PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Antipsychotic Prescribing in Care Homes Before and After Launch of a National Dementia Strategy: An Observational Study in English Institutions over a 4-year Period
AUTHORS	Szczepura, Ala; Wild, Deidre; Khan, Amir; Owen, David; Palmer, Tom; Muhammad, Tariq; Clark, Michael; Bowman, Clive

VERSION 1 - REVIEW

REVIEWER	Victor Molinari University of South Florida School of Aging Studies Tampa, Florida United States of America
REVIEW RETURNED	14-Sep-2015

GENERAL COMMENTS	This study evaluated the impact of the National Dementia Strategy on the prescribing patterns of antipsychotic medications in care homes (both nursing & residential) in the UK four years after its launch. Unfortunately, there were no differences in the numbers of those receiving anti-psychotic medications nor general prescribing patterns. The authors are to be commended for doing this large scale analysis on such an important topic. Their discussion rightfully focuses on the wide variability found in prescription practices across care homes, and how a 'prescribing culture' may permeate some care homes. My main concern is that the discussion is a bit weak regarding other possible reasons for such a negative outcome and would recommend a more robust discussion that entails some of the following points: 1) There is a healthy literature regarding how didactics per se (e.g., lectures, continuing education, or as in this case, promulgation of standards) do not necessarily influence changes in behavior. In this study, it is unclear how the UK standards were "rolled out", who read them, and in particular whether these standards incorporated an enforcement mechanism,. The authors need to provide these details. In the United State, the Centers for Medicaid and Medicare Services (CMS) recently have been successful at reducing prescription of anti-psychotic medications in nursing homes. The CMS initiative was well publicized with the scorecards of different states and different nursing homes showcased and open to public scrutiny. Perhaps more important there has been the expectation that if certain nursing homes did not reduce medication usage, that quality assurance measures might begin to incorporate penalties for not adhering to the directives. The CMS wisely started "with the carrot", but the threat of "the stick" may be necessary for assurance of change and its maintenance. 2) There is little discussion of how staff education may be a key to

addressing mental health and behavioral problems of Care Home residents. It has been proposed that such staff education must incorporate not only a didactic component but a supervised monitoring of skills component in order to be effective in improving outcomes. Trained mental health consultants as in the Karlin & Teri STAR-VA program may be necessary to oversee behavioral programs implemented by nursing aides.

3) One of the main ingredients for facility change is "buy-in" at the highest level of Care Home administration (see studies by Burgio et al.). Again, the question becomes what are the costs and what are the benefits of adhering to the UK standards. This reviewer is not familiar with the UK system, but in the USA it appears that at times there is cost-shifting in that nursing homes may not bear the brunt of paying for the untoward long-term sequelae (e.g., falls, heart problems etc.) of what appears at times to be the short term cheapest way (for the nursing homes) of addressing behavior problems by prescribing medications.

4) Again, this reviewer is not familiar with the UK system, but I expect that most probably those GPs prescribing anti-psychotics did not have advanced psychiatric training. Promoting geriatric psychiatry specialists may be an important way of adhering to safe prescribing practices. Some nursing homes in the USA are using pharmacists as the centerpiece to 'red flag' residents who may be inappropriately on ant-psychotic medication so that they can be monitored and targeted for reductions in medications (or being taken off entirely).

5) Related to #4, we should recognize that at times anti-psychotic medications are appropriate. In the USA, perhaps 10% of residents in nursing homes have Serious Mental Illness, and anti-psychotic medications may be an effective strategy to reduce psychotic episodes within an overall plan of care. Indeed, in the USA, 'black box' warnings have been implemented to try to prevent prescribing of anti-psychotic medications for the behavior problems per se of residents, but in some cases treating hallucinations and delusions of residents with dementia with anti-psychotic medications is appropriate.

6) "Culture change" perhaps is the most important way to reduce anti-psychotic medications. Creating a person-centered, homelike environment whereby staff workers know the residents and create the conditions for meaningful activities and relationships may do most to prevent behavioral symptoms or quickly address them. The American Association of Geriatric Psychiatry and the American Society of Clinical Pharmacists have issued position statements that specifically say that non-psychopharmacological measures should be utilized before medication are tried. The judicious use of anti-psychotic medications entails targeting specific symptoms and monitoring their effectiveness on a regular basis.

