PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association between COL11A1 (rs1337185) and ADAMTS5 (rs162509) Gene Polymorphisms and Lumbar Spine Pathologies in Chinese Han Population: an observational study
AUTHORS	Jiang, Hua; Yang, Qinghua; Jiang, Jie; Zhan, Xinli; Xiao, Zengming

VERSION 1 - REVIEW

REVIEWER	Guo Xiong Key Laboratory of Trace Elements and Endemic Diseases of Ministry of Health, School of Public Health, Health Science Center, Xi'an Jiaotong University, Xi'an, P. R. China.
REVIEW RETURNED	23-Jan-2017

GENERAL COMMENTS	The author conducted case-control study to validate the association of previously reported SNPs for lumbar disc degeneration. Association between environmental risk factors and specific lumbar spine pathologies were also evaluated. The language used is obscure and there are grammar issues and ambiguities throughout the manuscript. In addition, what has been listed in the table should not be rephrased too much in the text. Page5 line 139 The author stated that “400 normal control subjects were gender-matched adult” How did the 400 controls matched to 428 patients? The authors have not report the predefined thresholds of HW test. Page7 line198-200 The author stated that subgroup 3 has a higher job physical demand than other three subgroups. This may give the false implications that every combination of two subgroups have been tested for difference. However, since only one P-value was reported, I assumed the each subgroup is only tested for difference with the control group. While the subgrouping is an interesting topic, it may also lead to smaller sample sizes in each subgroups and I am worrying if the pathology-specific association still remains pathology-specific in larger sample size. Page10, line 283 The author stated that “as a large family of metalloproteases, ADAMTS-5 has”. Please check the expression. Page11 line 294-296 The author stated that “Therefore, it remains plausible that COL11A1 and ADAMTS5 genes may be involved in the etiology of LDD through interruption of the synthesis and metabolism balance of extracellular matrix.” It is not proper to propose such a hypothesis based on the context of the manuscript.

	Page12 321-323 The author stated that “The C allele of rs1337185 was risky for patients affected by the lumbar pathologies including disc herniation only, stenosis or/and spondylolisthesis. Please check the expression.

REVIEWER	John Loughlin Newcastle University Institute of Genetic Medicine UK
REVIEW RETURNED	27-Jan-2017

GENERAL COMMENTS	The authors have undertaken a genetic association analysis of six candidate SNPs in a cohort of patients with lumbar spine pathologies. They have studied the cohort as a whole and following subgroup stratification into four phenotypes. The case sample size overall is 428 patients and the control sample size is 400. The investigators have also tested a panel of environmental risk factors. The SNPs chosen came from a 2015 publication involving an Indian case-control cohort analysis. In the Discussion the authors introduce us to additional associations studies of the SNPs and of different SNPs but in the same genes. Reviewer Comments 1) Firstly, the sample size (as acknowledged by the authors) is small for a case-control study of what one presumes to be a multifactorial, polygenic trait. The positive p-values are consequently modest. Are the authors confident that their results are therefore robust and are not type 1 errors? 2) The authors have studied 6 SNPs in five phenotypes (unstratified and stratified four times). Would their p-values withstand multiple testing? 3) Although the ethnicity of the populations studied in this report (Han Chinese) and the manuscript that inspired it (an Indian cohort) are different, the fact that the authors considered the SNPs worthy of testing clearly suggests that they think there may be common genetic routes to the lumbar spine pathologies between the two ethnic groups. Would it not be worthwhile therefore to include meta-analyses? There are other relevant cohorts in which some of their tested SNPs have also been genotyped. A large scale meta-analysis may provide statistically more convincing results. 4) Related to the above, the authors do not explicitly state whether the allele at rs1337185 that shows associtiion in the Han Chinese is the same allele as in the Indian study. Are the associations in the same direction? This also applies to the rs162509 SNP. 5) Why were far more of the cases subjected to MRI and CT (78%) than the controls (16%). I presume it was part of the required clinical assessment of the cases? Does it impact on the ability of the investigators to confirm absence of disease phenotype in their controls? 6) The authors provide quite sparse details about the evaluation of environmental risk factors. For example, the hour/day spent driving or as a passenger. Is this in the last month, year, since childhood, etc? 7) If the four subgroup phenotypes that the authors split their cases in to are 'mutually exclusive', what justification is there for studying the cases as a whole? And why exactly are they mutually exclusive?

VERSION 1 – AUTHOR RESPONSE

Response to reviewer #1:
Question 1 (Q1): The author conducted case-control study to validate the association of previously reported SNPs for lumbar disc degeneration. Association between environmental risk factors and specific lumbar spine pathologies were also evaluated. The language used is obscure and there are grammar issues and ambiguities throughout the manuscript. In addition, what has been listed in the table should not be rephrased too much in the text.
Response 1 (R1): Thank you very much for your valuable comments. According to your advice, the inaccurate statements have been corrected in the revised manuscript. The English writing of this manuscript was polished by American Journal Experts.

