PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

(This paper received three reviews from its previous journal but only two reviewers agreed to published their review.)

ARTICLE DETAILS

TITLE (PROVISIONAL)	Living with disabling chronic pain: results from a face-to-face cross-sectional population-based study.
AUTHORS	Cabrera-León, Andrés; Cantero-Braojos, Miguel Angel; Garcia-Fernandez, Llenalia; Guerra de Hoyos, Juan Antonio

VERSION 1 – REVIEW

REVIEWER	Failde I University of Cadiz. Spain
REVIEW RETURNED	18-Dec-2017

GENERAL COMMENTS	The study includes information on secondary data of chronic pain and disability obtained from the Andalusian Health Survey that has not been frequently analyzed in other surveys. Despite the interest of this work, it presents important limitations because it aims to estimate Spanish Disabling Chronic Pain (DCP) prevalence by applying calibration adjustments, but using several variables from the Andalusia survey. The calibration method seems to be more appropriate to improve the precision of the parameters estimated from samples in that same population, but it is not so clear that it is adequate to estimate the Spanish prevalence from a sample of the Andalusian population. Although the population of Andalusia is an important part of the Spanish geography, it has sociocultural and economic characteristics that are different from other areas in the country, so the auxiliary variables used in the calibration (sex, age, educational level and employment status) may not be adequate. Specifically, the unemployment rate in Andalusia is higher than in other areas of Spain, and this factor can affect the prevalence estimated. On the other hand, it does not seem justified that the authors do not include Fibromyalgia (FM) and arthritis or rheumatism in the definition of CP, even though they are mentioned in the discussion as frequent causes of chronic pain. In the methodology the authors indicate that they have carried out a factorial analysis for the environmental quality ítems, but do not provide details about the procedure they have used. In the results section, the information corresponding to Figure 4a

	should be located below, as done for Figure 4b. This would facilitate the understanding of the comparisons described in the text. Paragraph 6 of the discussion is confusing and does not seem to be a justification for the results observed, especially the absence of difference between patients with DCP and without CP.

REVIEWER	Domenico Plantone ASL VCO - Ospedale San Biagio - Domodossola (VB), Italy
REVIEW RETURNED	14-Mar-2018

GENERAL COMMENTS	The paper is interesting and well written. The tables are really useful and add important details. I would just improve the discussion on the relation between depression/mood disorders and pain. From a neurological point of view, pain and depression are entwined in a complex relationship of situational and physiological connections that are not yet fully understood. [Chopra K, Arora V (2014)]. Depression has been identified as a predictor of pain and chronic pain is linked to depression in the general population. It has been highlighted that the pres- ence or severity of one condition cross-amplifies the other. At the same time, it is not easy to distinguish between disability due to chronic pain and disability due to other medical condition. I would better discuss also this aspect.

REVIEWER	Steven P. Cohen Johns Hopkins, USA
REVIEW RETURNED	23-Mar-2018

GENERAL COMMENTS	The authors have performed a large, cross-sectional study evaluating the prevalence of chronic pain, and chronic disabling pain, in Spanish adults. The relatively large size and face-to-face interviews are a plus, though the cross-sectional nature has its limitations (poor recall, overemphasis on recent events). It would also be helpful to know how representative this sample is of "Spanish adults" (including whether it includes undocumented immigrants), and why Spanish adults would be different than other populations studied (e.g. Kurita et al. Pain 2012 (Denmark), Walters et al. A & A 2017 (India and Nepal), Voerman et al. Eur J Pain 2015 (Dutch adolescents), in addition to the examples provided in the discussion). Downsides include that "disabling" in this context is subjective, and that the type of pain (or duration) was not identified (or at least not reported). Other comments 1. The abstract should stand alone, and therefore OR's and CI's for major findings would be helpful (inadequate sleep is also not a habit). 2. Please provide some information on how long the interview lasted, and the 32% who "didn't respond" (i.e. did they refuse, were they not home, was any person in the household OK (e.g. the man or woman), or was a certain person 'targeted'? 3. It is not clear whether the prevalence rates here estimate point, annual or lifetime (unlikely the latter).

	4. Table 1: There does not seem to be any reference to any type of neuropathic pain in row 2. 5. Any information on how "chronic" was defined (the IASP defines it as pain that persists longer than normal healing, while others use a specific cutoff such as 3 months). 6. In table 2, the overall p-values are reported, but we don't know where the differences lie as many sections have multiple breakdowns. 7. Was the "interview" structured in any way (notwithstanding table 1). 8. In the discussion, I would suggest "sectionalizing" parts. 9. Discussion, para 2, 1st sentence ("The simplest example..."): Please rewrite. 10. In figure 2 (as well as the discussion), the authors note that certain chronic diseases such as fibromyalgia and arthritis) are higher in the DCP population than the non-CP population. However, no one without pain will be diagnosed with these condition, so the CP prevalence should be 100% (who would get x-rays to diagnose arthritis without pain?). 11. I would discuss generalization, remark how this could affect practice or health initiatives, and areas for further study.