REVIEWER	Sube Banerjee Brighton and Sussex Medical School I was involved in the development of the National Dementia Strategy. I have been involved in developing policy and in evaluations of the use of antipsychotic medication. I have been employed by the Department of Health.
REVIEW RETURNED	01-Oct-2015

| **GENERAL COMMENTS** | This is a study that investigates an important area where there is need for good quality data. The paper is well written and presented overall. It is a useful contribution to the literature and with appropriate attention to the comments made below would warrant publication.

The main issue with the paper that the authors should address is the generalisability of the sample of care homes. The number of homes, 211 at baseline and 616 at follow up, is impressive but is only a very small proportion of the number of care homes in England. If they were a random sample of care homes, it would indeed be possible to make the strong inferences that are made. However these are not likely to be a random sample, instead they are those that used the EMM system.
The authors need to provide a clear statement on the source of the homes and that came to use the EMM system. This is necessary to see if there is any selection bias or clustering. If for example the homes were all part of one or a few large providers then the study might then be looking at the response of a single corporation to the NDS rather than necessarily a reflection of what is happening nationally (despite the national reach of the sample) since they may share elements of the same culture. The broader the membership of the EMMs group, the more wide the generalisability of the data.

The inclusion of the Cohort C analysis of the baseline homes alone is positive methodologically and the similarity between the old homes and those newly joining supports the use of the aggregate data.

The discussions and conclusion elements of the paper are worded strongly and it would be useful for the authors to include some consideration of the limitations in inference that come from the dataset they have analysed. There is for example no consideration of selection bias, clustering effects or how and why the data for care homes presented seem to vary markedly from the results of all primary care presented in the first paragraph of the discussion. These would be useful additions to the paper and would strengthen its impact. |
| --- | --- |

| **REVIEWER** | Hans Wouters
University of Groningen - The Netherlands |
| --- | --- |
| **REVIEW RETURNED** | 05-Oct-2015 |

| **GENERAL COMMENTS** | This was a very interesting article to read. The authors have done commendable efforts to describe the rationale of the study in a succinct manner and to present the findings in a clear concise and straightforward manner. To me, the key messages were presented in a logical order. I am also convinced that the paper will capture most readers' attention until the very end. I have only some minor issues. I hope these will be useful for the authors:

1. Abstract -> Objectives: I do not think that the study design permits causal reasoning. Please rephrase "To assess the impact of a National Dementia Strategy (NDS)..." as "To assess associations between the launch of the National Dementia Strategy (NDS) and antipsychotic prescribing"

2. P3. Line 35 -> "Long term use of antipsychotic drugs is also |
| --- | --- |

associated with
many adverse effects including a significant deterioration in verbal fluency and cognitive function". Please clarify. Do you mean global cognitive function? As a matter of fact, verbal fluency belongs to the cognitive domain of language/ semantic memory and as such would be a kind of cognitive function. Cognitive function refers to global cognitive function with regard to memory, executive function, language, visuospatial processing etc. Please consult e.g. "Neuropsychological Assessment (by Lezak)".

3. P. 4 Line 3 -> "To date UK research on medication use in care homes remains limited". Please refer to Patterson et al. who conducted a randomized controlled trial in Northern-Ireland.

4. P.4 Line 29 -> Could there have been a short-term effect of the NDS say after one year in 2010? Have you got data about the year 2010? If so, present findings from 2010.

5. P.4 Line 49 -> "...due to the time consuming nature of this process, such
analysis was limited to the licensed agent (risperidone)". Why was this the case? Wasn't it possible to do this in an automated manner?

6. P.4. -> Is a Kolmogorov-Smirnov test adequate? I am not sure whether it can handle dependent data? Since I suspect the data from the 2 time points are - at least in part - collected from the same homes, the authors should verify whether a Cochran's Q or Kendall's tests is more appropriate. Alternatively the authors could also opt for generalized linear mixed models.

