BMJ Open

Oral health of high-cost patients: findings from a large-scale health check-up database in South Korea

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-032446
Article Type:	Research
Date Submitted by the Author:	19-Jun-2019
Complete List of Authors:	Kim, Yeonkook; Chungbuk National University,
Keywords:	High-cost patients, oral health, oral health behavior, self-reported oral symptoms

Title: Oral health of high-cost patients: findings from a large-scale health check-up database in South Korea

Y. Joseph Kim, Ph.D.[1]

Author affiliation

[1] Chungbuk National University, Cheongju, Chungbuk, Rep. of Korea.

Correspondence to Dr. Y. Joseph Kim; yjk15@chungbuk.ac.kr

Keywords: High-cost patients, oral health, oral health behavior, self-reported oral symptoms

Word count: 3,243

ABSTRACT

Objective: To examine the oral health conditions and oral health behavior of high-cost patients and evaluate oral health measures as predictors of future high-cost patients.

Design: A retrospective, population-based cohort study using administrative healthcare records.

Setting: The National Health Insurance Service (NHIS) medical check-up database (a.k.a. NHIS - national health screening cohort database) in South Korea.

Participants: 131,549 individuals who received biennial health check-ups including dental check-ups in 2011 or 2012, aged 49–88.

Primary outcome measures: Current and subsequent year high-cost patient status.

Results: High-cost patients, on average, incur higher dental costs, suffer more from periodontal disease, brush their teeth less and use secondary oral hygiene products less. Some of the self-reported oral health behaviors and oral symptom variables show statistically significant associations with subsequent year high-cost patient indicators, even after adjusting for demographic, socio-economic, medical conditions, and prior healthcare cost and utilization.

Conclusions: We demonstrate that oral health measures are associated with an increased risk of becoming a high-cost patient.

Strengths and limitations of this study

- This is the first study to provide empirical evidence of the oral health characteristics of high-cost patients.
- This study is conducted using a large-scale health check-up cohort database.
- The results of this study show associations between oral health measures and future high-cost patient status but do not explain why the association exists or imply a causal relationship.
- This study is limited by the observation period of two consecutive years (i.e. we predict the outcome in 2012 using the data from 2011 and we predict the outcome in 2013 using the data from 2012).

INTRODUCTION

Studies have shown that a small portion of the population uses the majority of healthcare resources, highlighting the importance of predicting high-cost patients by recognizing the characteristics and patterns of their healthcare utilization for targeted interventions and for the modelling of patient risk factors for more equitable healthcare reimbursements.[1] If certain patient characteristics are predictive of highly persistent use, it may be possible to offer more cost-effective alternatives to frequent primary care visits, including disease management, case management, group visits, and patient education.[2] Wammes et al. conducted an extensive systematic review of the characteristics and healthcare utilization of high-cost patients.[3] However, to the best of our knowledge, the oral health characteristics of high-cost patients have not been studied.

Poor oral health has been shown to be associated with increased risk of various diseases and mortality.[4–7] For example, studies show associations between periodontal disease and a higher risk of Alzheimer's disease, coronary heart disease, rheumatoid arthritis, glomerulonephritis, inflammatory bowel disease and prostate cancer.[8–10] Other studies show that poor oral hygiene is associated with hypertension, diabetes mellitus, cardiovascular disease, and head and neck cancers.[11–14] Still other studies have found a positive association between poor oral health and mortality.[4,7,15]

Given that poor oral health is associated with many of chronic and severe diseases, we conjecture that high-cost patients have poor oral health. To test this hypothesis, using a large-scale health check-up database in Korea, we analyze the oral health of high-cost patients. One main reason to study the characteristics of high-cost patients is to predict future high-cost patients for targeted interventions. We evaluate oral health measures as predictors in the high-cost patient predictions. In short, our objective is to examine the oral health conditions and oral health behaviors of current high-cost patients and to evaluate which oral health measures identify future high-cost patients.

MATERIALS AND METHODS

Participants and setting

We analyze the National Health Insurance Service (NHIS) medical check-up database (also known as the NHIS - national health screening cohort (NHIS-HEALS)) of Korea.[16] The NHIS is a uniform insurance policy administrator that covers all residents of Korea. The NHIS datasets represent the entire population in Korea. Therefore, this database is a suitable source of information for nationwide population-based studies.[17]

To construct the NHIS-HEALS database, a sample cohort was first selected from the 2002 and 2003 health screening participants (aged between 40 and 79 in 2002) and followed up through 2013.[16] In 2002, this cohort has 514,866 health check-up participant, which is a 10% random sample of the total number of health check-up participants in 2002 and 2003. The NHIS-HEALS database consists of four databases: an eligibility database, a medical check-up database, a claims database, and a healthcare provider database. To use the most recent data in the database, we construct the study database with a medical check-up data set that contains information on examinees who received biennial medical and dental check-ups in 2011 or 2012.

The medical check-up data is generated through the National Health Screening Program (NHSP), which is a free national medical check-up program for NHIS beneficiaries who are 40 or older.[18] NHIS enrollees eligible for the NHSP are required to have a medical check-up biennially. The NHSP consists of a set of laboratory tests, a dental examination, and questionnaires on self-reported health behavior. Our study sample consists of 131,549 enrollees who participated in the NHSP in 2011 or 2012 aged from 49 to 88. For those enrollees who participated in both years, we only use 2011 observation to avoid duplication.[18] Since the majority of high cost patients are 50 years and older in Korea and many other countries, we use a large sample of that population segment in this study.[1,3,19] Using a join key in the database, we extract their claims and eligibility information from the claims and eligibility datasets, respectively, and merge them with the medical check-up information.

Patient involvement

A de-identified population healthcare data is used in this study. Thus, patients were not involved in the development of the research question, the outcome measures or the study design.

Statistical methods

Variables

Following prior studies, we define a 'high-cost' patient as an individual in a sample who is in the upper decile of annual healthcare expenditures.[1,3] We create a binary indictor for high-cost patients using all of the data in the database. For example, the medical check-up database contains 487,835 enrollees in 2011, and we tag enrollees who are in the upper decile of annual healthcare expenditures as high-cost patients. See online supplementary file1 for further details.

To understand the oral health conditions of high cost patients, we compare the characteristics and health behavior of these patients with those of a non-high-cost group. We compare demographic factors (gender and age), socioeconomic (income level) information, healthcare expenditures, and the total inpatient length of stay (LOS) as a measure of healthcare utilization.

We examine the three lifestyle risk factors of body mass index (BMI), pack-years smoked, and physical activity, all of which are associated with future healthcare expenditures.[20] Pack-years smoked values are calculated by multiplying the number of packs of cigarettes smoked per day by the number of years the person has smoked. The physical activity measure is from a questionnaire asking about the number of days a check-up participant exercises at a moderate level for more than 30 minutes in activities such as fast walking, playing tennis, and riding a bicycle.

Prior studies point to a high prevalence of multiple chronic conditions to explain high-cost patients' high utilization rates.[3] We examine the chronic conditions of high-cost patients. Disease information in our claims dataset is coded based on the International Classification of Diseases-Tenth Revision (ICD-10) coding scheme.

We select six oral health measures from the dental check-up and questionnaire data in the database based on findings from prior studies and on the statistical significance of the association with high-cost patient status. The first three variables are the prevalence of periodontal diseases, the frequency of toothbrushing, and the use of floss or interdental brushes. The last three are self-reported oral symptoms: tongue or inside-cheek pain (yes/no), difficulty in enunciation due to teeth/denture/gum conditions (yes, slightly, or no) and difficulty in chewing food due to teeth/denture/gum conditions (yes/no). We exclude participants with missing data for toothbrushing frequency (n=1305; 0.99%), use of floss/interdental brushes (n=554; 0.42%), tongue or inside-cheek pain (n=472; 0.36%), difficulty in enunciation (n=4042; 3.07%) and difficulty in chewing food (n=3511; 2.67%) from the corresponding analyses.

Statistical Analyses

We conduct a chi-square test and a t-test to compare the characteristics and health behaviors of the high-cost patient group with those of the non-high-cost group in years 2011 and 202. Because healthcare expenditure variables exhibit marked positive skewness, with a few high-cost patients and many low- or zero-expenditure healthcare users, we perform both a t-test and a Wilcoxon rank sum test.

