarrhea risk of ten month old infants in the USA. *Maternal and Child Nutrition*.
- Eisenberg J.N.S., Scott J.C. & Porco T. (2007) Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *American Journal of Public Health* **97** (5), 846–52.
- Engell R.E. & Lim S.S. (2013) Does clean water matter? An updated meta-analysis of water supply and sanitation interventions and diarrhoeal diseases. *The Lancet* **381**, S44.
- Esrey S.A. & Habicht J.P. (1986) Epidemiologic evidence for health benefits from improved water and sanitation in developing-countries. *Epidemiologic Reviews* **8** (1), 117–128.
- Esrey S.A., Habicht J.P. & Casella G. (1992) The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. *American Journal of Epidemiology* **135** (6), 659–666.
- Habicht J.P. (2004) Expert consultation on the optimal duration of exclusive breastfeeding – the process, recommendations, and challenges for the future in L Pickering, *et al.* eds. *Advances in Experimental Medicine and Biology* **554**, 79–87.
- Habicht J.P. (2000) The association between prolonged breastfeeding and poor growth – what are the implications? *Advances in Experimental Medicine and Biology* **478**, 193–200.
- Habicht J.P., DaVanzo J. & Butz W.P. (1986) Does breastfeeding really save lives, or are apparent benefits due to biases? *American Journal of Epidemiology* **123** (2), 279–290.
- Habicht J.P., DaVanzo J. & Butz W.P. (1988) Mother's milk and sewage: their interactive effects on infant mortality. *Pediatrics* **81** (3), 456–461.
- Habicht J.P. & Pelto G.H. (2014) From biological to program efficacy: promoting dialogue among the research, policy, and program communities. *Advances in Nutrition* **5** (1), 7–34.
- Habicht J.P., Victora C.G. & Vaughan J.P. (1999) Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology* **28** (1), 10–18.
- Hesselmar B., Sjöberg F., Saalman R., Åberg N., Adlerberth I., & Wold A.E. (2013) Pacifier cleaning practices and risk of allergy development. *Pediatrics* **131** (6), e1829–e1837.
- Hill A.B. (1965) The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* **58**, 295–300.
- Humphrey J.H., Jones A.D., Manges A., Mangwadu G., Maluccio J.A., Mbuya M.N. *et al.* (2015) The sanitation hygiene infant nutrition efficacy (SHINE) trial: rationale, design, and methods. *Clinical Infectious Diseases* **61** (suppl 7), S685–S702.
- Krieger N. (1994) Epidemiology and the web of causation: has anyone seen the spider? *Social Science and Medicine* **39** (7), 887–903.
- Kuhn L, Aldrovandi GM, Sinkala M, Kankasa C, Semrau K, Mwiya M, Kasonde P, *et al.* (2008) Effects of early, abrupt weaning on HIV-free survival of children in Zambia. *The New England Journal of Medicine* **359** (2), 130–41.
- Kuhn, T., 2012. The structure of scientific revolutions 50th anniv, Chicago 60673.
- MacMahon B. & Pugh T.F. (1975) *Epidemiology Principles and Methods*. Little, Brown and Company: Boston.
- Miller D. (1985) *Popper Selections*. Princeton University Press: Princeton, New Jersey.
- Pelto G.H., Zhang Y. & Habicht J.-P. (2010) Premastication: the second arm of infant and young child feeding for health and survival? *Maternal and Child Nutrition* **6** (1), 4–18.
- Porta M. (2008) *A Dictionary of Epidemiology*, 5th edn. Oxford University Press: Oxford.
- Rakhmanina N. *et al.* (2011) Premastication of food by caregivers of HIV-exposed children – nine U.S. Sites, 2009–2010. *Morbidity and Mortality Weekly Report* **60** (9), 273–275.
- Rothman K.J. (2002) *Epidemiology*. Oxford University Press, USA, New York.
- Susser M. (1991) What is a cause and how do we know one? A grammar for pragmatic epidemiology. *American Journal of Epidemiology* **133** (7), 635–648.