7. Table 1 -> There is a huge difference in number of homes between baseline and follow-up. In the beginning of the Method section (p.4 Line 30), the authors explain that "Because uptake of the EMM system increased over this period, a data sub-set was extracted for a cohort of care homes with the EMM system in place throughout the 4-year period (Cohort C)". Do you mean that the uptake of HOMES IN THE EMM SYSTEM INCREASED? Otherwise, I don't understand this big difference. I recommend that the authors explain this a bit more.

8. Page 14 Line 40 -> "Evidence on a direct relationship between staffing levels and antipsychotic use is currently lacking". The findings by Zuidema et al. International Psychogeriatrics, 2011 would be interesting in this regard.

9. First general comment on conclusion -> Perhaps the NDS decelerated the increase of antipsychotics?

10. Second general comment on conclusion -> Overall the NDS seemed to be unsuccessful. However, given the huge variation between areas/homes, might it have been possible that the NDS was actually successful for some homes and unsuccessful for other homes? It would be worthwhile to do some more detailed analyses on this matter.

- END: THANK YOU -

REVIEWER	Peter Watson Medical Research Council UK
REVIEW RETURNED	23-Oct-2015

GENERAL COMMENTS	Antipsychotic prescribing in care homes before and after launch of a national dementia strategy: an observational study in English institutions over a 4-year period. Bmjopen-2015-009882
	A few queries here: firstly there appears to be no omnibus direct testing of differences between the rate ratios displayed in Table 3 on page 10 instead, seemingly, relying on the (non)overlap of confidence intervals about the rate ratios. Omnibus tests and also corrections for multiple testing are used to guard against the dangers of obtaining spurious false positive results in multiple testing particularly when doing pairwise comparisons of groups. The overlap of confidence intervals can also be misleading when judging if there are differences between groups.
	I also think multilevel modelling would be a better analysis than the one presented in this paper since it directly utilises the inherent nesting of residents within rest homes and is used frequently in the medical and other literature for this type of data and could simply use the number of prescriptions per resident or a prescribed/not prescribed dichotomy as the outcome response to see if these vary with predictors without the need for the use of arbitrary divisions into quintiles.
	It is also not explicitly stated how the conclusions in the abstract on page 2 and in the results on page 7 are arrived at. Assuming results use the non-overlapping confidence approach to assess whether the Q5/Q1 rate varies across levels of a factor there still appear to be omissions and incorrect conclusions from the results in Table 3 on page 10.
	Page 5, line 38. I would also query, when assessing variations in prescription rates of individual residents as done in Table 3 on page 10, the assumption used in placing all residents into the same quintile based solely on a pooled estimate of prescriptions rates as evinced by a total prescription rate in their care home. Surely not all residents would be high prescribers of antipsychotics just because they are in a rest home which has an overall high number of prescriptions compared to other care homes?
	In more detail:
	In Table 3 you are comparing proportions of resident characteristics in those residents with low and high numbers of prescribed antipsychotic prescriptions. A chi-square-test is usually used for directly comparing odds across different levels of a factor (or Fisher's exact if expected cell sizes are small). The usual chi-square tests of frequencies could compare the proportion of people who feature in the two most extreme prescribing quintiles who are in in the highest prescribing quintile across levels and then, if this is statistically significant, an odds ratio test and confidence interval could be used to compare and illustrate the odds of being in the high prescribing interval to the low prescribing quintile in one level (e.g. age 65-74 if age is the factor) as opposed to another level (e.g. 75-84 if age is the factor).

Table 3 on page 10 could also be made clearer by clarifying what the frequencies for Q1 and Q5 refer to. I assume, for example, that the frequencies under resident characteristics are numbers of residents whereas the frequencies under care home characteristics are those of care homes? The title of the table implies to me that both are used.

I also wonder in comparisons of rate ratios which are based on residents rather than care homes if the differences obtained in Table 3 on page 10 are been driven by just a handful of care homes who have large percentages of residents on antipsychotics. For example in Figure 1 on page 23 the cumulative distribution plots suggest most care homes have less than 40% of their residents on antipsychotics but that there are a few who have large proportions of residents on antipsychotics so that the numbers of residents in Table 3 who have high levels of prescribed antipsychotics may be coming from a relatively small number of care homes.