To evaluate oral health measures as predictors of high-cost patients, we measure the associations between oral health measures and future high-cost patient status (i.e., a binary indictor of high-cost patient status in the subsequent year) using multivariate logistic regression models. We predict the outcome in 2012 using the data from 2011, and we predict the outcome in 2013 using the data from 2012. We select covariates in our models based on prior high-cost patient and healthcare cost prediction studies.[1,20–22] We develop four different models. The first model adjusts for three confounders: age, gender and income. For the second model, we add the three lifestyle risk factors of BMI, pack-years smoked, and physical activity. To adjust for chronic conditions, we add 44 binary chronic disease indicators to the first model. We use the stepwise feature-selection procedure in SAS version 9.4 (SAS Institute, Cary, NC, USA) to eliminate insignificant and collinear variables. For the last model, we add

prior healthcare cost and LOS.

RESULTS

The general characteristics of the high-cost patients are presented in Table 1. The percentage of high-cost patients in our sample is 6.52%, as people who go to health check-ups tend to be younger and less likely to be high-cost patients. Consistent with prior studies, high costs are associated with increasing age.[3] The mean age for the non-high-cost group is 58, whereas the mean for the high-cost patient group is 64 in our sample data. As income increases, the likelihood of becoming a high-cost patient decreases. Enrollees in medical aid programs do not report income, and a much higher proportion of high-cost patients are in medical aid programs compared to those in the non-high-cost group (1.5% vs 0.24%). 7.35% of total high-cost patients are in the lowest income decile, whereas 6.45% of the non-high-cost group are in the same decile. On the other hand, 22.25% of total high-cost patients are in the highest income decile, whereas 24.32% of the non-high-cost group are in the same decile. The average total cost for the high-cost patient group in 2011 is $6,608, almost eight times higher than that for the non-high-cost group. The total average drug cost for high-cost patients is about 3.8 times higher than that for the non-high-cost group as of 2011. High-cost patients also incurred 1.4 times more, on average, in dental costs compared to the non-high-cost group. With regard to healthcare expenditure variables, we performed both a t-test and a Wilcoxon rank sum test and found statistically significant differences at the 0.1% level for both tests.

Inset [Table 1] here.

Next, we report three lifestyle risk factors for high-cost patients: BMI, pack-years smoked and physical activity. Consistent with the result in Leigh et al., high-cost patients have higher BMIs (24.32 vs. 24.02), smoke more (8.31 vs. 7.82) and exercise less (1.34 days vs.1.47 days).[20]

We report the average count of chronic diseases for high-cost patients in Table 1. On average, high-cost

patients have 4.87 chronic conditions while those in the non-high-cost group have 2.26. For example, more than half of our high-cost patients suffer from chronic lower back pain. Another common chronic disease for high-cost patients in our sample is hypertension. Approximately 39% of high-cost patients have hypertension, while 26% in the non-high-cost group suffer from the same disease. The prevalence of cancer is much higher among high-cost patients than among non-high-cost patients (22.04 % vs 3.03 %). Chronic ischemic heart disease is another chronic disease with a large disparity: 15.16% for high-cost patients vs. 3.23% for the non-high-cost group.

While this disparity of chronic disease conditions has been reported in numerous prior studies of high-cost patients, the oral health characteristics of high-cost patients have not been reported.[3] Table 2 presents the oral health condition, oral health behaviors and the three aforementioned self-reported oral symptoms.

Inset [Table 2] here.

First, we report the prevalence of periodontal diseases for the high-cost patients in Table 2. The prevalence of periodontal diseases is higher in the high-cost patient group than in the non-high-cost group (28.6 % vs 25.9%). The first oral health behavior variable is the frequency of toothbrushing per day. Higher proportions of high-cost patients report one time or less per day or two times per day, while a higher proportion of the non-high-cost group reports three times or more per day. The second oral health behavior is the use of floss or/and interdental brushes. Again, high-cost patients use secondary oral products less: 5.64% for the high-cost patient group and 7.2% for the non-high-cost group.

We report three self-reported oral symptoms. High-cost patients are more likely to report tongue or inside-cheek pain: 7.45% for the high-cost patient group and 6.40% for the non-high-cost group. The second symptom is difficulty in enunciation due to teeth/denture/gum conditions. A higher proportion of high-cost patients report this symptom compared to those in the non-high-cost group: 8.49% of high-cost patients reported this symptom, while 5.45% of the non-high-cost group did so. The third symptom is difficulty in chewing food due to teeth/denture/gum conditions, with 22% of the high-cost patient

group reporting this symptom and 16.57% of the non-high-cost group doing so.

Table 3 presents the result of the binominal logistic regression analyses. Poor oral health measures are associated with a higher likelihood of a person becoming a high-cost patient in the subsequent year. The univariate odds ratios for all of the oral health measures are significant at the 0.1% level. Even after we adjust for age, gender and income level in the first model, all oral health measures have odds ratios greater than one, showing that poor oral health increases the odds of becoming a high-cost patient in the following year. In the second model, where we additionally adjust for the three lifestyle risk factors of BMI, pack-years smoked, and physical activity, the odds ratios for oral health behavior become less significant. The odds ratio of the frequency of toothbrushing becomes statistically insignificant and the odds ratio for the use of floss or interdental brushes becomes less significant in the second model. However, the odds ratio for the prevalence of periodontal disease (OR=1.11 and 95% CI 1.06 to 1.16) and those for the self-reported symptoms remain statistically significant in the second model.

Inset [Table 3] here.

To adjust for a high prevalence of multiple chronic conditions in high-cost patients, we adjust for chronic conditions in our third model. When chronic disease variables are introduced in the model, the odds ratio for periodontal disease becomes statistically insignificant. To our final model, we add prior total healthcare cost and LOS, as they are powerful predictors of future high-cost patients.[1] The odds ratios for the variables of use of floss or interdental brushes, tongue or inside-cheek pain, and difficulty in enunciation are significant at the 1% level.

DISCUSSION

In the present study, we show a positive association between poor oral health and high-cost patient status. High-cost patients have a higher prevalence of periodontal diseases (28.6 % vs 25.9%). They are more likely to brush their teeth less than three times a day and are less likely to use floss or interdental brushes than the non-high-cost group. This positive association can partly be explained by prior studies

that show positive associations between periodontal disease and various chronic and severe diseases.[8–10] Other studies also find that the frequency of toothbrushing and the use of secondary oral hygiene products are associated with diseases such as hypertension, diabetes mellitus, cardiovascular disease, and head and neck cancers.[11–14] In a recent study, researchers found that flossing and brushing of interdental spaces may reduce the risk for new cardiovascular events among patients with coronary heart disease.[23] Moreover, our results show that the high-cost patient group is more likely to report tongue or inside-cheek pain (7.45% vs. 6.40%), difficulty in enunciation due to teeth/denture/gum conditions and difficulty in chewing food due to teeth/denture/gum conditions (22% vs. 16.57%) as compared to the non-high-cost group.

The prevalence of periodontal diseases shows a positive association with future high-cost patient status. However, when we adjust for chronic conditions, the statistical significance of the odds ratio disappears. The associations between the oral health behavior variables and future high-cost status are also statistically significant. However, the associations become less statistically significant when lifestyle risk factors are introduced in the model. This is due to the association between oral health behaviors and lifestyle risk factors, and the result is consistent with those in prior studies.[6,24,25] Kim et al. show that an increased frequency of toothbrushing and the number of secondary oral products used are associated with a lower BMI, less smoking and a higher level of physical activity.[6]

The associations between self-reported oral symptoms and future high-cost patient status remain statistically significant even after we adjust for demographic and socioeconomic measures, lifestyle risk behaviors, chronic diseases, and prior healthcare costs and utilization rates. This shows that self-reported oral symptoms are a potential new source of data for high-cost prediction modelling.

Strengths and limitation

To the best of our knowledge, this is the first study to focus on the oral health of high-cost patients. Prior studies of high-cost patients and healthcare cost predictions indicate a high prevalence of multiple (chronic) conditions and lifestyle risk factors.[3,20] However, the oral health characteristics of high-cost

patients have not been studied. Our study fills this void in the literature using a large sample of 131,549 enrollees. This study, however, does not explain or imply a causal relationship between oral health and the likelihood of becoming a high-cost patient. As Griffin et al. point out, oral diseases can have an impact on many aspects of general health, and health conditions can in turn have an impact on oral health.[26] Another limitation of this study is that we do not consider the long-term effects of oral health on healthcare expenditures. However, one-year prediction model is the most common study design in high-cost patient prediction literature.[3]

Policy and research implications

There are several implications pertaining to our work. Wammes et al. argue that high-cost patients make up the sickest and most complex populations.[3] They are a small portion of the population yet use the majority of healthcare resources. Accurately identifying high-cost patients and managing their care is a significant first step in improving quality levels and reducing population health costs.[1] Actively exploring data sources available for their identification and prediction is a requisite for achieving these goals. To this end, our analysis provides a new data source for high-cost prediction modelling. Moreover, we have shown that people with poor oral health are at an increased risk of becoming high-cost patients. In that oral and other chronic and severe diseases also share common risk factors, it is important to examine the interplay between these diseases and oral disease as well as their combined impact when health policymakers develop programs involving targeted interventions for high-cost patients and develop preventive measures at the population level.[26]

In conclusion, we demonstrate that oral health measures are associated with the risk of becoming a high-cost patient. Our results highlight the impact of oral health on healthcare costs and support the development of preventive measures at the population level.

