Causal inference and directed acyclic graph: An epidemiological concept much needed for oral submucous fibrosis

Naman R. Rao^a, Chandramani B. More^b, Rahi M. Brahmabhatt^{c,*}, Youbai Chen^{a,d}, Wai-kit Ming^{e,f}

^a Harvard Medical School, Harvard University, Boston, MA, USA

^b Department of Oral Medicine and Radiology, K. M. Shah Dental College and Hospital, Sumandeep Vidyapeeth University, Vadodara, Gujarat, India

^c College of Liberal Arts and Sciences, University of Illinois at Chicago, USA

^d Department of Plastic and Reconstructive Surgery, Chinese PLA General Hospital, Beijing, China

^e Faculty of Medicine, International School, Jinan University, Guangzhou, China

^f The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

ARTICLE INFO

Keywords:

Causal inference

Causal directed acyclic graph

Misconception

Oral submucous fibrosis research

ABSTRACT

Oral submucous fibrosis (OSMF), although already established as an oral potentially malignant disorder (OPMD), still stands over a weak bridge because of its controversial pathogenesis. There has been tremendous work on this disease since 1962, surprisingly, we are unsuccessful in finding the exact causation of OSMF. The potential cause for this is either a lack of systematically performed clinical observational studies or over-interpreted inferences of the presented results. Accordingly, the literature is piled with complex data that is being followed by emerging researchers. Hence, this conceptual paper is presented to focus and explain only the epidemiological concepts of causal inference and the construction of DAGs. These concepts will help to encode our subject matter knowledge and assumptions regarding the causal structure problem, classify the source of systematic bias, identify the potential confounders, potential issues in the study design, and guide the data analysis.

1. Introduction

Oral submucous fibrosis (OSMF) was first described by Schwartz et al. in 1952 as “Atropica Idiopathica Mucosae Oris” and formally defined as ‘oral submucous fibrosis’ by Pindborg and Sirsat in 1962 as ‘an insidious, chronic disease that affects any part of the oral cavity and sometimes the pharynx. Although occasionally preceded by, or associated with, the formation of vesicles, it is always associated with a juxta-epithelial inflammatory reaction followed by a fibroelastic change of the lamina propria and epithelial atrophy that leads to stiffness of the oral mucosa and causes trismus and an inability to eat’.^{1–3} Since 1962, various nomenclatures have been proposed in the scientific literature including Juxta-epithelial fibrosis, Idiopathic palatal fibrosis, Idiopathic scleroderma of mouth, Diffuse oral submucous fibrosis, Sclerosing stomatitis, and Oral fibrosis.⁴ Since the disorder is most commonly diagnosed as a clinical entity with a wide range of presentation of symptoms including pain, burning sensations, ulcerations, progressive restriction in mouth opening, blanching of the mucosa, depapillation of the tongue, loss of pigmentation and dysphonia, and hearing impairment in advanced cases (Fig. 1),^{6–9} recently, a clinical definition was proposed describing

OSMF as ‘a debilitating, progressive, irreversible collagen metabolic disorder induced by chronic chewing of areca nut and its commercial preparations; affecting the oral mucosa and occasionally the pharynx and esophagus; leading to mucosal stiffness and functional morbidity; and has a potential risk of malignant transformation’.⁵ As a well-documented disorder known for its high malignant transformation, it is considered a potentially malignant disorder of the oral cavity.^{4,10}

Etiopathogenesis till date includes areca nut chewing in any formulation,^{11–14} high uptake of chilies, infectious disease, copper-related oxidative stress, vitamin deficiency, increased levels of specific proteins,¹⁵ anemia, and genetic and immunologic predisposition.¹⁶ These factors were foremost identified from the observational data and limit its presentation to the level of “association”. However, many studies have unintentionally portrayed their study conclusion with the statement of “causality”. Accordingly, there exists complex literature on this disorder which often leaves the clinicians and researchers with confusion. The authors of this paper feel there is a strong need for developing critical epidemiological concepts that would enhance the differentiation between association and causal reports. This conceptual/expert opinion paper discusses different epidemiological approaches to study

* Corresponding author.