I feel a multilevel analysis would be better than the rate ratio analysis in Table 3 because the multilevel analysis would additionally take advantage of the nested nature of the data as evinced by several residents living in the same care home (ie nested). This analysis teases out both differences among residents and differences across care homes in prescription rates in the form of variance components. You would not need to use quintiles but just have number of prescriptions per resident or, alternatively, whether a prescription for an antipsychotic has been given for each resident at baseline as an outcome, resident number in care home and care home number which could yield the variance components. Resident characteristics could then be entered as predictors and you could then obtain chi-squares or even possibly F ratios which investigate such things as whether different types of residential homes (e.g. residential vs nursing vs dual registered as in Table 3 on page 10) have different prescription rates. Such an approach is taken in, for example, Barber ND et al. (2009) Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 18 341-346.

It is not mentioned how you reached the conclusions from the results in Table 3 on page 10. For example in the abstract (page 2, line 22) it states that care homes in the highest prescribing quintile are more likely to be located in a deprived area and a confidence interval given for the rate ratio. I suspect the conclusions are based upon whether a set of confidence intervals overlap (e.g. the confidence interval for deprived neighbourhood in Table 3 on page 10 does not overlap with that for non-deprived neighbourhoods). If this is the case I would query the absence of any omnibus test directly comparing the ratios in deprived areas with those in non-deprived areas. Similarly there is no direct comparison of the three home types or three GP practice types or any other levels of categories in Table 3 on page 10 where conclusions about differences are made on page 9 (lines 11-21). Wolfe R and Hanley J (2002) mention that although non-overlapping confidence intervals are indicative of a difference, differences between pairs of groups can still exist even if their confidence intervals overlap which is another reason why we usually do a more formal test (e.g. chi-square) to assess differences in estimates.

Using the non-overlapping confidence interval approach the authors surprisingly don't mention on page 9 that dual registered home types have high numbers of residents with low numbers of prescriptions for antipsychotics compared to residential and nursing homes. Instead they say (page 9, line 13) that there is no clear difference in rate ratios across types of institution. Similarly in the abstract on page 2, line 24 at the end of the results paragraph it mentions that those with lower rates of prescriptions for antipsychotics are more likely to be served by a single GP practice yet from Table 3 on page 10 line 30 there is an overlap between the confidence intervals for prescription rates involving 1 GP practice (line 29) with that for 2-3 GP practices (line 30) so assuming the authors are using non-overlap of confidence intervals as their basis for assessing differences I wondered how they had reached this conclusion. Usually to avoid obtaining two many false positive responses we only perform pairwise comparisons if the overall test of group differences is statistically significant. In the case of comparing the GP practices in Table 3 on page 10 one could argue one should firstly perform an omnibus test to see if all three groups differ on rate ratios (akin to a F test in an anova) and only if this overall test is statistically significant compare the three pairs of practices with an appropriate correction factor to the p-values such as Bonferroni or Sidak for performing multiple comparisons.

Other comments

Is there any evidence of different approaches being used inside care homes to decide whether residents are prescribed antipsychotics?

On line 10 of page 14 it mentions large variations in 'antipsychotic use' amongst care homes. Is this saying there is variation in the rate of prescribing antipsychotics among care homes and where is this stated in the results? Contrary to this, I notice on line 8 of page 7 that nursing and residential homes are stated as exhibiting similar PP [prescription] rates at both baseline and 48 months which seems to imply this rate of variation is not, at least, between these two types of care homes. I do notice in Table 3 on page 10 that dual registered homes have a smaller rate ratio than both residential and nursing homes but since this is not reported in the results section I assume the authors didn't use this to support this statement about differences in antipsychotic prescriptions rates amongst care homes.

Did you consider looking at the change in resident characteristics across quintiles instead of comparing the two extremes?

Page 12, I don't find this table particularly useful because I don't know what or where the areas are from looking at the table. Could you use more descriptive regional area terms such as Scotland, Cumbria to differentiate between them?