Acknowledgments I would like to thank Ms. Boram Park and Ms. Kyung-Ran Lee at the NHIS for providing answers to numerous questions on the NHIS medical check-up database used in this study.

Author contributions YJK conceptualized and designed the study, collected and analyzed the data, and wrote the manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None.

Ethics approval The Chungbuk National University Institutional Review Board, The Seoul National University Institutional Review Board.

Data sharing statement Information about accessing study data is available at: https:// nhiss.nhis.or.kr.

References

1. Tamang S, Milstein A, Sørensen HT, et al. Predicting patient ' cost blooms ' in Denmark : a longitudinal population-based study. *BMJ Open*. 2017;7(1):e011580. doi:10.1136/bmjopen-2016-011580
2. Naessens JM, Baird MA, Van Houten HK, Vanness DJ, Campbell CR. Predicting persistently high primary care use. *Ann Fam Med*. 2005;3(4):324-330. doi:10.1370/afm.352
3. Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open*. 2018;8(9):e023113. doi:10.1136/bmjopen-2018-023113
4. Adolph M, Darnaud C, Thomas F, et al. Oral health in relation to all-cause mortality: The IPC cohort study. *Sci Rep*. 2017;7(October 2016):3-8. doi:10.1038/srep44604
5. Sabbah W, Mortensen LH, Sheiham A, Batty D. Oral health as a risk factor for mortality in middle-aged men: The role of socioeconomic position and health behaviours. *J Epidemiol Community Health*. 2013;67(5):392-397. doi:10.1136/jech-2012-201340
6. Kim Y-H, Kim D-H, Lim KS, et al. Oral health behaviors and metabolic syndrome: the 2008–2010 Korean National Health and Nutrition Examination Survey. *Clin Oral Investig*. 2014;18(5):1517-1524. doi:10.1007/s00784-013-1112-2
7. Joshy G, Arora M, Korda RJ, Chalmers J, Banks E. Is poor oral health a risk marker for incident cardiovascular disease hospitalisation and all-cause mortality? Findings from 172 630 participants from the prospective 45 and Up Study. *BMJ Open*. 2016;6(8):e012386. doi:10.1136/bmjopen-2016-012386
8. Galgut PN. Periodontal disease and poor health outcomes. *Bmj*. 2010;340(jun17 1):c2735-c2735. doi:10.1136/bmj.c2735
9. Lee JH, Kweon HHI, Choi JK, Kim YT, Choi SH. Association between periodontal disease and prostate cancer: results of a 12-year longitudinal cohort study in South Korea. *J Cancer*. 2017;8(15). doi:10.7150/jca.20532
10. Gaur S, Agnihotri R. Alzheimer's disease and chronic periodontitis: Is there an association? *Geriatr Gerontol Int*. 2015;15(4):391-404. doi:10.1111/ggi.12425
11. Kuwabara M, Motoki Y, Ichiura K, et al. Association between toothbrushing and risk factors for cardiovascular disease: A large-scale, cross-sectional Japanese study. *BMJ Open*. 2016;6(1). doi:10.1136/bmjopen-2015-009870
12. de Oliveira C, Watt R, Hamer M. Toothbrushing, inflammation, and risk of cardiovascular disease: results from Scottish Health Survey. *Bmj*. 2010;340(may27 1):c2451-c2451. doi:10.1136/bmj.c2451
13. Farquhar DR, Divaris K, Mazul AL, Weissler MC, Zevallos JP, Olshan AF. Poor oral health affects survival in head and neck cancer. *Oral Oncol*. 2017;73(July):111-117. doi:10.1016/j.oraloncology.2017.08.009
14. Darnaud C, Thomas F, Pannier B, Danchin N, Bouchard P. Oral health and blood pressure: The IPC cohort. *Am J Hypertens*. 2015;28(10):1257-1261. doi:10.1093/ajh/hpv025
15. Kim JK, Baker LA, Davarian S, Crimmins E. Oral health problems and mortality. *J Dent Sci*. 2013;8(2):115-120. doi:10.1016/j.jds.2012.12.011

16. Seong SC, Kim YY, Park SK, et al. Cohort profile: The National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. *BMJ Open*. 2017;7(9). doi:10.1136/bmjopen-2017-016640

17. Lee J, Lee JS, Park S-H, Shin SA, Kim K. Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol*. 2016:dyv319. doi:10.1093/ije/dyv319

18. Kim HB, Lee SA, Lim W. Knowing is not half the battle: impacts of information from the national health screening program in Korea. *IZA Discuss Pap*. 2017;(10650).

19. Kim ST, Seong SC. *National Health Insurance Statistical Yearbook*. Seoul: Health Insurance Review & Assessment Service, National Health Insurance Service; 2015.

20. Leigh JP, Hubert HB, Romano PS. Lifestyle risk factors predict healthcare costs in an aging cohort. *Am J Prev Med*. 2005;29(5):379-387. doi:10.1016/j.amepre.2005.08.005

21. Fleishman JA, Cohen JW. Using information on clinical conditions to predict high-cost patients. *Health Serv Res*. 2010;45(2):532-552. doi:10.1111/j.1475-6773.2009.01080.x

22. Leininger LJ, Friedsam D, Voskuil K, DeLeire T. Predicting high-need cases among new Medicaid enrollees. *Am J Manag Care*. 2014;20(9):e399-407. http://www.ncbi.nlm.nih.gov/pubmed/25364876.

23. Reichert S, Schlitt A, Beschow V, et al. Use of floss/interdental brushes is associated with lower risk for new cardiovascular events among patients with coronary heart disease. *J Periodontal Res*. 2015;50(2):180-188. doi:10.1111/jre.12191

24. Tada A, Matsukubo T. Relationship Between Oral Health Behaviors and General Health Behaviors in a Japanese Adult Population. *J Public Health Dent*. 2003;63(4):250-254. doi:10.1111/j.1752-7325.2003.tb03508.x

25. Kim H-J, Kim Y-H, Cho K-H, et al. Oral health behaviors and bone mineral density in South Korea: the 2008–2010 Korean National Health and Nutrition Examination Survey. *J Bone Miner Metab*. 2016;34(2):225-233. doi:10.1007/s00774-015-0669-z

26. Griffin SO, Jones JA, Brunson D, Griffin PM, Bailey WD. Burden of oral disease among older adults and implications for public health priorities. *Am J Public Health*. 2012;102(3):411-418. doi:10.2105/AJPH.2011.300362