E-mail addresses: rahi.brbbhatt@gmail.com, rbbrahm2@uic.edu (R.M. Brahmabhatt).

<https://doi.org/10.1016/j.jobcr.2020.06.008>

Received 29 May 2020; Received in revised form 11 June 2020; Accepted 21 June 2020

Available online 03 July 2020

2212-4268/ © 2020 Craniofacial Research Foundation. Published by Elsevier B.V. All rights reserved.

Fig. 1. Clinical presentation of oral submucous fibrosis.

Legend: Panel A shows the blanching of palatal mucosa and Panel B shows the presence of blanching and thin fibrous bands in the left buccal mucosa.

the relationship between an exposure and the outcome, consequently helping clinicians and researchers/investigators to differentiate between the association and causation, design appropriate clinical studies, and report apt scientific conclusions. Authors urge the researchers to read the book *Causal Inference* by Hernan M. A. and Robins J. M. (2019) and *Epidemiology*, an introduction by Kenneth J. Rothman to learn the in-depth concept for better applicability.

2. Interpreting association

The primary purpose of all epidemiological studies is to identify the statistical association between disease and exposure. It is essential to take a step ahead and interpret this association relationship. Furthermore, association can be artifactual, non-causal, or causal.¹⁷ This paper focuses on how a causal association relation can be established. We observe that many articles report a conclusion based on the crude analysis. This conclusion holds no scientific value if the data is not analyzed with stratum-specific analysis, which evaluates for the confounder and effect modifier.¹⁷ It is crucial to understand the following terms for interpretation: Confounder- Confounding occurs when the relationship between an exposure (risk factor) and an outcome (disease) is misrepresented (distorted) because each of these variables is also related to a third variable, known as the confounder. Effect modifier- Interaction occurs with a covariate if the stratum-specific measures of association differ from each other when the confounder defines the strata. The goal of this paper is to enumerate the approach or methodology to assess causal link between exposure and outcome and the general framework on how it can be applied to OSMF (Fig. 2).

3. What is the causal association?

Kenneth J. Rothman defines cause as “an act or event or state of nature which initiates or permits, alone or in conjunction with other causes, a sequence of events resulting in an effect”.¹⁸ Alternately, a cause can be defined “as a condition, characteristic or event which precedes the disease, and in the absence of which the disease would not have occurred at all or would not have occurred until the passage of a certain amount of time”.¹⁹ For example, OSMF will not occur in absence of areca nut and/or its by product consumption.

Most often, there is not only one sole cause for a disease; instead, a disease occurrence is a result of action and interaction of various causal factors at the appropriate time.¹⁹ Every individual cause required to produce disease is called a component cause, but together their collection is known as the sufficient cause model of a disease or the causal pie model.¹⁹ When each component of the adequate cause mechanism

is present, it will inevitably lead to the occurrence of the desired outcome or disease. It is relevant to note that there can be more than one sufficient cause model for any disease.²⁰ However, as science advances our knowledge, component causes can be added or deleted from an adequate cause model (Fig. 3).

An important implication of the causal pie model is the concept of multi-causality, which states that a specific event is the result of the joint action of several component causes.¹⁹ For simplicity, let us take the hypothetical example of Jones, who developed OSMF, to understand the concept of multi-causality. Jones have been chewing areca nut and its byproducts for five years, therefore playing a crucial role in developing OSMF. The causal pie model for Jones' OSMF condition may involve components such as high uptake of chilies, infectious diseases, vitamin deficiency, increase level of specific proteins, anemia, genetic and immunological predisposition, and/or other unknown causes (Fig. 4). In nearly all cases, the component causes of a disease do not work simultaneously; they gather over a period, where the last component cause completes the causal pie. In the case of Jones, the earliest acting component cause would be the areca nut and its byproduct chewing habits. According to Rothman, the time interval between the earliest acting component cause and the completion of the causal pie by the last-acting component cause is known as the induction time for the disease. In contrast, the latent period of a disease is the time interval between the onset of a disease and its detection by diagnostic tests or symptoms.