It would be useful to state the equation for the delta method used for the variance of the rate ratios in Table 3 (page 10) perhaps in a small appendix. One could quote the general formula and, perhaps, show how it is applied to an example in this paper.

Reference
Wolfe R and Hanley J (2002) If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2. CMAJ 166 65-66

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1:

OVERALL:…" commended for doing this large scale analysis on such an important topic."
We thank reviewer 1 for this commendation. We especially welcome comments from the US perspective.

…..(R1) would recommend a "discussion that entails some of the following points":
We have tried to incorporate some of these points, in a succinct manner, adding refs:

Q1. "whether these standards incorporated an enforcement mechanism. Also, Medicaid and Medicare (USA) reduced antipsychotics in nursing homes with scorecards of different settings showcased & open to public scrutiny"
A1. This is now made clear for international readers i.e. para 4 p3 "were incorporated as guidelines, without any enforcement systems"; and p15 that "the National Dementia Strategy did not include any enforcement mechanisms and antipsychotic prescribing patterns in UK care homes are not open to public scrutiny"

Q2. "staff education may be key to addressing mental health and behavioral problems of residents"
A2. We agree & have highlighted the importance of this para 2 p14 "organisational factors such as leadership and investment in staff development" and added a reference.

Q3. "main ingredients for facility change is "buy-in" at the highest level of Care Home administration"
A3. We agree about the importance of leadership/ senior management role and have added text para 2 p14 (see A2 above).

Q4. "Also, in the USA it appears that at times there is cost-shifting from nursing homes may not bear cost of long-term sequelae (e.g. falls) so choose cheapest way (for the nursing homes) of addressing behavior problems."
A4. In our experience cost-shifting of this type is less likely to be observed in the UK.

Q5. "most probably those GPs prescribing anti-psychotics did not have advanced psychiatric training Promoting geriatric psychiatry specialists may be important for safe prescribing practices"
A5. It is true that UK GPs would generally not have this type of advanced training. However, community pharmacists are also important. We have added regular review "by GPs or community pharmacists" (para 1 p14).

Q6. "Culture change" perhaps is the most important way to reduce anti-psychotic medications. American Association of Geriatric Psychiatry & American Society of Clinical Pharmacists have issued position statements saying non-psychopharmacological measures to be utilized before medication tried
A6. We wholeheartedly agree – but culture change is not only difficult to achieve (linked to factors such as A2 & A3 above), it is also difficult to sustain over a longer period of time.
We were very interested to hear about the position statements issued by the AAGP & ASCP in the US.

REVIEWER 2:

OVERALL: "The paper is well written and presented overall"." (R2)
We thank reviewer 2 for this comment. Our aim was to present a complex analysis in a succinct & clear manner.

Q1. address the generalisability of the sample of care homes. The number of homes, 211 at baseline and 616 at follow up, is impressive but is only a very small proportion of the number of care homes in England.

A1. Of course, although our sample is 'impressive' it represents only a small proportion of the total UK care homes. In terms of demonstrating generalisability, we are unable to report comparable national data (e.g. for Table 2). This is not available from the relevant national body i.e. the Care Quality Commission Ref [32]

Q2. If for example the homes were all part of one or a few large providers then the study might then be looking at the response of a single corporation to the NDS….The broader the membership of the EMMs group, the more wide the generalisability of the data.

A2. We can confirm that homes in the sample were not "part of one or a few large providers" because we recorded whether homes were part of a chain.

Exploration of chain size and other care home characteristics such as type of ownership (commercial, voluntary/charity, local authority) will be the focus of a further article.

Q3. The inclusion of the Cohort C analysis of the baseline homes alone is positive methodologically and the similarity between the old homes and those newly joining supports the use of the aggregate data.

A3. It is reassuring to receive this comment.

Q4. The discussions and conclusion elements of the paper are worded strongly and it would be useful for the authors to include some consideration of the limitations in inference that come from the dataset they have analysed

A4. Bearing in mind the challenging targets set by the NDA, we do not consider that our findings are worded too strongly. However, we have added to Strengths & Limitations box (p2) "One limitation of this observational study is the lack of comparable national data to demonstrate generalisability of the study sample"

Q5. include some consideration of …..how and why the data for care homes presented seem to vary markedly from the results of all primary care presented in the first paragraph of the discussion.