Q2: Page 5 line 139 The author stated that “400 normal control subjects were gender-matched adult” How did the 400 controls matched to 428 patients?
R2: Thank you for your careful review of the manuscript. The male to female ratio in the lumbar disc degeneration group is 226/202 (1.1), and that in the control group is 224/176 (1.2). This modification can be seen in Line 139-141 on Page 5.

Q3: The authors have not report the predefined thresholds of HW test.
R3: Thank you very much for your advice. The predefined thresholds of the HW test were added to the genetic association analysis section. Please refer to Line 215-216 on Page 8.

Q4: Page 7 line 198-200 The author stated that Subgroup 3 has a higher job physical demand than other three subgroups. This may give the false implications that every combination of two subgroups have been tested for difference. However, since only one P-value was reported, I assumed the each subgroup is only tested for difference with the control group. While the subgrouping is an interesting topic, it may also lead to smaller sample sizes in each subgroups and I am worrying if the pathology-specific association still remains pathology-specific in larger sample size.
R4: According to your suggestion, the physical workload was compared in the four subgroups using an analysis of variance. The result showed that Subgroup 3 had a higher physical workload than did other three subgroups (all $P<0.01$). There was no statistical significance among Subgroups 1, 2 and 4 ($P<0.05$). This correction can be seen in Line 205-207 on Page 8. It is generally recognized that lumbar disc degeneration (LDD) is a complex and multifactorial disease. The definition of LDD is based on the various types of MRI features, such as loss of signal intensity, narrowing of disc space, osteophyte formation, annular tears, endplate sclerosis and disc bulging. In this study, we classified the LDD patients into four subgroups by using the highly selective radiographic features. This approach allowed us to subgroup patients according to the specific lumbar spine pathologies that have a pre-defined clinical significance. However, stratification by pathological subgroups reduced the number of subjects who were compared, which may result in a weakened statistical power. We acknowledge that this is a limitation of our study and would like to enlarge the group of cases in future work to confirm our results in a larger cohort of subjects. This information has been added as the limitation of our study, which can be seen in Line 338-342 on Page 13.

Q5: Page 10 line 283 The author stated that “as a large family of metalloproteases, ADAMTS-5 has”. Please check the expression.
R5: Thank you very much for your comments. The correction can be seen in Line 302 on Page 11.

Q6: Page 11 line 294-296 The author stated that “Therefore, it remains plausible that COL11A1 and ADAMTS5 genes may be involved in the etiology of LDD through interruption of the synthesis and metabolism balance of extracellular matrix.” It is not proper to propose such a hypothesis based on the context of the manuscript.
R6: Thank you for your advice. A modification can be seen in the discussion section. Please refer to Line 313-315 on Page 12.

Q7: Page 12 line 321-323 The author stated that “The C allele of rs1337185 was risky for patients affected by the lumbar pathologies including disc herniation only, stenosis or/and spondylolisthesis. Please check the expression.
R7: The correction can be seen in the discussion section. Please refer to Line 348-350 on Page 13.

Response to reviewer #2:
Question 1 (Q1): Firstly, the sample size (as acknowledged by the authors) is small for a case-control study of what one presumes to be a multifactorial, polygenic trait. The positive p-values are consequently modest. Are the authors confident that their results are therefore robust and are not type 1 errors?
Response 1 (R1): Thank you for your comments. We acknowledge that a limitation of this study is its sample size. An increase in the number of cases and controls will be necessary to strengthen the statistical power and substantiate differences. In the future, we are interested in enlarging the group of cases and controls to confirm these results in a larger cohort of subjects. This information was added as a limitation of our study, which can be seen in Line 338-342 on Page 13.

Q2: The authors have studied 6 SNPs in five phenotypes (unstratified and stratified four times). Would their p-values withstand multiple testing?
R2: According to your suggestion, Bonferroni correction was used for p-value adjustment by R Software 3.3.2 for Windows. After Bonferroni correction, rs1337185 and rs162509 showed significant associations with disc herniation, and rs1337185 remained significant association with stenosis and spondylolisthesis. However, there was no significant interaction between rs162509 and patients with degenerative scoliosis after Bonferroni correction. The modifications can be seen in the abstract, statistical analysis, genetic association analysis, and discussion section. Please refer to Line 53, 54, 56 on Page 2, Line 177-180 on Page 7, Line 227-229, 235, 237-239 on Page 9, Line 280-284 on Page 10-11.