REVIEWER	Martine J. Barons University of Warwick, UK
REVIEW RETURNED	09-Apr-2018

GENERAL COMMENTS	This study investigates the prevalence, comorbidities and demographic factors associated with those experiencing disabling chronic pain, which the authors define as disability with diagnoses of illness with ‘pain’ in the name. This is a well-written article. The statistical analyses are appropriate and comparators appropriate. Page 5: please include exclusion criteria. Page 7 and following. Some of the diseases listed as comorbidities, such as osteoporosis and arthritis are known to be painful despite not having pain in the name. Indeed, some of the pain included in the list may be caused by these diseases. You address this briefly in passing on page 14. Please give fuller explanation why such diseases to not amount to pain. Please do sensitivity analyses where appropriate including / excluding the disease from the definition of DCP. Page 7 Table 1 Likert scale (spelling) Page 9 Lower educational attainment influences employment status and type of work. Please comment on the interaction of these factors. Page 14 Could nDCP lead to DCP in time? Please comment on this in relation to higher age in DCP, referring the disability rates between age groups. Page 14 Osteoporosis is more common in females. To what extent does this explain multimorbidity frequency in women? Page 15 Depression may be caused by disability with chronic pain. The good point about the need for psychosocial services should be repeated in the summary.

REVIEWER	David Steinsaltz Department of Statistics University of Oxford United Kingdom
REVIEW RETURNED	15-Apr-2018

GENERAL COMMENTS	Overall, I think the design and the analysis make sense. I do have some concerns. Choice of model: The categories here are clearly ordered — nDCP is clearly between DCP and nCP — but multinomial logistic regression does not take this into account, at least, not as usually defined. If not, there will be a substantial loss of power, and possible bias. Also, it makes it hard to interpret the results in the way that seems to be desired, as describing factors responsible for the distinction between disabling and non-disabling chronic pain. Thus, for example, age appears as a factor predicting DCP, but not nDCP. This has a different interpretation if (as I understand it) the model treats these as two completely separate categories than if DCP is seen as an intensification of CP from nDCP. But it seems to me that the conclusions that are being drawn — and that one naturally would want to draw — are based on the latter interpretation. The factor analysis (called here “factorial analysis”) is not described in any detail. It’s not clear what or how many covariates are being summarised, or how the choice of two factors was made. Model diagnostics: I’d like to see a bit more detail about this. The procedure of “re-entering” variables whose removal changes the other estimates by more than 30% seems unusual to me. I understand the idea, but I’m not sure exactly how it’s supposed to work, and the only citation for it is a general article on confounding that doesn’t seem to mention logistic regression or this particular correction specifically. It’s not clear how potential collinearity of covariates was tested for. Model presentation: Table 2 presents some results of the model fitting, but the presentation is not very clear. Is this supposed to represent a final model? It includes variables “with $p< 0.2$”, but the “significance level” is stated to be 0.05. And Social Class is included, with $p=0.68$. Or are these just the p-values of the initial fit, before removing non-significant factors. Sampling design: I’m not an expert on sampling, but (or, perhaps, therefore) I would have liked to have seen an explanation of the claimed design effect. I would think it would depend on a number of assumptions/ estimates. I’m also not sure what the “$p=q=0.5$” refers to. The statement is made that “ the sample design was considered throughout”, but it looks as though the confidence intervals in Supplementary Tables 1 and 2 have not been corrected for the design effect. For example, with 6207 individuals sampled and no design effect, the width of a 95% confidence interval for sex would be 2.48%, and the stated confidence interval has width 2.4%. In Figure 4, and elsewhere, the presentation of numbers was inconsistent with regard to significant figures. Given the size of the study, it seems to me that 2 places after the decimal point would be appropriate. The order of limits of the confidence intervals in Figure 4 are also inconsistent — sometimes the lower bound is first, sometimes the upper.)

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

Reviewer Name: Failde I
Institution and Country: University of Cadiz. Spain
Please state any competing interests or state 'None declared': None declared

Comments to the Author:
The study includes information on secondary data of chronic pain and disability obtained from the Andalusian Health Survey that has not been frequently analyzed in other surveys.

COMMENT 1: You use a stratified random design, but the statistics section does not discuss weighting of data in the analysis. Despite the interest of this work, it presents important limitations because it aims to estimate Spanish Disabling Chronic Pain (DCP) prevalence by applying calibration adjustments, but using several variables from the Andalusia survey. The calibration method seems to be more appropriate to improve the precision of the parameters estimated from samples in that same population, but it is not so clear that it is adequate to estimate the Spanish prevalence from a sample of the Andalusian population.

RESPONSE 1: Thank you for your comment. We considered both the sample design and the sample weights throughout the statistical analysis (descriptive, bivariate and multivariate analysis). We used the advanced sampling module of SPSS as well as an approximation of sampling with replacement. Individual case weight was used to adjust for municipality's population [36] following the method described in the Andalusian Health Survey [19]. Comments about this were included in the last two paragraphs of the Statistical analysis section.

With respect to the Spanish prevalence of DCP, calibration adjustments can be also used to extrapolate the estimations of a survey, as in the present article, with the weightings providing sample 'estimates' for the totals of the auxiliary variables that match known population totals for these variables (Särndal CE. The calibration approach in survey theory and practice. Surv Methodol 2007; 33:99–119). Furthermore, calibration adjustments increased, not only the accuracy of the estimation for the Spanish DCP prevalence, but also the validity of its generalization from a smaller area, i.e. Andalusia, to a larger one, i.e. Spain. This approach is being increasingly adopted in social sciences, especially in studies with small samples or with coverage and non-response problems. We discussed that in the manuscript "Cabrera-Leon A, Lopez-Villaverde V, Rueda M, Moya-Garrido MN. Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study. Human Reproduction; 2015; 30(11):2677-85." [30].