A5. The two sets of data are very different.

For primary care, the cited article analysed antipsychotic prescriptions at the point when dementia was first diagnosed. Our findings are not for newly diagnosed cases. We would therefore expect antipsychotic use to be higher. Also, since the NDS is reported to have led to increased diagnosis in the community, cases are presumably identified at an even earlier stage. We do say that "A trend towards earlier diagnosis may explain the lower rate of antipsychotic prescribing reported in 2011 in the first study" (para 1 p13).

We would not like to speculate further.

REVIEWER 3:

OVERALL: "This was a very interesting article to read. The authors have done commendable efforts to describe the rationale of the study in a succinct manner and to present the findings in a clear concise and straightforward manner." (R3)

We are glad that reviewer 3 found the article "very interesting" and that he thought we presented our findings in a clear manner for the reader.

Q1. Abstract: "Objectives: Please rephrase "To assess the impact of a National Dementia Strategy (NDS) on antipsychotic prescribing..." as "To assess associations between the launch of the National Dementia Strategy (NDS) and antipsychotic prescribing"

A1. Because the other three reviewers did not suggest rephrasing the Abstract, we have decided to

leave this text unchanged.

Q2. P3. Line 35: -"Long term use of antipsychotic drugs is also associated with many adverse effects including a significant deterioration in verbal fluency and cognitive function". Please clarify. Do you mean global cognitive function?

A2. The type of cognitive function is unclear. We have therefore replaced with a more significant conclusion by the authors of 'increasing concerns about serious adverse effects including mortality' (para 3 p3)

Q3. P. 4 Line 3: "To date UK research on medication use in care homes remains limited". Please refer to Patterson et al. who conducted a randomized controlled trial in Northern-Ireland.

A3. We thank the reviewer for this information. The Patterson at al RCT evaluated the impact in nursing homes in Northern Ireland of a United States model of pharmaceutical care on prescribing of inappropriate psychoactive medications and falls. We have added the Patterson et al. reference (final para p3). However, we would still argue that UK research on medication use in care homes is limited.

Q4. Could there have been a short-term effect of the NDS say after one year in 2010? Have you got data about the year 2010? If so, present findings from 2010.

A4. Regrettably, we do not have data for 2010. However, we consider it very unlikely that there would be a significant short-term effect in 2010 which would then disappear, especially since the NDS recommended a reduction in prescribing levels 'over a period of 3 years'.

Q5. P.4 Line 49: "...due to the time consuming nature of this process, such analysis was limited to the licensed agent (risperidone)". Why was this the case? Wasn't it possible to do this in an automated manner?

A5. It was not possible to estimate LOE in an automated manner. Data on sequential prescriptions were not necessarily recorded in a consistent manner & had to be checked, a time consuming process.

Q6. P.4: Is a Kolmogorov-Smirnov test adequate? I am not sure whether it can handle dependent data? Since I suspect the data from the 2 time points are - at least in part - collected from the same homes, the authors should verify whether a Cochran's Q or Kendall's tests is more appropriate. Alternatively the authors could also opt for generalized linear mixed models

A6. The two-sample Kolmogorov-Smirnov test is appropriate here because it investigates whether the two data samples come from the same distribution without making any assumption about what that distribution is.

Q7. (R3) In the beginning of the Method section (p.4 Line 30), the authors explain that "Because uptake of the EMM system increased over this period, a data sub-set was extracted for a cohort of care homes with the EMM system in place throughout the 4-year period (Cohort C)". Do you mean that the uptake of HOMES IN THE EMM SYSTEM INCREASED? Otherwise, I don't understand this big difference. I recommend that the authors explain this a bit more.

A7. Yes – the number of homes using the EMM system did increase. We have made this slightly clearer (just in case other readers are confused) by changing the text to "Because the number of care homes which had implemented the EMM system increased over this period" (Methods, para 1).

Q8. Page 14 Line 40: "Evidence on a direct relationship between staffing levels and antipsychotic use is currently lacking". The findings by Zuidema et al. International Psychogeriatrics, 2011 would be interesting in this regard.