Q3: Although the ethnicity of the populations studied in this report (Han Chinese) and the manuscript that inspired it (an Indian cohort) are different, the fact that the authors considered the SNPs worthy of testing clearly suggests that they think there may be common genetic routes to the lumbar spine pathologies between the two ethnic groups. Would it not be worthwhile therefore to include meta-analyses? There are other relevant cohorts in which some of their tested SNPs have also been genotyped. A large scale meta-analysis may provide statistically more convincing results.
R3: Thanks for your valuable comments. We completely agree with your opinion that a meta-analysis could provide more credible information by improving the statistical power of the association analysis. A computerized literature search was conducted using the “PubMed” database on February 21, 2017. The association of rs1337185 and disc degeneration was reported in two studies (Refs. 17, 23), and the associations of other SNPs (rs5275, rs5277, rs7575934, rs3213718 and rs162509) and disc degeneration were only tested in the Indian cohort. Unfortunately, neither study provided detailed information regarding the genotype and allele frequencies of these SNPs. We could not conduct a meta-analysis due to a lack of original data in these studies. Although a meta-analysis was not included in our study, we believe that your advice will undoubtedly benefit our following studies. This information is added as the limitation of our study, which can be seen in Line 342-344 on Page 13.

Q4: Related to the above, the authors do not explicitly state whether the allele at rs1337185 that shows associtio in the Han Chinese is the same allele as in the Indian study. Are the associations in the same direction? This also applies to the rs162509 SNP.
R4: We reviewed the paper of the Indian cohort. However, no detailed results were reported for the risk or protective alleles in rs1337185 and rs162509. Due to the lack of original data, we did not evaluate the difference in genetic associations between the Han Chinese population and Indian population. This information has been added in Line 284-287 on Page 11.

Q5: Why were far more of the cases subjected to MRI and CT (78%) than the controls (16%). I presume it was part of the required clinical assessment of the cases? Does it impact on the ability of the investigators to confirm absence of disease phenotype in their controls?
R5: Thank you for your careful review. In control group, 336 (336/400, 84%) of the adults were subjected to MRI, and 64 (64/400, 16%) of the adults were subjected to CT. Conversely, all patients underwent an MRI scan in the lumbar disc degeneration group. This modification can be seen in the subjects and data collection section. Please refer to Line 141-143 on Page 5-6.

Q6: The authors provide quite sparse details about the evaluation of environmental risk factors. For example, the hour/day spent driving or as a passenger. Is this in the last month, year, since childhood, etc.?
R6: The correction can be seen in the environmental factors evaluation section. Please refer to Line 155-159 on Page 6.

Q7: If the four subgroup phenotypes that the authors split their cases in to are 'mutually exclusive', what justification is there for studying the cases as a whole? And why exactly are they mutually exclusive?
R7: Thank you very much for your comments. The causes and pathogenesis of lumbar disc degeneration (LDD) are poorly understood. LDD seems to result from the combined insult of environmental and genetic factors. Regardless of the initiating causative factor, the final common pathway leading to LDD is end-plate damage. The term LDD as a general phenotype has been used synonymously with a multitude of terminologies, such as disc dehydration, reduction in disc height, discopathies, disc bulge, disc herniation, spinal stenosis, spondylolisthesis, osteophyte formation, and scoliosis. For these reasons, defining special phenotypes for disc degeneration-related studies is challenging. We selected four highly selective phenotypes that could be clearly diagnosed based on clinical and radiological findings and evaluated previous candidate gene studies for LDD. This may help us identify the role played by each of the SNPs in their association with the various disc pathologies.
1. Martin MD, Boxell CM, Malone DG. Pathophysiology of lumbar disc degeneration: a review of the literature. Neurosurg Focus 2002;13(2):E1.
2. Kao PY, Chan D, Samartzis D, et al. Genetics of lumbar disk degeneration: technology, study designs, and risk factors. Orthop Clin North Am 2011;42(4):479-86, vii.
3. Rajasekaran S, Kanna RM, Senthil N, et al. Phenotype variations affect genetic association studies of degenerative disc disease: conclusions of analysis of genetic association of 58 single nucleotide polymorphisms with highly specific phenotypes for disc degeneration in 332 subjects. Spine J 2013;13(10):1309-20.
4. Battie MC, Lazary A, Fairbank J, et al. Disc degeneration-related clinical phenotypes. Eur Spine J 2014;23 Suppl 3:S305-14.
5. Nakki A, Battie MC, Kaprio J. Genetics of disc-related disorders: current findings and lessons from other complex diseases. Eur Spine J 2014;23 Suppl 3:S354-63.

VERSION 2 – REVIEW

REVIEWER	Xiong Guo Key Laboratory of Trace Elements and Endemic Diseases of National Health and Family Planning Commission, School of Public Health, Health Science Center, Xi'an Jiaotong University
REVIEW RETURNED	16-Mar-2017

GENERAL COMMENTS	Thank you for replying the issues.

REVIEWER	Professor John Loughlin Newcastle University, UK.
REVIEW RETURNED	29-Mar-2017

GENERAL COMMENTS	Thank you for the changes made. I think the manuscript is much improved.