We have added a (marked) text with the previous comment in the ninth paragraph of the discussion section, and the following new phrase and two references at the end of the first paragraph of the Statistical analysis section:

The 'sampling' R package was used for the sample design and calibration weightings in estimations of DCP prevalence, and 'samplingVarEst' package for its variance estimation.

[31] Tillé Y, Matei A. R Package sampling: survey sampling 2015:76. https://cran.r-project.org/web/packages/sampling/sampling.pdf (accessed Aug 2016).

[32] Escobar-Lopez E, Barrios-Zamudio E. SamplingVarEst: Sampling Variance Estimation 2015. https://cran.r-project.org/web/packages/samplingVarEst/samplingVarEst.pdf (accessed 9 Aug 2016).

COMMENT 2: Although the population of Andalusia is an important part of the Spanish geography, it has sociocultural and economic characteristics that are different from other areas in the country, so the auxiliary variables used in the calibration (sex, age, educational level and employment status) may not be adequate. Specifically, the unemployment rate in Andalusia is higher than in other areas of Spain, and this factor can affect the prevalence estimated.

RESPONSE 2: Thank you for your comment. It is true, Spain and Andalusia have different sociocultural and economic characteristics. That is the reason why we considered not only sex and age as calibration variables, but also educational level and employment status. All those auxiliary variables are considered in the new calibrated weights which include information both from the Andalusian sample and the Spanish census. Thus, the extrapolated prevalence of DCP from Andalusia to Spain is representative, at least, for all those variables. We discussed results about bias in direct population estimations for the calibration variables with respect to their census values in the publication "Cabrera-León A, Rueda M, Cantero-Braojos M. Calibrated prevalence of disabling chronic pain: a face-to face cross-sectional population study in Southern Spain. BMJ Open 2017; 7:e014033; doi:10.1136/bmjopen-2016-014033". [59]

In addition, the fact that those variables were associated with the study variable (DCP), provided better results in terms of accuracy and validity of the estimations. That is shown in the multivariate model. We also discussed it in the previous publication [59].

To make this clearer, we have added (marked) texts in the ninth paragraph of the Discussion section.

COMMENT 3: On the other hand, it does not seem justified that the authors do not include Fibromyalgia (FM) and arthritis or rheumatism in the definition of CP, even though they are mentioned in the discussion as frequent causes of chronic pain.

RESPONSE 3: Thank you for your comment. The survey does not gather information about pain due to other conditions. However, it does gather information about a wide range of chronic conditions (Table 1), which have been crossed with the study variable DCP. That way, the relationship of chronic pain with specific chronic diseases (e.g. fibromyalgia, arthritis, Figure 1) as well as with multimorbidity (Figure 2) was studied.

Pain is the first cause to explore in the diagnosis of certain chronic diseases such as fibromyalgia or arthritis. But these diseases can appear in a state of remission and the interviewee can report no pain during the last three months, although they reported arthritis, arthrosis, fibromyalgia, ... In that sense, Brown et al. (2008)) [45] reported an apparent dissociation between clinical remission and continuous structural deterioration in rheumatoid arthritis. In addition, Hannan, Felson, Pincus (2000)) [46] observed that only 47% of people with stage 2-4 of radiographic arthropathy reported knee pain, and 61% reported that their doctor had diagnosed arthrosis. Brinjikji et al. (Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol. 2015;36(4):811-6) reported similar results, imaging findings of spinal degeneration are frequently present in asymptomatic individuals, and that result increases with age. Many degenerative features based on images are probably part of the natural process of ageing and they are not associated with pain, although they have been diagnosed at some point. These imaging findings should be interpreted in the context of the patient's clinical condition. Some recent findings in the scientific literature call into question the current diagnostic criteria for chronic pain. Moreover, there are new contributions related to pain such as central sensitization (Beattie et al, 2005, Woolf, 2007, Lee, Nassikas, Clauw, 2011, Finan et al., 2012). That is why it is so important to consider chronic pain as a disease in itself (European Federation IASP Chapters, FIC´S, International Association for the Study of Pain. EFIC´S

Declaration on chronic pain as a major healthcare problem a disease in its own right. 2004), as well as to consider those chronic diseases which presumably produce pain as multimorbidities.

COMMENT 4: In the methodology the authors indicate that they have carried out a factorial analysis for the environmental quality items, but do not provide details about the procedure they have used.

RESPONSE 4: The enviromental variables considered in the Factor Analysis are described in Table 1. They were related to noise, smell, air pollution, industry, green areas, delinquency/insecurity and heavy traffic. They were measured on a scale from 1 to 3, where 1 is “a lot” and 3 is “nothing”. Categorical Principal Component Analysis (CATPCA) was performed and two components were chosen because, as a general rule, the eigenvalue for a dimension should be larger than 1. The proportion of variance explained by the loadings was 61%.