A8. Thanks, we have added Zuidema reference (sentence 3, p14). In this Dutch nursing homes study, it is reported that staff distress plus other aspects of the nursing home environment are associated with psychotropic drug use. We therefore consider this reference sits better with other evidence on

staff distress.

Q9. First general comment on conclusion: Perhaps the NDS decelerated the increase of antipsychotics

A9. Since there was no control group not exposed to the National Dementia Strategy, there is no way of confirming deceleration of otherwise increase in prescribing rates.

However, although this reviewer's comments focus primarily on prescribing rates, it is also important to note that our analyses also show a lack of change in response to NDS recommendations about type of antipsychotic and length of exposure.

Q10. Second general comment on conclusion: might it have been possible that the NDS was actually successful for some homes and unsuccessful for other homes? It would be worthwhile to do some more detailed analyses on this matter.

A10. The question of individual 'success' is complex. In terms of prescribing rates, some homes started with very low PP values at baseline and therefore limited room for improvement (see Figure 1), whereas others have very high rates which might be reduced. Some analysis will be included in a further article.

Finally, 'success' is not confined solely to prescribing rates (see A9 above).

REVIEWER 4:

Q1. there appears to be no omnibus direct testing of differences between the rate ratios displayed in Table 3 on page 10...... I also think multilevel modelling would be a better analysis than the one presented in this paper.

could simply use the number of prescriptions per resident or a prescribed/not prescribed dichotomy

A1. The reviewer is correct that there is no omnibus testing of the rate ratios. However the rate ratios can be investigated by comparing whether the 95% confidence intervals overlap. The suggestion to use multilevel modelling is interesting but we prefer a basic reporting of rates in this first instance.

Also, this paper does not only examine whether the challenging target of a 'two thirds reduction over a period of 3 years' as recommended by Prof Banerjee (R2) was achieved. Equally important are analysis of the types of antipsychotics prescribed, and the length of exposure (versus the NDS recommendations).

Q2. Page 5, line 38. Surely not all residents would be high prescribers of antipsychotics just because they are in a rest home which has an overall high number of prescriptions compared to other care homes?

A2. It is doctors not residents who are 'high prescribers'. Also, we are not reporting 'overall number of prescriptions compared to other care homes'. Instead, we report point-prevalence (PP) i.e. the percentage of residents in a care home prescribed at least one antipsychotic at a specific time point.

Q3. In more detail: A chi-square-test is usually used for directly comparing odds across different levels of a factor (or Fisher's exact if expected cell sizes are small).

A3. Chi-squared tests measure the association between the two categorical variables. The odds ratio is one possible measure of that association. Comparisons between the rate ratios we report is possible because we provide confidence intervals for the rate ratios. There is no need to perform the chi-squared test and report an odds ratio if that is significant since this would be selective reporting.

Q4. Table 3 on page 10 could also be made clearer by clarifying what the frequencies for Q1 and Q5 refer to.

A4. On P5 (under Characteristics of high/low prescribing institutions) we explain that "care homes were placed into quintiles based on their baseline PP level; if an organisation was placed in a particular quintile, all residents in that home were placed in the same quintile". Frequencies are

number of residents. The other three reviewers did not identify the need for greater clarity.

Q5. In more detail: I also wonder….. the numbers of residents in Table 3 who have high levels of prescribed antipsychotics may be coming from a relatively small number of care homes.
A5. Table 3 does not show the 'numbers of residents who have high levels of prescribed antipsychotics'. In fact, almost no resident receives high levels (doses) of antipsychotics (0.3% residents overall, see Table 2). If this comment refers to the number of antipsychotics per resident, then residents are also very rarely "(0.7% at baseline and 1.67% at 48 months) prescribed more than one antipsychotic at the same time." (see para 2 p7).

Q6. In more detail: I feel a multilevel analysis would be better than the rate ratio analysis in Table 3
A6. See response A1 above.