Table 1 Current high-cost patient characteristics

Patient characteristics	Non high cost	High cost	p Value*
N	122,974	8,575	
%	93.48%	6.52%	
% female	40.8%	47.6%	
Age (years)	58.45 ± 0.02	63.83 ± 0.09	<.0001
Income decile (%)			<.0001
Medical Aid	0.24 (0.01)	1.5 (0.13)	
1	6.45 (0.07)	7.35 (0.28)	
2-5	24.61 (0.12)	23.25 (0.46)	
6-9	44.39 (0.14)	45.64 (0.54)	
10	24.32 (0.12)	22.25 (0.45)	
Total Cost (TC) ($)			
Avg. TC 2011	$858 ± 3	$6,608 ± 80	<.0001
Avg. TC 2012	$936 ± 4	$7,043 ± 90	<.0001
Avg. RX cost 2011	$353 ± 2	$1,342 ± 22	<.0001
Avg. RX cost 2012	$367 ± 2	$1,255 ± 26	<.0001
Avg. Dental cost 2011	$44 ± 0	$63 ± 4	<.0001
Avg. Dental cost 2012	$48 ± 0	$70 ± 3	<.0001
Healthcare Utilization			
Total inpatient length of stay (days)	0.59 (0.01)	16.31 (0.30)	<.0001
Lifestyle risk factors			
BMI (kg/m^2)	24.02 ± 0.01	24.32 ± 0.03	<.0001
Pack-years smoked	7.82 ± 0.04	8.31 ± 0.17	<.0001
Physical activity (days/week)	1.47± 0.01	1.34 ± 0.02	<.0001
Chronic Conditions			
Chronic condition count	2.26 ± 0.01	4.87 ± 0.03	<.0001
Chronic low back pain (%)	33.32 (0.13)	58.93 (0.53)	<.0001
Hypertension (%)	25.75 (0.12)	38.55 (0.53)	<.0001
Osteoarthrosis (%)	15.92 (0.1)	37.14 (0.52)	<.0001
Severe vision reduction (%)	11.22 (0.09)	30.95 (0.5)	<.0001
Lipid metabolism disorders (%)	8.06 (0.08)	10.78 (0.33)	<.0001
Prostatic hyperplasia (%)	6.07 (0.07)	15.88 (0.39)	<.0001
Thyroid dysfunction (%)	4.58 (0.06)	8.06 (0.29)	<.0001
Neuropathies (%)	3.9 (0.06)	11.11 (0.34)	<.0001
Cancers (%)	3.03 (0.05)	22.04 (0.45)	<.0001
Chronic ischemic heart disease (%)	3.23 (0.05)	15.16 (0.39)	<.0001
Cerebral ischemia/ Chronic (%)	2.14 (0.04)	11.69 (0.35)	<.0001
Hemorrhoids (%)	2.25 (0.04)	4.01 (0.21)	<.0001
Depression (%)	1.45 (0.03)	6.05 (0.26)	<.0001
Severe hearing loss (%)	1.34 (0.03)	2.92 (0.18)	<.0001

Rheumatoid arthritis/Chronic (%)	1.22 (0.03)	3.84 (0.21)	<.0001
Cardiac arrhythmias (%)	1.09 (0.03)	3.36 (0.19)	<.0001
Somatoform disorders (%)	0.82 (0.03)	2.2 (0.16)	<.0001
Dementia (%)	0.45 (0.02)	2.93 (0.18)	<.0001
Renal insufficiency (%)	0.28 (0.02)	3.1 (0.19)	<.0001
Cardiac insufficiency (%)	0.28 (0.02)	1.59 (0.13)	<.0001
Parkinson's disease (%)	0.17 (0.01)	1.82 (0.14)	<.0001

Values are presented as mean ± standard error or percentages with standard error in parentheses.

* p Values are obtained by Chi-square test/t-test/Wilcoxon rank sum test.

Table 2 Oral health measures of current high-cost patients

Patient characteristics	Non high cost	High cost	p Value*
Oral Condition (%)			
Periodontal disease	25.89 (0.12)	28.62 (0.49)	<.0001
Oral health behavior (%)			
1. Frequency of toothbrushing			<.0001
≤1 time per day	8.26 (0.08)	10.95 (0.34)	
2 times per day	42.27 (0.14)	45.45 (0.54)	
≥3 times per day	49.48 (0.14)	43.6 (0.54)	
2. Use of floss/interdental brush	7.2 (0.1)	5.64 (0.1)	<.0001
Self-reported oral symptom (%)			
1.Tongue or inside-cheek pain	6.40 (0.07)	7.45 (0.22)	<.0001
2. Difficulty in enunciation due to teeth/denture/gum conditions			<.0001
Yes	5.45 (0.07)	8.49 (0.31)	
Slightly	46.47 (0.14)	48.69 (0.55)	
No	48.08 (0.14)	42.82 (0.54)	
3. Difficulty in chewing food due to teeth/denture/gum conditions	16.57 (0.11)	22.00 (0.45)	<.0001

Values are presented as percentages with standard error in parentheses.
* p Values are obtained by Chi-square test.

Table 3 Odds ratios for future high-cost patients

Patient characteristics	Univariate OR (95% CI)	Model 1[a] Multivariate OR	Model 2[b] Multivariate OR	Model 3[c] Multivariate OR	Model 4[d] Multivariate OR
Oral Condition					
Periodontal disease	1.13 (1.08-1.18)***	1.10 (1.05-1.16)***	1.11 (1.06-1.16)***	1.03 (0.98-1.09)	1.02 (0.96-1.07)
Oral health behavior					
Toothbrushing (times/day)					
≤1	1.47 (1.37-1.57)***	1.07 (0.99-1.15)	1.02 (0.94-1.10)	1.06 (0.98-1.15)	1.05 (0.97-1.14)
2	1.22 (1.17-1.27)***	1.07 (1.02-1.12)**	1.04 (0.99-1.08)	1.05 (1.00-1.10)*	1.04 (0.99-1.10)
≥3	Reference				
Use of floss/interdental brush	1.29 (1.23-1.34)***	1.07 (1.03-1.12)**	1.05 (1.01-1.1)*	1.08 (1.04-1.14)**	1.06 (1.01-1.12)**
Self-reported oral symptom					
Tongue or inside-cheek pain	1.20 (1.13-1.28)***	1.22 (1.14-1.3)***	1.20 (1.13-1.28)***	1.13 (1.05-1.21)***	1.11 (1.04-1.20)**
Difficult in enunciation due to teeth/denture/gum conditions					
Yes	1.74 (1.61-1.89)***	1.30 (1.2-1.41)***	1.09 (1.04-1.14)**	1.21 (1.11-1.32)***	1.15 (1.05-1.26)**
Slightly	1.23 (1.18-1.28)***	1.10 (1.05-1.15)***	1.27 (1.17-1.38)***	1.08 (1.03-1.13)***	1.08 (1.03-1.13)**
No	Reference				
Difficult in chewing food due to teeth/denture/gum conditions	1.34 (1.28-1.41)***	1.12 (1.07-1.19)***	1.11 (1.05-1.17)**	1.09 (1.03-1.16)**	1.06 (1.00-1.12)*

Data are odds ratio (95% confidence interval).

a Adjustment for age, gender and income

b Adjustment for age, gender, income and lifestyle risk factors (BMI, , pack-years smoked, and physical activity)

c Adjustment for age, gender, income and chronic disease

d Adjustment for age, gender, income, chronic disease, prior total healthcare cost and LOS

*p<0.05; **p<0.01; ***p<0.001

eSUPPLEMENT

- Data count by year:

Data count \ Year	2011	2012	2013
NHIS - national health screening cohort	487,835	483,421	478,740
High cost patients (top 10%)	48,784	48,342	47,874
Medical and Dental check-up participants	76,269	77,689	74,263

- Sample size by year:

	2,011	2012	Total
Sample size	76,269	55,280	131,549

Reporting checklist for cohort study.

Based on the STROBE cohort guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cohort reporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	2
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	#3	State specific objectives, including any prespecified hypotheses	2
Study design	#4	Present key elements of study design early in the paper	2
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4

Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up.	4
	#6b	For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	5-6
Bias	#9	Describe any efforts to address potential sources of bias	4
Study size	#10	Explain how the study size was arrived at	4
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	4-6
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	6
	#12b	Describe any methods used to examine subgroups and interactions	4-6
	#12c	Explain how missing data were addressed	6
	#12d	If applicable, explain how loss to follow-up was addressed	N/A
	#12e	Describe any sensitivity analyses	N/A
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	N/A
	#13b	Give reasons for non-participation at each stage	N/A
	#13c	Consider use of a flow diagram	N/A

Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	6-9
	#14b	Indicate number of participants with missing data for each variable of interest	5-6
	#14c	Summarise follow-up time (eg, average and total amount)	2
Outcome data	#15	Report numbers of outcome events or summary measures over time. Give information separately for exposed and unexposed groups if applicable.	2,5
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6-9
	#16b	Report category boundaries when continuous variables were categorized	6-9
	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	6-9
Key results	#18	Summarise key results with reference to study objectives	9-10
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	10
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	9-11
Generalisability	#21	Discuss the generalisability (external validity) of the study results	4
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	12

BMJ Open

Oral health of high-cost patients and evaluation of oral health measures as predictors for high-cost patients in South Korea: a population-based cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-032446.R1
Article Type:	Research
Date Submitted by the Author:	06-Aug-2019
Complete List of Authors:	Kim, Yeonkook; Chungbuk National University,
Primary Subject Heading:	Health informatics
Secondary Subject Heading:	Health services research
Keywords:	High-cost patients, oral health, oral health behavior, self-reported oral symptoms

Title: Oral health of high-cost patients and evaluation of oral health measures as predictors for high-cost patients in South Korea: a population-based cohort study

Yeonkook Joseph Kim, Ph.D.[1]

Author affiliation

[1] Chungbuk National University, Cheongju, Chungbuk, Rep. of Korea.