We have included the reference in Table 1 when Factor Analysis is referred to in the Statistical section. We have also added a new text in Table 1 based on the previous paragraph.

COMMENT 5: In the results section, the information corresponding to Figure 4a should be located below, as done for Figure 4b. This would facilitate the understanding of the comparisons described in the text.

RESPONSE 5: Thank you. Both Figures (now 4 and 5) share the same information located below them.

COMMENT 6: Paragraph 6 of the discussion is confusing and does not seem to be a justification for the results observed, especially the absence of difference between patients with DCP and without CP.

RESPONSE 6: Thank you for your comment. We believe you are referring to the absence of differences between patients without DCP (i.e. nDCP) and without CP. As a result, to make this clearer, we have rewritten the corresponding phrase in paragraph 6 of the discussion.

We hope our responses and the changes made address your comments.

Reviewer 2:

Reviewer Name: Domenico Plantone
Institution and Country: ASL VCO - Ospedale San Biagio - Domodossola (VB), Italy
Please state any competing interests or state ‘None declared’: None declared

Comments to the Author:
The paper is interesting and well written. The tables are really useful and add important details.

COMMENT 1: I would just improve the discussion on the relation between depression/mood disorders and pain. From a neurological point of view, pain and depression are entwined in a complex relationship of situational and physiological connections that are not yet fully understood. [Chopra K, Arora V (2014)]. Depression has been identified as a predictor of pain and chronic pain is linked to depression in the general population. It has been highlighted that the presence or severity of one condition cross-amplifies the other.

RESPONSE 1: Thank you for your comment. We have made changes in the seventh paragraph of the Discussion section and added the Chopra and Arora reference with the following information: “We consider these results with caution for two reasons. Firstly, from a neurological point of view, pain and depression interact in a complex relationship of situational and physiological connections that is not yet fully understood [53]. Secondly, depression and anxiety were measured together in our study, through the same variable. Despite this, the association between DCP and those mental disorders highlights the need for psychosocial services in chronic pain management [37]”.

COMMENT 2: At the same time, it is not easy to distinguish between disability due to chronic pain and disability due to other medical condition. I would better discuss also this aspect.

RESPONSE 2: Thank you very much for you comment. The definition of disability includes people who were limited in their activity by each one of the reported chronic diseases (Table 1). In that sense, people interviewed were asked about disability related to each chronic disease (including chronic pains). Therefore, the disability is due to the corresponding chronic disease or chronic pain, not to other medical condition.

To make this clearer, we have added (marked) texts in the Variable section and in the second paragraph of the Discussion section.

We hope this addresses your comment.

Reviewer 3:

Reviewer Name: Steven P. Cohen
Institution and Country: Johns Hopkins, USA
Please state any competing interests or state ‘None declared’: None

Comments to the Author:
The authors have performed a large, cross-sectional study evaluating the prevalence of chronic pain, and chronic disabling pain, in Spanish adults.

COMMENT 1: The relatively large size and face-to-face interviews are a plus, though the cross-sectional nature has its limitations (poor recall, overemphasis on recent events).

RESPONSE 1: We agree with you. We have added your comment and a reference [58] in the ninth paragraph of the Discussion section.

COMMENT 2: It would also be helpful to know how representative this sample is of "Spanish adults" (including whether it includes undocumented immigrants).

RESPONSE 2: Thank you for your comment. Spain and Andalusia have different sociocultural and economic characteristics. That is the reason why we applied calibration adjustments (Särndal CE. The calibration approach in survey theory and practice. Surv Methodol 2007; 33:99–119). These techniques increased not only the accuracy of the estimation for the Spanish DCP prevalence, but also the validity of its generalization from a smaller area, i.e. Andalusia, to a larger one, i.e. Spain. This approach is being increasingly adopted in social sciences, especially in studies with small samples or with coverage and non-response problems. We discussed that in the manuscript “Cabrera-Leon A, Lopez-Villaverde V, Rueda M, Moya-Garrido MN. Calibrated prevalence of infertility in 30- to 49-year-old women according to different approaches: a cross-sectional population-based study. Human Reproduction; 2015; 30(11):2677-85[30].

We considered sex, age, educational level and employment status as calibration variables. All of them are considered in the new calibrated weights which include information both from the Andalusian sample and the Spanish census. Thus, the extrapolated prevalence of DCP from Andalusia to Spain is representative, at least, for all those variables. We discussed results about bias of direct population estimations for the calibration variables with respect to their census values in the publication "Cabrera-León A, Rueda M, Cantero-Braojos M. Calibrated prevalence of disabling chronic pain: a face-to face cross-sectional population study in Southern Spain. BMJ Open 2017; 7:e014033; doi:10.1136/bmjopen-2016-014033"[59].

To make this clearer, we have added (marked) texts in the ninth paragraph of the Discussion section, and a new phrase and two references at the end of the first paragraph of the Statistical analysis section.

COMMENT 3: Why Spanish adults would be different than other populations studied (e.g. Kurita et al. Pain 2012 (Denmark), Walters et al. A & A 2017 (India and Nepal), Voerman et al. Eur J Pain 2015 (Dutch adolescents), in addition to the examples provided in the discussion).