4. Causal directed acyclic graph (DAG)

Greenland et al. first introduced the use of DAGs for epidemiological research in 1999.²⁰ However, Merchant and Pitiphat championed their use for dental research in 2002.²¹ A directed acyclic graph (DAG) is a non-parametric representation of the causal relationship between an exposure and outcome.²² The term acyclic in DAG refers to the fact that a directed path never forms a closed loop.²⁰ In addition to depicting factors influencing the exposure or outcome, DAGs also incorporate mediators that are a result of the exposure or outcome as well as variables that are causal to other factors depicted in the DAG.²²

Causal DAG construction: An analysis model consisting of a set of assumptions is a prerequisite to start constructing any DAG.²⁰ A variable in a DAG is called node or vertex, and any two variables with a relationship are connected by a line called arc or edge, which can include an arrowhead.²¹ An arc directly connects adjacent variables in a DAG.²⁰ A single arrowhead represents a direct effect of one variable over the other, but not the other way around.²⁰ In Fig. 5, P and Q are adjacent variables while P and R are not; P has a direct effect on Q,

Fig. 2. Systematic method for reporting effective inference.

which is not mediated by any other variable. Variable P is called an ancestor or parent for variables Q and S, meaning Q and S are the descendants of P.²⁰ Absence of a single-headed arrow between any two variables denotes a null effect assumption and no causal effect between those variables,²⁰ which is denoted by the relations between S and R as well as Q and S. Some variables in a DAG can be represented by two-headed or bi-directional arrows indicating a common ancestor whereas a non-directional arc between two variables can indicate that they are

related due to reasons other than sharing an ancestor or affecting one other.²⁰ For example, Q and R may have a common ancestor who is not depicted here.

A path is an uninterrupted route joining the nodes along the arcs in a DAG irrespective of the direction of the arrowheads between the variables.²² When a path can be traced along with single-headed arrows, which exclusively enter through the tail and exit through the head, it is called a directed or causal path.^{20,22} A path can be classified

Fig. 3. Two possible causal pie models for the same outcome. Fig. 2A shows a causal pie model with components A, B, C, D, E causing outcome. Fig. 2B Shows a causal pie model with A, B, C, G, H causing same outcome as shown in Fig. 2A.

Fig. 4. Hypothetical causal pie model for Jones' OSMF condition.

Fig. 5. Causal DAG construction.

as front or back door, and as an open or closed path.²¹ Front door, a causal path, is depicted when the arrowheads point from the exposure to the outcome, while a backdoor path, non-causal path, has more than one arrowhead pointing towards the exposure rather than away from it.^{19,22} A front door path where all the arrows flow in the same direction is considered open whilst in a closed front door path the direction of at least one arrow is changed.²² Variables present between exposure and outcome on an open causal path are called mediators. Other than the arrowhead(s) pointing to the exposure and the arrowhead pointing to the outcome, all the other arrows on an open backdoor path flow in the same direction.²² (Table 1).

DAGs provide a scientific way to encode our subject-matter knowledge and assumptions about the causal structure of a problem, identify the potential confounders, classify sources of systematic bias, identify the potential issues in the study design, and guide data

Table 1	
-Examples of causal DAG.	
Examples of Causal DAG's	
	Open front door A→B→C
	Closed front door A→B↔D→C
	Open backdoor D←C→B→E
	Closed backdoor D←A→C←B→E

analysis.²³

5. Conclusion

To treat OSMF or any other disease, knowing its pathology is necessary. As observed with the literature to date, it is difficult to conclude a cause for OSMF due to limited use of appropriate causal inference methodologies and reliance on naive or univariate associations. Overall, this paper discusses the established concepts of causal inference and its DAG formation, which enlighten the researchers to understand the differentiation between association and causation, develop an appropriate study design, and the appropriate method of presenting results. In order to address the etiopathogenesis of OSMF,

appropriate causal methodologies should be adopted which will help in understanding the risk factors, biology and potential interventions for prevention and/or clinical management.

Funding sources

Self.

Declaration of competing interest

Nil.