Correspondence to Dr. Y. Joseph Kim; yjk15@snu.ac.kr

Keywords: High-cost patients, oral health, oral health behavior, self-reported oral symptoms

Word count: 3,398

ABSTRACT

Objective: To examine the oral health conditions and oral health behavior of high-cost patients and evaluate oral health measures as predictors of future high-cost patients.

Design: A retrospective, population-based cohort study using administrative healthcare records.

Setting: The National Health Insurance Service (NHIS) medical check-up database (a.k.a. NHIS - national health screening cohort database) in South Korea.

Participants: 131,549 individuals who received biennial health check-ups including dental check-ups in 2011 or 2012, aged 49–88.

Primary outcome measures: Current and subsequent year high-cost patient status.

Results: High-cost patients, on average, incur higher dental costs, suffer more from periodontal disease, brush their teeth less and use secondary oral hygiene products less. Some of the self-reported oral health behaviors and oral symptom variables show statistically significant associations with subsequent year high-cost patient indicators, even after adjusting for demographic, socio-economic, medical conditions, and prior healthcare cost and utilization.

Conclusions: We demonstrate that oral health measures are associated with an increased risk of becoming a high-cost patient.

Strengths and limitations of this study

▪ This is the first study to provide empirical evidence of the oral health characteristics of high-cost patients.

▪ This study is conducted using a large-scale health check-up cohort database.

▪ We compare current patient characteristics and oral health conditions of high-cost patients with those of the non-high-cost group.

▪ We evaluate six oral health measures as predictors for the following year high-cost patient status using multivariate logistic regression models.

▪ This study is limited by one-year prediction period (i.e. we predict the outcome in 2012 using the data from 2011 and we predict the outcome in 2013 using the data from 2012).

INTRODUCTION

Studies have shown that a small portion of the population uses the majority of healthcare resources, highlighting the importance of predicting high-cost patients by recognizing the characteristics and patterns of their healthcare utilization for targeted interventions and for the modelling of patient risk factors for more equitable healthcare reimbursements.[1] If certain patient characteristics are predictive of highly persistent use, it may be possible to offer more cost-effective alternatives to frequent primary care visits, including disease management, case management, group visits, and patient education.[2] Wammes et al. conducted an extensive systematic review of the characteristics and healthcare utilization of high-cost patients.[3] However, to the best of our knowledge, the oral health characteristics of high-cost patients have not been studied.

Poor oral health has been shown to be associated with increased risk of various diseases and mortality.[4–7] For example, studies show associations between periodontal disease and a higher risk of Alzheimer's disease, coronary heart disease, rheumatoid arthritis, glomerulonephritis, inflammatory bowel disease and prostate cancer.[8–10] Other studies show that poor oral hygiene is associated with hypertension, diabetes mellitus, cardiovascular disease, and head and neck cancers.[11–14] Still other studies have found a positive association between poor oral health and mortality.[4,7,15]

Given that poor oral health is associated with many of chronic and severe diseases, we conjecture that high-cost patients have poor oral health. To test this hypothesis, using a large-scale health check-up database in Korea, we analyze the oral health of high-cost patients. One main reason to study the characteristics of high-cost patients is to predict future high-cost patients for targeted interventions. We evaluate oral health measures as predictors in the high-cost patient predictions. In short, our objective is to examine the oral health conditions and oral health behaviors of current high-cost patients and to evaluate which oral health measures identify future high-cost patients.

MATERIALS AND METHODS

Participants and setting

We analyze the National Health Insurance Service (NHIS) medical check-up database (also known as the NHIS - national health screening cohort (NHIS-HEALS)) of Korea.[16] The NHIS is a uniform insurance policy administrator that covers all residents of Korea. The NHIS datasets represent the entire population in Korea. Therefore, this database is a suitable source of information for nationwide population-based studies.[17]

To construct the NHIS-HEALS database, a sample cohort was first selected from the 2002 and 2003 health screening participants (aged between 40 and 79 in 2002) and followed up through 2013.[16] In 2002, this cohort has 514,866 health check-up participant, which is a 10% random sample of the total number of health check-up participants in 2002 and 2003. The NHIS-HEALS database consists of four databases: an eligibility database, a medical check-up database, a claims database, and a healthcare provider database. To use the most recent data in the database, we construct the study database with a medical check-up data set that contains information on examinees who received biennial medical and dental check-ups in 2011 or 2012.

The medical check-up data is generated through the National Health Screening Program (NHSP), which is a free national medical check-up program for NHIS beneficiaries who are 40 or older.[18] NHIS enrollees eligible for the NHSP are required to have a medical check-up biennially. The NHSP consists of a set of laboratory tests, a dental examination, and questionnaires on self-reported health behavior. Our study sample consists of 131,549 enrollees who participated in the NHSP in 2011 or 2012 aged from 49 to 88. For those enrollees who participated in both years, we only use 2011 observation to avoid duplication.[18] Since the majority of high cost patients are 50 years and older in Korea and many other countries, we use a large sample of that population segment in this study.[1,3,19] Using a join key in the database, we extract their claims and eligibility information from the claims and eligibility datasets, respectively, and merge them with the medical check-up information (For more information on how to use the database, please visit the NHIS website: nhiss.nhis.or.kr).

Patient involvement

A de-identified population healthcare data is used in this study. Thus, patients were not involved in the development of the research question, the outcome measures or the study design.

Statistical methods

Variables

Following prior studies, we define a 'high-cost' patient as an individual in a sample who is in the upper decile of annual healthcare expenditures.[1,3] We create a binary indictor for high-cost patients using all of the data in the database. For example, the medical check-up database contains 487,835 enrollees in 2011, and we tag enrollees who are in the upper decile of annual healthcare expenditures as high-cost patients. See online supplementary file1 for further details.

To understand the oral health conditions of high cost patients, we compare the characteristics and health behavior of these patients with those of a non-high-cost group. We compare demographic factors (gender and age), socioeconomic (income level) information, healthcare expenditures, and the total inpatient length of stay (LOS) as a measure of healthcare utilization.

We examine the three lifestyle risk factors of body mass index (BMI), pack-years smoked, and physical activity, all of which are associated with future healthcare expenditures.[20] Pack-years smoked values are calculated by multiplying the number of packs of cigarettes smoked per day by the number of years the person has smoked. The physical activity measure is from a questionnaire asking about the number of days a check-up participant exercises at a moderate level for more than 30 minutes in activities such as fast walking, playing tennis, and riding a bicycle.

Prior studies point to a high prevalence of multiple chronic conditions to explain high-cost patients' high utilization rates.[3] We examine the chronic conditions of high-cost patients. Disease information in our claims dataset is coded based on the International Classification of Diseases-Tenth Revision (ICD-10) coding scheme.

We select six oral health measures from the claims data, the dental check-up and questionnaire data in the database based on findings from prior studies (first three variables) and on the statistical significance of the association with high-cost patient status (next three variables). The first three variables are the prevalence of periodontal diseases, the frequency of toothbrushing, and the use of floss or interdental brushes. The last three are self-reported oral symptoms: tongue or inside-cheek pain (yes/no), difficulty in enunciation due to teeth/denture/gum conditions (yes, slightly, or no) and difficulty in chewing food due to teeth/denture/gum conditions (yes/no). We exclude participants with missing data for age (n=0; 0%), BMI (n= 3,184; 2.42%), pack-years smoked (n= 3,749; 2.85%), physical activity (n=3,336; 2.54%), toothbrushing frequency (n=1,305; 0.99%), use of floss/interdental brushes (n=554; 0.42%), tongue or inside-cheek pain (n=472; 0.36%), difficulty in enunciation (n=4,042; 3.07%) and difficulty in chewing food (n=3,511; 2.67%) from the corresponding analyses.

Statistical Analyses

We conduct a chi-square test and a t-test to compare the characteristics and health behaviors of the high-cost patient group with those of the non-high-cost group in years 2011 and 202. Because healthcare expenditure variables exhibit marked positive skewness, with a few high-cost patients and many low- or zero-expenditure healthcare users, we perform both a t-test and a Wilcoxon rank sum test.