RESPONSE 3: Thank you for your comment. Epidemiological studies related to chronic pain need to research much more cultural and ethnic diversity both in and between countries. According to Peacock and Pate (Peacock S, Pate Sh. Cultural Influences on Pain. Rev Pain, 2008; 1(2): 6–9), pain is a private experience, and social, cultural and psychological factors have an influence on its behaviour, form and on the social environment in which it occurs. The distribution of a health system's services varies depending on the country, and pain treatment can vary depending on the culture and ethnicity. Populations are changing and therefore it is very important to focus on the local health professionals' knowledge of their patients' different ethnic backgrounds. Peacock and Pate reported that immigrants will integrate and receive the provided health care and its experience of pain depending on their level of acculturation. In Spain, the immigrant and ethnic population is growing although not at the same level as other countries, such as Germany, UK, Luxemburg, Austria or Denmark. We are very interested in this area of work and we will consider it seriously in future studies.

COMMENT 4: Downsides include that "disabling" in this context is subjective, and that the type of pain (or duration) was not identified (or at least not reported).

RESPONSE 4: The concept of disability has been aligned with the definition of disability according to WHO (the second paragraph of the Discussion section). Thus, it encompasses impairments, activity limitations and participation restrictions. In relation to impairments (problems in body function/structure), interviewers asked individuals (at home, face-to-face) whether a doctor or a nurse had told them that they suffered from one or more chronic diseases (Table 1). Activity limitation and participation restrictions are constructed as population who declared, when asked about each of the chronic diseases listed, that they were limited in their activity. This is explained in the first paragraph of Variables section.

With respect to the duration of CP, it is reported in the second paragraph of the discussion. The IASP definition [21] is included in the criteria of chronic pain [20]: "The basis of the definition of CP in this study is the medical or healthcare professionals' diagnosis (reports of more than 3 months suffering the chronic disease that included the word 'pain') [20,21]." This was included in the second paragraph of the Discussion section.

To make this clearer, we have added the duration in the definition of CP in Table 1.

COMMENT 5: The abstract should stand alone, and therefore OR's and CI's for major findings would be helpful (inadequate sleep is also not a habit).

RESPONSE 5: Following your comment, we have added the OR´s and CIs in the abstract. In addition, we have also changed the term “inadequate sleep” to “sleeping≤7h”.

COMMENT 6: Please provide some information on how long the interview lasted, and the 32% who "didn't respond" (i.e. did they refuse, were they not home, was any person in the household OK (e.g. the man or woman), or was a certain person 'targeted'?

RESPONSE 6: Thank you for your suggestion. The average time of the interview was 28.84 minutes with a SD of 6.8 and a median of 30 minutes. The 32% no respondent was due to refusal to participate once the household had been contacted.

We have added that text at the end of the first paragraph of the Sample and data collection section.

COMMENT 7: It is not clear whether the prevalence rates here estimate point, annual or lifetime (unlikely the latter).

RESPONSE 7: The prevalence estimations are point, requiring at least 3 months of suffering the pain to be defined as chronic.

COMMENT 8: Table 1: There does not seem to be any reference to any type of neuropathic pain in row 2.

RESPONSE 8: Thank you for your observation. No distinction is made between nociceptive and neuropathic pain in this survey. We have included the following comment at the end of the second paragraph of the Discussion section: “Finally, our study did not gather information to analyse neuropathic, nociceptive or dysfunctional pain because this is not essential information as these entities are considered as different points on the same continuum [38]”.

COMMENT 9: Any information on how "chronic" was defined (the IASP defines it as pain that persists longer than normal healing, while others use a specific cutoff such as 3 months).

RESPONSE 9: Please see our response to comment 4.

COMMENT 10: In table 2, the overall p-values are reported, but we don't know where the differences lie as many sections have multiple breakdowns.

RESPONSE 10: Thank you for your comment. We have made changes in table 2. We hope this addresses your comment.

COMMENT 11: Was the "interview" structured in any way (notwithstanding table 1).

RESPONSE 11: The interview was structured as follows: a first part with items referring to the participant’s house, sex, age, relationship with the other residents and mental or physical limitations of each resident. A second part with questions referring to the quality of the local environment and the condition of the house. The next part included items related to general health, participant’s limitations due to pain, tiredness, physical or emotional problems, chronic diseases, accidents, consumption of

medicine, visits to the general practitioner and specialised doctors as well as general use of Medical care and hospital assistance. A fourth part referred to tobacco and alcohol consumption. The fifth to sleeping and eating habits, and physical activity. The sixth part of the interview referred to medical care such as vaccination, cholesterol level and blood pressure tests, contraceptive methods, solar protection and anthropometric measures. The seventh part covered road safety measures taken by the interviewee such as the use of seat belt. After that, there was a battery of questions for participants over 65 years old related to dependency. The last part concerned socio-demographic characteristics and working conditions; this was followed by the interview's closure in which the date, duration, time and day of week of its realization were recorded.

COMMENT 12: In the discussion, I would suggest "sectionalizing" parts.

RESPONSE 12: Thank for your suggestion. Done.

COMMENT 13: Discussion, part 2, 1st sentence ("The simplest example..."): Please rewrite.

RESPONSE 13: Done.