References

- Pindborg JJ, Sirsat SM. Oral submucous fibrosis. *Oral Surg Oral Med Oral Pathol*. 1966;22(6):764–779.
- More C, Shilu K, Gavli N, Rao NR. Etiopathogenesis and clinical manifestations of oral submucous fibrosis, a potentially malignant disorder: an update. *Int J Curr Res*. 2018;10:71816–71820 07.
- Rao NR, Villa A, More CB, Jayasinghe RD, Kerr AR, Johnson NW. Oral submucous fibrosis: a contemporary narrative review with a proposed inter-professional approach for an early diagnosis and clinical management. *J Otolaryngol - Head Neck Surg*. 2020;49(1).
- Waal I. Historical perspective and nomenclature of potentially malignant or potentially premalignant oral epithelial lesions with emphasis on leukoplakia—some suggestions for modifications Oral. *Surg Oral Med Oral Pathol Oral Radiol*. 2018;125:577–581.
- More CB, Rao NR. Proposed clinical definition for oral submucous fibrosis. *J Oral Biol Craniofac Res*. 2019;9(4):311–314.
- Tilakaratne WM, Ekanayaka RP, Warnakulasuriya S. Oral submucous fibrosis: a historical perspective and a review on etiology and pathogenesis. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2016;122(2):178–191.
- More C, Pawar R, Rao N, Shah P, Gavli N. Oral ulcer: an overview with an emphasis on differential diagnosis. *Int J Oral Health Sci Adv*. 2015;3(4):1–13.
- Warnakulasuriya S, Kerr AR. Oral submucous fibrosis: a review of the current management and possible directions for novel therapies. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2016;122(2):232–241.
- More C, Shah P, Rao N, Pawar R. Oral submucous fibrosis: an overview with evidence-based management. *Int J Oral Health Sci Adv*. 2015;3(3):40–49.
- Adalja C, Adalja C, More C, Rao N. Role of stem cell therapy in oral premalignant and malignant lesions. *Int J Curr Res*. 2016;8(12):43880–43883.
- More CB, Rao NR, Hegde R, Brahmabhatt RM, Shrestha A, Kumar G. Oral submucous fibrosis in children and adolescents: Analysis of 36 cases. *J Indian Soc Pedod Prev Dent*. 2020;38(2):190–199. https://doi.org/10.4103/JISPPD.JISPPD_173_20.
- Ahmad MS, Ali SA, Ali AS, Chaubey KK. Epidemiological and etiological study of oral submucous fibrosis among gutkha chewers of Patna, Bihar, India. *J Indian Soc Pedod Prev Dent*. 2006;24(2):84–89.
- More CB, Rao NR, More S, Johnson NW. Reasons for initiation of areca nut and related products in patients with oral submucous fibrosis within an endemic area in Gujarat, India. *Substance use and Misuse*. 2020;55(9):1413–1421. <https://doi.org/10.1080/10826084.2019.1660678>.
- More CB, Gavli N, Chen Y, Rao NR. A novel clinical protocol for therapeutic intervention in oral submucous fibrosis: an evidence based approach. *J Oral Maxillofac Pathol*. 2018;22(3):382–391.
- Anila Nambodiripad PC. Cystatin C: its role in pathogenesis of OSMF. *J Oral Biol Craniofac Res*. 2014;4(1):42–46.
- More C, Peter R, Nishma G, Chen Y, Rao N. Association of Candida species with oral submucous fibrosis and oral leukoplakia: a case control study. *Ann Clin Lab Res*. 2018;6(3):1–5. <https://doi.org/10.21767/2386-5180.100248>.
- Advisory Committee to the Surgeon General of the Public Health Service. Smoking and health; 1964 282-189). P.H.S publication number 1103. Washington, DC: Public Health Service.
- Rothman KJ. *Reviews and Commentary CAUSES*. vol. 141. 1995; 1995.
- Rothman KJ, Greenland S. Causation and causal inference in Epidemiology. *Am J Publ Health*. 2005;95(S1).
- Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. January 1999;10(1):37–48.
- Merchant AT, Pitiphat W. Directed acyclic graphs (DAGs): an aid to assess confounding in dental research. *Community Dent Oral Epidemiol*. 2002;30(6):399–404.
- Akinkugbe AA, Sharma S, Ohrbach R, Slade GD, Poole C. Directed acyclic graphs for oral disease research. *J Dent Res*. 2016;95(8):853–859.
- Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC; 2019:1–133.