To evaluate oral health measures as predictors of high-cost patients, we measure the associations between oral health measures and future high-cost patient status (i.e., a binary indictor of high-cost patient status in the subsequent year) using multivariate logistic regression models. We predict the outcome in 2012 using the data from 2011, and we predict the outcome in 2013 using the data from 2012. We select covariates in our models based on prior high-cost patient and healthcare cost prediction studies.[1,20–22] We develop four different models. The first model adjusts for three confounders: age, gender and income. For the second model, we add the three lifestyle risk factors of BMI, pack-years smoked, and physical activity. For the third model, we adjust for chronic conditions. Wammes et al.

point to a high prevalence of multiple (chronic) conditions to explain high-cost patients' utilization.[3] We add 46 major chronic disease indicators selected in van den Busschet et al. to the first model.[23] Since some of these chronic disease indicators show high associations, we use the stepwise feature-selection procedure in SAS version 9.4 (SAS Institute, Cary, NC, USA) to eliminate insignificant and collinear variables. For the last model, we add prior healthcare cost and LOS.

RESULTS

The general characteristics of the high-cost patients are presented in Table 1. The percentage of high-cost patients in our sample is 6.52%, as people who go to health check-ups tend to be younger and less likely to be high-cost patients. Consistent with prior studies, high costs are associated with increasing age.[3] The mean age for the non-high-cost group is 58, whereas the mean for the high-cost patient group is 64 in our sample data. As income increases, the likelihood of becoming a high-cost patient decreases. Enrollees in medical aid programs do not report income, and a much higher proportion of high-cost patients are in medical aid programs compared to those in the non-high-cost group (1.5% vs 0.24%). 7.35% of total high-cost patients are in the lowest income decile, whereas 6.45% of the non-high-cost group are in the same decile. On the other hand, 22.25% of total high-cost patients are in the highest income decile, whereas 24.32% of the non-high-cost group are in the same decile. The average total cost for the high-cost patient group in 2011 is $6,608, almost eight times higher than that for the non-high-cost group. The total average drug cost for high-cost patients is about 3.8 times higher than that for the non-high-cost group as of 2011. High-cost patients also incurred 1.4 times more, on average, in dental costs compared to the non-high-cost group. With regard to healthcare expenditure variables, we performed both a t-test and a Wilcoxon rank sum test and found statistically significant differences at the 0.1% level for both tests.

Inset [Table 1] here.

Next, we report three lifestyle risk factors for high-cost patients: BMI, pack-years smoked and physical

activity. Consistent with the result in Leigh et al., high-cost patients have higher BMIs (24.32 vs. 24.02), smoke more (8.31 vs. 7.82) and exercise less (1.34 days vs.1.47 days).[20]

We report the average count of chronic diseases for high-cost patients in Table 1. On average, high-cost patients have 4.87 chronic conditions while those in the non-high-cost group have 2.26. For example, more than half of our high-cost patients suffer from chronic lower back pain. Another common chronic disease for high-cost patients in our sample is hypertension. Approximately 39% of high-cost patients have hypertension, while 26% in the non-high-cost group suffer from the same disease. The prevalence of cancer is much higher among high-cost patients than among non-high-cost patients (22.04 % vs 3.03 %). Chronic ischemic heart disease is another chronic disease with a large disparity: 15.16% for high-cost patients vs. 3.23% for the non-high-cost group.

While this disparity of chronic disease conditions has been reported in numerous prior studies of high-cost patients, the oral health characteristics of high-cost patients have not been reported.[3] Table 2 presents the oral health condition, oral health behaviors and the three aforementioned self-reported oral symptoms.

Inset [Table 2] here.

First, we report the prevalence of periodontal diseases for the high-cost patients in Table 2. We use the ICD 10 code (K05) from the claims data to identify those who were treated for periodontal diseases. The prevalence of periodontal diseases is higher in the high-cost patient group than in the non-high-cost group (28.6 % vs 25.9%). The first oral health behavior variable is the frequency of toothbrushing per day. Higher proportions of high-cost patients report one time or less per day or two times per day, while a higher proportion of the non-high-cost group reports three times or more per day. The second oral health behavior is the use of floss or/and interdental brushes. Again, high-cost patients use secondary oral products less: 5.64% for the high-cost patient group and 7.2% for the non-high-cost group.

We report three self-reported oral symptoms. High-cost patients are more likely to report tongue or

inside-cheek pain: 7.45% for the high-cost patient group and 6.40% for the non-high-cost group. The second symptom is difficulty in enunciation due to teeth/denture/gum conditions. A higher proportion of high-cost patients report this symptom compared to those in the non-high-cost group: 8.49% of high-cost patients reported this symptom, while 5.45% of the non-high-cost group did so. The third symptom is difficulty in chewing food due to teeth/denture/gum conditions, with 22% of the high-cost patient group reporting this symptom and 16.57% of the non-high-cost group doing so.

Table 3 presents the result of the binominal logistic regression analyses. Poor oral health measures are associated with a higher likelihood of a person becoming a high-cost patient in the subsequent year. The univariate odds ratios for all of the oral health measures are significant at the 0.1% level. Even after we adjust for age, gender and income level in the first model, all oral health measures have odds ratios greater than one, showing that poor oral health increases the odds of becoming a high-cost patient in the following year. In the second model, where we additionally adjust for the three lifestyle risk factors of BMI, pack-years smoked, and physical activity, the odds ratios for oral health behavior become less significant. The odds ratio of the frequency of toothbrushing becomes statistically insignificant and the odds ratio for the use of floss or interdental brushes becomes less significant in the second model. However, the odds ratio for the prevalence of periodontal disease (OR=1.11 and 95% CI 1.06 to 1.16) and those for the self-reported symptoms remain statistically significant in the second model.

Inset [Table 3] here.

To adjust for a high prevalence of multiple chronic conditions in high-cost patients, we adjust for chronic conditions in our third model. When chronic disease variables are introduced in the model, the odds ratio for periodontal disease becomes statistically insignificant. To our final model, we add prior total healthcare cost and LOS, as they are powerful predictors of future high-cost patients.[1] The odds ratios for the variables of use of floss or interdental brushes, tongue or inside-cheek pain, and difficulty in enunciation are significant at the 1% level.

DISCUSSION

In the present study, we show a positive association between poor oral health and high-cost patient status. High-cost patients have a higher prevalence of periodontal diseases (28.6 % vs 25.9%). They are more likely to brush their teeth less than three times a day and are less likely to use floss or interdental brushes than the non-high-cost group. This positive association can partly be explained by prior studies that show positive associations between periodontal disease and various chronic and severe diseases.[8–10] Other studies also find that the frequency of toothbrushing and the use of secondary oral hygiene products are associated with diseases such as hypertension, diabetes mellitus, cardiovascular disease, and head and neck cancers.[11–14] In a recent study, researchers found that flossing and brushing of interdental spaces may reduce the risk for new cardiovascular events among patients with coronary heart disease.[24] Moreover, our results show that the high-cost patient group is more likely to report tongue or inside-cheek pain (7.45% vs. 6.40%), difficulty in enunciation due to teeth/denture/gum conditions and difficulty in chewing food due to teeth/denture/gum conditions (22% vs. 16.57%) as compared to the non-high-cost group.

The prevalence of periodontal diseases shows a positive association with future high-cost patient status. However, when we adjust for chronic conditions, the statistical significance of the odds ratio disappears. The associations between the oral health behavior variables and future high-cost status are also statistically significant. However, the associations become less statistically significant when lifestyle risk factors are introduced in the model. This is due to the association between oral health behaviors and lifestyle risk factors, and the result is consistent with those in prior studies.[6,25,26] Kim et al. show that an increased frequency of toothbrushing and the number of secondary oral products used are associated with a lower BMI, less smoking and a higher level of physical activity.[6]

The associations between self-reported oral symptoms and future high-cost patient status remain statistically significant even after we adjust for demographic and socioeconomic measures, lifestyle risk behaviors, chronic diseases, and prior healthcare costs and utilization rates. This shows that self-

reported oral symptoms are a potential new source of data for high-cost prediction modelling.