COMMENT 14: In figure 2 (as well as the discussion), the authors note that certain chronic diseases such as fibromyalgia and arthritis) are higher in the DCP population than the non-CP population. However, no one without pain will be diagnosed with these condition, so the CP prevalence should be 100% (who would get x-rays to diagnose arthritis without pain?).

RESPONSE 14: We agree that pain is the first cause to explore in the diagnosis of the referred diseases. But these diseases can appear in a state of remission and the interviewee can report no pain during the last three months, although they reported arthritis, arthrosis, fibromyalgia,... In that sense, Brown et al. (2008) [45] reported an apparent dissociation between clinical remission and continuous structural deterioration in rheumatoid arthritis. In addition, Hannan, Felson, Pincus (2000) [46] observed that only 47% of people with stage 2-4 of radiographic arthropathy reported knee pain, and 61% reported that their doctor had diagnosed arthrosis. Brinjikji et al. (Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol. 2015 Apr;36(4):811-6. doi: 10.3174/ajnr.A4173) reported similar results: imaging findings of spinal degeneration were frequently present in asymptomatic individuals, and that result increased with age. Many degenerative features based on images are probably part of the natural process of ageing and they are not associated with pain, although they have been diagnosed at some point. These imaging findings should be interpreted in the context of the patient's clinical condition. Some recent findings in the scientific literature call into question the current diagnostic criteria for chronic pain. Furthermore, there are new contributions related to pain such as central sensitization (Beattie et al, 2005, Woolf, 2007, Lee, Nassikas, Clauw, 2011, Finan et al., 2012). That is why it is so important to consider chronic pain as a disease in itself (European Federation IASP Chapters, FIC´S, International Association for the Study of Pain. EFIC´S Declaration on chronic pain as a major healthcare problem a disease in its own right. 2004), as well as to consider those chronic diseases which presumably produce pain as multimorbidities.

We hope this addresses your comment.

COMMENT 15: I would discuss generalization, remark how this could affect practice or health initiatives, and areas for further study.

RESPONSE 15: Following your recommendation, we have made some changes in the last paragraph of the Discussion section with the following text: Therefore, it is a disease that could affect medical practices and political health programmes, as well as future research areas.

Reviewer: 4

Reviewer Name: Martine J. Barons
Institution and Country: University of Warwick, UK
Please state any competing interests or state ‘None declared’: ‘None declared’

Comments to the Author:
This study investigates the prevalence, comorbidities and demographic factors associated with those experiencing disabling chronic pain, which the authors define as disability with diagnoses of illness with ‘pain’ in the name.
This is a well-written article. The statistical analyses are appropriate and comparators appropriate.

COMMENT 1: Please include exclusion criteria.

RESPONSE 1: Thank you for your suggestion. We have included them at the end of the first paragraph of the Design section.

COMMENT 2: Some of the diseases listed as comorbidities, such as osteoporosis and arthritis are known to be painful despite not having pain in the name. Indeed, some of the pain included in the list may be caused by these diseases. You address this briefly in passing on page 14. Please give fuller explanation why such diseases to not amount to pain. Please do sensitivity analyses where appropriate including/excluding the disease from the definition of DCP.

RESPONSE 2: Thank you for your comment. Those diseases can happen in a state of remission and the interviewee can report no pain in the last three months, although they reported osteoporosis, arthritis, arthrosis, fibromyalgia.... In that sense, Brown et al. (2008) [45] reported an apparent dissociation between clinical remission and continuous structural deterioration in rheumatoid arthritis. In addition, Hannan, Felson, Pincus (2000) [46] observed that only 47% of people with stage 2-4 of radiographic arthropathy reported knee pain, and 61% reported that their doctor had diagnosed arthrosis. Brinjikji et al. (Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol. 2015;36(4):811-6) reported similar results, imaging findings of spinal degeneration are frequently present in asymptomatic individuals, and that result increases with age. Many degenerative features based on images are probably part of the natural process of ageing and they are not associated with pain, although they have been diagnosed at some point. These imaging findings should be interpreted in the context of the patient's clinical condition. Some recent findings in the scientific literature call into question the current diagnostic criteria for chronic pain. Furthermore, there are new contributions related to pain such as central sensitization (Beattie et al, 2005, Woolf, 2007, Lee, Nassikas, Clauw, 2011, Finan et al., 2012). That is why it is so important to consider chronic pain as a disease in itself (European Federation IASP Chapters, FIC´S, International Association for the Study of Pain. EFIC´S Declaration on chronic pain as a major healthcare problem a disease in its own right. 2004), as well as to consider those chronic diseases which presumably produce pain as multimorbidities.

With respect to a sensibility analysis, we think it would not be necessary as the diseases mentioned are not included in the definition of DCP. Nevertheless, we analysed prevalences of DCP, nDCP and nCP in people with and without fibromyalgia, osteoporosis and arthritis, obtaining similar prevalences between those diseases. DCP prevalence is around 10% in people without fibromyalgia, osteoporosis or arthritis, nDCP prevalence is 5.7% and nCP prevalence 84.3%. Those prevalences are similar to those from general population. With respect to that population who does have those diseases, 53.9% does not have CP among people with fibromyalgia, 55.9% with arthritis y 46.9% with osteoporosis.