Strengths and limitation

To the best of our knowledge, this is the first study to focus on the oral health of high-cost patients. Prior studies of high-cost patients and healthcare cost predictions indicate a high prevalence of multiple (chronic) conditions and lifestyle risk factors.[3,20] However, the oral health characteristics of high-cost patients have not been studied. Our study fills this void in the literature using a large sample of 131,549 enrollees. This study, however, does not explain or imply a causal relationship between oral health and the likelihood of becoming a high-cost patient. As Griffin et al. point out, oral diseases can have an impact on many aspects of general health, and health conditions can in turn have an impact on oral health.[27] Another limitation of this study is that we do not consider the long-term effects of oral health on healthcare expenditures. However, one-year prediction model is the most common study design in high-cost patient prediction literature.[3]

Policy and research implications

There are several implications pertaining to our work. Wammes et al. argue that high-cost patients make up the sickest and most complex populations.[3] They are a small portion of the population yet use the majority of healthcare resources. Accurately identifying high-cost patients and managing their care is a significant first step in improving quality levels and reducing population health costs.[1] Actively exploring data sources available for their identification and prediction is a requisite for achieving these goals. To this end, our analysis provides a new data source for high-cost prediction modelling. Moreover, we have shown that people with poor oral health are at an increased risk of becoming high-cost patients. In that oral and other chronic and severe diseases also share common risk factors, it is important to examine the interplay between these diseases and oral disease as well as their combined impact when health policymakers develop programs involving targeted interventions for high-cost patients and develop preventive measures at the population level.[27]

In conclusion, we demonstrate that oral health measures are associated with the risk of becoming a high-cost patient. Our results highlight the impact of oral health on healthcare costs and support the development of preventive measures at the population level.

Acknowledgments I would like to thank Ms. Boram Park and Ms. Kyung-Ran Lee at the NHIS for providing answers to numerous questions on the NHIS medical check-up database used in this study.

Author contributions YJK conceptualized and designed the study, collected and analyzed the data, and wrote the manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None.

Ethics approval The Seoul National University Institutional Review Board.

Data sharing statement Information about accessing study data is available at: https:// nhiss.nhis.or.kr.

References

1. Tamang S, Milstein A, Sørensen HT, et al. Predicting patient ' cost blooms ' in Denmark : a longitudinal population-based study. *BMJ Open*. 2017;7(1):e011580. doi:10.1136/bmjopen-2016-011580
2. Naessens JM, Baird MA, Van Houten HK, Vanness DJ, Campbell CR. Predicting persistently high primary care use. *Ann Fam Med*. 2005;3(4):324-330. doi:10.1370/afm.352
3. Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open*. 2018;8(9):e023113. doi:10.1136/bmjopen-2018-023113
4. Adolph M, Darnaud C, Thomas F, et al. Oral health in relation to all-cause mortality: The IPC cohort study. *Sci Rep*. 2017;7(October 2016):3-8. doi:10.1038/srep44604
5. Sabbah W, Mortensen LH, Sheiham A, Batty D. Oral health as a risk factor for mortality in middle-aged men: The role of socioeconomic position and health behaviours. *J Epidemiol Community Health*. 2013;67(5):392-397. doi:10.1136/jech-2012-201340
6. Kim Y-H, Kim D-H, Lim KS, et al. Oral health behaviors and metabolic syndrome: the 2008–2010 Korean National Health and Nutrition Examination Survey. *Clin Oral Investig*. 2014;18(5):1517-1524. doi:10.1007/s00784-013-1112-2
7. Joshy G, Arora M, Korda RJ, Chalmers J, Banks E. Is poor oral health a risk marker for incident cardiovascular disease hospitalisation and all-cause mortality? Findings from 172 630 participants from the prospective 45 and Up Study. *BMJ Open*. 2016;6(8):e012386. doi:10.1136/bmjopen-2016-012386
8. Galgut PN. Periodontal disease and poor health outcomes. *Bmj*. 2010;340(jun17 1):c2735-c2735. doi:10.1136/bmj.c2735
9. Lee JH, Kweon HHI, Choi JK, Kim YT, Choi SH. Association between periodontal disease and prostate cancer: results of a 12-year longitudinal cohort study in South Korea. *J Cancer*. 2017;8(15). doi:10.7150/jca.20532
10. Gaur S, Agnihotri R. Alzheimer's disease and chronic periodontitis: Is there an association? *Geriatr Gerontol Int*. 2015;15(4):391-404. doi:10.1111/ggi.12425
11. Kuwabara M, Motoki Y, Ichiura K, et al. Association between toothbrushing and risk factors for cardiovascular disease: A large-scale, cross-sectional Japanese study. *BMJ Open*. 2016;6(1). doi:10.1136/bmjopen-2015-009870
12. de Oliveira C, Watt R, Hamer M. Toothbrushing, inflammation, and risk of cardiovascular disease: results from Scottish Health Survey. *Bmj*. 2010;340(may27 1):c2451-c2451. doi:10.1136/bmj.c2451
13. Farquhar DR, Divaris K, Mazul AL, Weissler MC, Zevallos JP, Olshan AF. Poor oral health affects survival in head and neck cancer. *Oral Oncol*. 2017;73(July):111-117. doi:10.1016/j.oraloncology.2017.08.009
14. Darnaud C, Thomas F, Pannier B, Danchin N, Bouchard P. Oral health and blood pressure: The IPC cohort. *Am J Hypertens*. 2015;28(10):1257-1261. doi:10.1093/ajh/hpv025
15. Kim JK, Baker LA, Davarian S, Crimmins E. Oral health problems and mortality. *J Dent Sci*. 2013;8(2):115-120. doi:10.1016/j.jds.2012.12.011

16. Seong SC, Kim YY, Park SK, et al. Cohort profile: The National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. *BMJ Open*. 2017;7(9). doi:10.1136/bmjopen-2017-016640

17. Lee J, Lee JS, Park S-H, Shin SA, Kim K. Cohort Profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. *Int J Epidemiol*. 2016:dyv319. doi:10.1093/ije/dyv319

18. Kim HB, Lee SA, Lim W. Knowing is not half the battle: impacts of information from the national health screening program in Korea. *IZA Discuss Pap*. 2017;(10650).

19. Kim ST, Seong SC. *National Health Insurance Statistical Yearbook*. Seoul: Health Insurance Review & Assessment Service, National Health Insurance Service; 2015.

20. Leigh JP, Hubert HB, Romano PS. Lifestyle risk factors predict healthcare costs in an aging cohort. *Am J Prev Med*. 2005;29(5):379-387. doi:10.1016/j.amepre.2005.08.005

21. Fleishman JA, Cohen JW. Using information on clinical conditions to predict high-cost patients. *Health Serv Res*. 2010;45(2):532-552. doi:10.1111/j.1475-6773.2009.01080.x

22. Leininger LJ, Friedsam D, Voskuil K, DeLeire T. Predicting high-need cases among new Medicaid enrollees. *Am J Manag Care*. 2014;20(9):e399-407. http://www.ncbi.nlm.nih.gov/pubmed/25364876.

23. Bussche H van den. Which chronic diseases and disease combinations are specific to multimorbidity in the elderly? Results of a claims data based cross-sectional study in Germany. *BMC Public Health*. 2011;18(4):322-332. doi:10.1159/000122413

24. Reichert S, Schlitt A, Beschow V, et al. Use of floss/interdental brushes is associated with lower risk for new cardiovascular events among patients with coronary heart disease. *J Periodontal Res*. 2015;50(2):180-188. doi:10.1111/jre.12191

25. Tada A, Matsukubo T. Relationship Between Oral Health Behaviors and General Health Behaviors in a Japanese Adult Population. *J Public Health Dent*. 2003;63(4):250-254. doi:10.1111/j.1752-7325.2003.tb03508.x

26. Kim H-J, Kim Y-H, Cho K-H, et al. Oral health behaviors and bone mineral density in South Korea: the 2008–2010 Korean National Health and Nutrition Examination Survey. *J Bone Miner Metab*. 2016;34(2):225-233. doi:10.1007/s00774-015-0669-z

27. Griffin SO, Jones JA, Brunson D, Griffin PM, Bailey WD. Burden of oral disease among older adults and implications for public health priorities. *Am J Public Health*. 2012;102(3):411-418. doi:10.2105/AJPH.2011.300362

Table 1 Current high-cost patient characteristics (from years 2011 and 2012)

Patient characteristics	Non high cost	High cost	p Value*
N	122,974	8,575	
%	93.48%	6.52%	
% female	40.8%	47.6%	
Age (years)	58.45 ± 0.02	63.83 ± 0.09	<.0001
Income decile (%)			<.0001
Medical Aid	0.24 (0.01)	1.5 (0.13)	
1	6.45 (0.07)	7.35 (0.28)	
2-5	24.61 (0.12)	23.25 (0.46)	
6-9	44.39 (0.14)	45.64 (0.54)	
10	24.32 (0.12)	22.25 (0.45)	
Total Cost (TC) ($)			
Avg. TC 2011	$858 ± 3	$6,608 ± 80	<.0001
Avg. TC 2012	$936 ± 4	$7,043 ± 90	<.0001
Avg. RX cost 2011	$353 ± 2	$1,342 ± 22	<.0001
Avg. RX cost 2012	$367 ± 2	$1,255 ± 26	<.0001
Avg. Dental cost 2011	$44 ± 0	$63 ± 4	<.0001
Avg. Dental cost 2012	$48 ± 0	$70 ± 3	<.0001
Healthcare Utilization			
Total inpatient length of stay (days)	0.59 (0.01)	16.31 (0.30)	<.0001
Lifestyle risk factors			
BMI (kg/m^2)	24.02 ± 0.01	24.32 ± 0.03	<.0001
Pack-years smoked	7.82 ± 0.04	8.31 ± 0.17	<.0001
Physical activity (days/week)	1.47± 0.01	1.34 ± 0.02	<.0001
Chronic Conditions			
Chronic condition count	2.26 ± 0.01	4.87 ± 0.03	<.0001
Chronic low back pain (%)	33.32 (0.13)	58.93 (0.53)	<.0001
Hypertension (%)	25.75 (0.12)	38.55 (0.53)	<.0001
Osteoarthrosis (%)	15.92 (0.1)	37.14 (0.52)	<.0001
Severe vision reduction (%)	11.22 (0.09)	30.95 (0.5)	<.0001
Lipid metabolism disorders (%)	8.06 (0.08)	10.78 (0.33)	<.0001
Prostatic hyperplasia (%)	6.07 (0.07)	15.88 (0.39)	<.0001
Thyroid dysfunction (%)	4.58 (0.06)	8.06 (0.29)	<.0001
Neuropathies (%)	3.9 (0.06)	11.11 (0.34)	<.0001
Cancers (%)	3.03 (0.05)	22.04 (0.45)	<.0001
Chronic ischemic heart disease (%)	3.23 (0.05)	15.16 (0.39)	<.0001
Cerebral ischemia/ Chronic (%)	2.14 (0.04)	11.69 (0.35)	<.0001
Hemorrhoids (%)	2.25 (0.04)	4.01 (0.21)	<.0001
Depression (%)	1.45 (0.03)	6.05 (0.26)	<.0001
Severe hearing loss (%)	1.34 (0.03)	2.92 (0.18)	<.0001

Rheumatoid arthritis/Chronic (%)	1.22 (0.03)	3.84 (0.21)	<.0001
Cardiac arrhythmias (%)	1.09 (0.03)	3.36 (0.19)	<.0001
Somatoform disorders (%)	0.82 (0.03)	2.2 (0.16)	<.0001
Dementia (%)	0.45 (0.02)	2.93 (0.18)	<.0001
Renal insufficiency (%)	0.28 (0.02)	3.1 (0.19)	<.0001
Cardiac insufficiency (%)	0.28 (0.02)	1.59 (0.13)	<.0001
Parkinson's disease (%)	0.17 (0.01)	1.82 (0.14)	<.0001

Values are presented as mean ± standard error or percentages with standard error in parentheses.

* p Values are obtained by Chi-square test/t-test/Wilcoxon rank sum test.

Table 2 Oral health measures of current high-cost patients

Patient characteristics	Non high cost	High cost	p Value*
Oral Condition (%)			
Periodontal disease	25.89 (0.12)	28.62 (0.49)	<.0001
Oral health behavior (%)			
1. Frequency of toothbrushing			<.0001
≤1 time per day	8.26 (0.08)	10.95 (0.34)	
2 times per day	42.27 (0.14)	45.45 (0.54)	
≥3 times per day	49.48 (0.14)	43.6 (0.54)	
2. Use of floss/interdental brush	7.2 (0.1)	5.64 (0.1)	<.0001
Self-reported oral symptom (%)			
1.Tongue or inside-cheek pain	6.40 (0.07)	7.45 (0.22)	<.0001
2. Difficulty in enunciation due to teeth/denture/gum conditions			<.0001
Yes	5.45 (0.07)	8.49 (0.31)	
Slightly	46.47 (0.14)	48.69 (0.55)	
No	48.08 (0.14)	42.82 (0.54)	
3. Difficulty in chewing food due to teeth/denture/gum conditions	16.57 (0.11)	22.00 (0.45)	<.0001

Values are presented as percentages with standard error in parentheses.
* p Values are obtained by Chi-square test.

Table 3 Odds ratios for future high-cost patients

Patient characteristics	Univariate OR (95% CI)	Model 1[a] Multivariate OR	Model 2[b] Multivariate OR	Model 3[c] Multivariate OR	Model 4[d] Multivariate OR
Oral Condition					
Periodontal disease	1.13 (1.08-1.18)***	1.10 (1.05-1.16)***	1.11 (1.06-1.16)***	1.03 (0.98-1.09)	1.02 (0.96-1.07)
Oral health behavior					
Toothbrushing (times/day)					
≤1	1.47 (1.37-1.57)***	1.07 (0.99-1.15)	1.02 (0.94-1.10)	1.06 (0.98-1.15)	1.05 (0.97-1.14)
2	1.22 (1.17-1.27)***	1.07 (1.02-1.12)**	1.04 (0.99-1.08)	1.05 (1.00-1.10)*	1.04 (0.99-1.10)
≥3	Reference				
Use of floss/interdental brush	1.29 (1.23-1.34)***	1.07 (1.03-1.12)**	1.05 (1.01-1.1)*	1.08 (1.04-1.14)**	1.06 (1.01-1.12)**
Self-reported oral symptom					
Tongue or inside-cheek pain	1.20 (1.13-1.28)***	1.22 (1.14-1.3)***	1.20 (1.13-1.28)***	1.13 (1.05-1.21)***	1.11 (1.04-1.20)**
Difficult in enunciation due to teeth/denture/gum conditions					
Yes	1.74 (1.61-1.89)***	1.30 (1.2-1.41)***	1.09 (1.04-1.14)**	1.21 (1.11-1.32)***	1.15 (1.05-1.26)**
Slightly	1.23 (1.18-1.28)***	1.10 (1.05-1.15)***	1.27 (1.17-1.38)***	1.08 (1.03-1.13)***	1.08 (1.03-1.13)**
No	Reference				
Difficult in chewing food due to teeth/denture/gum conditions	1.34 (1.28-1.41)***	1.12 (1.07-1.19)***	1.11 (1.05-1.17)**	1.09 (1.03-1.16)**	1.06 (1.00-1.12)*

Data are odds ratio (95% confidence interval).

a Adjustment for age, gender and income

b Adjustment for age, gender, income and lifestyle risk factors (BMI, , pack-years smoked, and physical activity)

c Adjustment for age, gender, income and chronic disease

d Adjustment for age, gender, income, chronic disease, prior total healthcare cost and LOS

*p<0.05; **p<0.01; ***p<0.001

eSUPPLEMENT

- Data count by year:

Data count \ Year	2011	2012	2013
NHIS - national health screening cohort	487,835	483,421	478,740
High cost patients (top 10%)	48,784	48,342	47,874
Medical and Dental check-up participants	76,269	77,689	74,263

- Sample size by year:

	2,011	2012	Total
Sample size	76,269	55,280	131,549

Reporting checklist for cohort study.

Based on the STROBE cohort guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cohort reporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	2
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	2
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	#3	State specific objectives, including any prespecified hypotheses	2
Study design	#4	Present key elements of study design early in the paper	2
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4

Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up.	4
	#6b	For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	5-6
Bias	#9	Describe any efforts to address potential sources of bias	4
Study size	#10	Explain how the study size was arrived at	4
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	4-6
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	6
	#12b	Describe any methods used to examine subgroups and interactions	4-6
	#12c	Explain how missing data were addressed	6
	#12d	If applicable, explain how loss to follow-up was addressed	N/A
	#12e	Describe any sensitivity analyses	N/A
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	N/A
	#13b	Give reasons for non-participation at each stage	N/A
	#13c	Consider use of a flow diagram	N/A

Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	6-9
	#14b	Indicate number of participants with missing data for each variable of interest	5-6
	#14c	Summarise follow-up time (eg, average and total amount)	2
Outcome data	#15	Report numbers of outcome events or summary measures over time. Give information separately for exposed and unexposed groups if applicable.	2,5
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	6-9
	#16b	Report category boundaries when continuous variables were categorized	6-9
	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	6-9
Key results	#18	Summarise key results with reference to study objectives	9-10
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	10
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	9-11
Generalisability	#21	Discuss the generalisability (external validity) of the study results	4
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	12