Regarding DCP, results are 39.7%, 37.6% and 49.4%, respectively. And with respect to nDCP, those are 6.4%, 6.5% y 3.7%, respectively.

We hope this addresses your comment.

COMMENT 3: Table 1 Likert scale (spelling)

RESPONSE 3: Thank you. Done.

COMMENT 4: Lower educational attainment influences employment status and type of work. Please comment on the interaction of these factors.

RESPONSE 4: In accordance with your comment, we have added text in the eighth paragraph of the Discussion section.

COMMENT 5: Could nDCP lead to DCP in time? Please comment on this in relation to higher age in DCP, referring the disability rates between age groups.

RESPONSE 5: Thank you very much for your comment. It is true, nDCP could lead to DCP in time, especially in middle age and over. This can be seen in Figure 1 where nDCP prevalence is quite similar in the highest age groups, regardless of gender, while DCP prevalence presents much higher differences. Moreover, as shown in supplementary Table 4, the change over time of nDCP into DCP could be much faster among people with other chronic diseases.

In accordance with the previous paragraph, we have added new text at the end of the third paragraph in the Discussion section.

COMMENT 6: Osteoporosis is more common in females. To what extent does this explain multimorbidity frequency in women?

RESPONSE 6: Thank you for your comment. The prevalence of osteoporosis in men is 0.16% (0.07-0.38%), while it is almost 4% in women (3.96%; 3.35-4.68%). Moreover, the prevalence of osteoporosis in men with DCP is 0.5%, with nDCP 0.8% and without CP 0.1%, while in women it is 12.3% (9.8-15.3%), 1.6% (0.6-4.2%) and 2.4% (1.9-3.1%), respectively. Thus, women with DCP have a much higher prevalence of osteoporosis when compared to women with nDCP or without CP.

As a result, we have rewritten the sixth paragraph in the Discussion section.

COMMENT 7: Depression may be caused by disability with chronic pain. The good point about the need for psychosocial services should be repeated in the summary.

RESPONSE 7: Thank you for your suggestion. Done.

Reviewer: 5

Reviewer Name: David Steinsaltz
Institution and Country: Department of Statistics, University of Oxford, United Kingdom
Please state any competing interests or state ‘None declared’: none declared

Comments to the Author:
Overall, I think the design and the analysis make sense. I do have some concerns.

COMMENT 1: Choice of model: The categories here are clearly ordered — nDCP is clearly between DCP and nCP — but multinomial logistic regression does not take this into account, at least, not as usually defined. If not, there will be a substantial loss of power, and possible bias. Also, it makes it hard to interpret the results in the way that seems to be desired, as describing factors responsible for the distinction between disabling and non-disabling chronic pain. Thus, for example, age appears as a factor predicting DCP, but not nDCP. This has a different interpretation if (as I understand it) the model treats these as two completely separate categories than if DCP is seen as an intensification of CP from nDCP. But it seems to me that the conclusions that are being drawn — and that one naturally would want to draw — are based on the latter interpretation.

RESPONSE 1: Thank you for your comment. There were two reason why ordinal regression was not used. Firstly, the choice between ordinal and multinomial logistic regression depends on having a clearly ordered response. This was not our case, because our response is a classification of a person depending on their limitation and not on an ordered scale from 1 to 3 (e.g. degree of disease). Neither is it a clear degree of disease; the definition and classification of DCP depended on several chronic pains and their limitations on activity. Therefore, not all classified persons within the same category of DCP have the same diseases and so no ordinal scale could be imposed on the different categories. Thus, the response is not truly ordered. Secondly, the assumption of an ordinal logistic regression should satisfy the assumption of proportional odds, this did not hold with our data. So, using a multinomial regression was more appropriate for our data.

We hope this addresses your comment.

COMMENT 2: The factor analysis (called here “factorial analysis”) is not described in any detail. It’s not clear what or how many covariates are being summarised, or how the choice of two factors was made.

RESPONSE 2: The enviromental variables considered in the Factor Analysis are described in Table 1. They were related to noise, smell, air pollution, industry, green areas, delinquency/insecurity and heavy traffic. They were measured on a scale from 1 to 3, where 1 is “a lot” and 3 is “nothing”. Then, Categorical Principal Component Analysis (CATPCA) was performed and two components were chosen because, as a general rule, the eigenvalue for a dimension should be larger than 1. The proportion of variance explained by the loadings was 61%.

We have included the reference in Table 1 when Factor Analysis is referred to in the Statistical section. We have also added a new text in Table 1 based on the previous paragraph.

COMMENT 3: Model diagnostics: I’d like to see a bit more detail about this. The procedure of “re-entering” variables whose removal changes the other estimates by more than 30% seems unusual to me. I understand the idea, but I’m not sure exactly how it’s supposed to work, and the only citation for it is a general article on confounding that doesn’t seem to mention logistic regression or this particular correction specifically. It’s not clear how potential collinearity of covariates was tested for.

RESPONSE 3: The 30% change in parameter estimation was used to detect possible confounding, as explained in the reference given [33]. Another reference is Bursac, Z., Gauss, C. H., Williams, D. K., & Hosmer, D. W. (2008). Purposeful selection of variables in logistic regression. Source Code for Biology and Medicine, 3, 17. http://doi.org/10.1186/1751-0473-3-17. With respect to collinearity, it was checked by studying, covariates correlation (rho>0.7) and checking parameter correlations. The association between those included in the model was lower than 0.3.

We have added new text at the end of the second paragraph of the Statistical section.

COMMENT 4: Model presentation: Table 2 presents some results of the model fitting, but the presentation is not very clear. Is this supposed to represent a final model? It includes variables “with p< 0.2”, but the “significance level” is stated to be 0.05. And Social Class is included, with p=0.68. Or are these just the p-values of the initial fit, before removing non-significant factors.

RESPONSE 4: We apologize for the misunderstanding. We have rewritten the footnote of that Table 2 in order to make this clearer as well as including a new phrase in the second paragraph of the Statistical analysis section.

COMMENT 5: Sampling design: I’m not an expert on sampling, but (or, perhaps, therefore) I would have liked to have seen an explanation of the claimed design effect. I would think it would depend on a number of assumptions/ estimates. I’m also not sure what the “p=q=0.5” refers to. The statement is made that “the sample design was considered throughout”, but it looks as though the confidence intervals in Supplementary Tables 1 and 2 have not been corrected for the design effect. For example, with 6207 individuals sampled and no design effect, the width of a 95% confidence interval for sex would be 2.48%, and the stated confidence interval has width 2.4%.

RESPONSE 5: Thank you. Each estimation of prevalence has its own design effect. In our study, there were good results as it was lower than 1.7 in any of the prevalence estimations. With respect to the parameters p and q, they take the quantity of 0.5 when considering the highest possible variability in dichotomous items (which are the most common ones used to estimate a prevalence). This means that the sample size will be high enough so as to estimate a prevalence with the lowest possible variability.
Finally, you are right about not considering the sample design in the supplementary tables 1 and 2. The reason was to present the sample sizes, their distribution on the main sociodemographic variables, and information about the missing data. Thus, the sample design is not applicable here. We have changed the titles of these tables and added a footnote in order to clarify that no sample design was considered.

COMMENT 6: In Figure 4, and elsewhere, the presentation of numbers was inconsistent with regard to significant figures. Given the size of the study, it seems to me that 2 places after the decimal point would be appropriate. The order of limits of the confidence intervals in Figure 4 are also inconsistent — sometimes the lower bound is first, sometimes the upper.)

RESPONSE 6: Thank you very much for your comment and suggestions for Figure 4. We have corrected the order of limits of the confidence intervals and rewritten the numbers adding 2 places after the decimal point instead of 3.

We hope our responses and the changes made address your comments.

VERSION 2 – REVIEW

REVIEWER	Domenico Plantone Institution and Country: ASL VCO - Ospedale San Biagio - Domodossola (VB), Italy
REVIEW RETURNED	05-Jun-2018

GENERAL COMMENTS	The authors addressed all the comments therefore I propose to accept the revised version of the manuscript.

REVIEWER	Martine J. Barons University of Warwick, UK
REVIEW RETURNED	05-Jun-2018

GENERAL COMMENTS	The authors have addressed my concerns satisfactorily and I have no further comments

REVIEWER	David Steinsaltz Department of Statistics University of Oxford United Kingdom
REVIEW RETURNED	03-Jun-2018

GENERAL COMMENTS	It's not necessary, but I would suggest restructuring Table 2. It took me a while to understand that the variable "disabling chronic pain" with outcomes Yes or No is a way of reporting the outcomes DCP and nDCP. Since these are the primary outcomes being predicted from the model -- and the decision has been made to model these as unordered categorical outcomes -- effectively there are just two separate models. I think it would make more sense to have two separate subtables, clearly indicating that these are coefficients for predicting these two distinct outcomes. Also, once it has been decided to choose one category as reference, it's confusing to then flip it, as was done in reporting "A significantly higher likelihood of DCP (adjusted for age and sex) was found in those belonging to manual labour social classes (ORmanual=1.26)" -- a reader will have difficulty relating this to the results reported on the table, as this is the reciprocal of the listed lower likelihood (.794) in those belonging to non-manual-labour classes. (Indeed, it would be helpful to state more prominently at the beginning of where selected results from the tables are being highlighted, which table they are being drawn from.)

VERSION 2 – AUTHOR RESPONSE

Comments to the Author (Reviewer 5):

Comment 1: I think it would make more sense to have two separate subtables, clearly indicating that these are coefficients for predicting these two distinct outcomes.

Comment 2: Also, once it has been decided to choose one category as reference, it's confusing to then flip it, as was done in reporting "A significantly higher likelihood of DCP (adjusted for age and sex) was found in those belonging to manual labour social classes (ORmanual=1.26)" -- a reader will have difficulty relating this to the results reported on the table, as this is the
reciprocal of the listed lower likelihood (.794) in those belonging to non-manual-labour classes.
(Indeed, it would be helpful to state more prominently at the beginning of where selected results from the tables are being highlighted, which table they are being drawn from.)

Responses to comments 1 and 2: We are pleased to submit a revised version of Table 2. Following the useful comment of the reviewer 5, we have produced two subtables as well as we have changed the reference category for the ‘Social class (short version)’ variable. We hope our changes made address his comment.