Q7. In more detail: there is no direct comparison of the three home types or three GP practice types or any other levels of categories in Table 3 on page 10 ….. although non-overlapping confidence intervals are indicative of a difference, differences between pairs of groups can still exist even if their confidence intervals overlap
A7. We are aware that differences between groups may still exist, even if their confidence intervals overlap. However, this paper is not aiming to present an in-depth analysis of prescribing rates alone (see response in A1 above). In terms of adherence to NDS guidance other indicators of quality & adherence - such as the type of antipsychotic prescribed, dosage, length of treatment, and which antipsychotic is prescribed first are equally important.

Q8. In more detail: surprisingly don't mention on page 9 that dual registered home types have high numbers of residents with low numbers of prescriptions for antipsychotics compared to residential and nursing homes.
A8. This comment is unclear i.e. 'numbers of residents with low numbers of prescriptions for antipsychotics'. Virtually all residents received only one prescribed antipsychotic (see response A5 above).

Q9. Other comments: On line 10 of page 14 it mentions large variations in 'antipsychotic use' amongst care homes. Is this saying there is variation in the rate of prescribing antipsychotics among care homes and where is this stated in the results?
A9. The large variation in levels of 'antipsychotic use' among care homes should be clear from Figure 1 – see range on axis 'Percentage of residents on antipsychotics'.

Q10. Other comments: I do notice in Table 3 on page 10 that dual registered homes have a smaller rate ratio than both residential and nursing homes but since this is not reported in the results section I assume the authors didn't use this to support this statement about differences in antipsychotic prescriptions rates amongst care homes
A10. We have not examined this group in detail because dual registered homes are difficult to categorise. They vary across a wide spectrum, from residential homes with a small number of nursing beds to nursing homes with a few residential places. Because we did not have a breakdown of bed types, drawing any conclusions about these homes is problematic.

Q11. Other comments: Page 12, I don't find this table particularly useful because I don't know what or where the areas are from looking at the table. Could you use more descriptive regional area terms such as Scotland.
A11. Locations are not identified to preserve anonymity. However, the areas listed represent a level at which prescribing patterns might eventually be monitored. Collapsing these into larger "more descriptive regional terms … such as Scotland" would not be particularly helpful.

Q12. Other comments: It would be useful to state the equation for the delta method used for the variance of the rate ratios in Table 3 (page 10) perhaps in a small appendix. One could quote the general formula and, perhaps, show how it is applied to an example in this paper.

A12. This is the standard delta-method formula for the standard error of a rate ratio (see for example Kirkwood and Sterne, Essential Medical Statistics 2nd edition 2003), i.e. if the two by two table entries are: a, b, c, d; then the standard error of the log rate ratio is given by the square root of:

$1/a - 1/(a+c) + 1/b - 1/(b+d)$.

VERSION 2 – REVIEW

REVIEWER	Victor Molinari University of South Florida School of Aging Studies Tampa, Florida, USA
REVIEW RETURNED	05-Feb-2016

GENERAL COMMENTS	The authors appear to have addressed all my major concerns. However, I believe it should be clarified that that although the National Dementia Strategy encouraged the more judicious use of anti-psychotic medications for those with dementia, many residents of care homes do not have dementia, and prescribing anti-psychotic medications for those residents with primary diagnoses of psychosis e.g., schizophrenia or bipolar disorder (~ 10% in the U.S.) may well be evidence based practice. One minor point - please review p.13, lines 28-32: it is stated that there is more prescribing of anti-psychotics in other countries versus the USA, but the numbers for Canada appear to contradict this. Please clarify.

REVIEWER	Sube Banerjee Brighton and Sussex Medical School, UK I have been but am no longer employed by the Department of Heath for England. I was the co-lead of the National Dementia Strategy and of the "Time for Action" report that is the subject of this paper.
REVIEW RETURNED	10-Mar-2016

GENERAL COMMENTS	The paper remains a useful contribution. There are four points that I believe continue to be worth addressing. 1. The response to a point made before (R2 Q2/A2 in the response) is less than satisfactory. Given the small sample size relative to the total, the issue of generalisability is not a minor one. I think that the data are difficult to interpret without a presentation and analysis of the provider type. These data are available, the authors state "Exploration of chain size and other care home characteristics such as type of ownership (commercial, voluntary/charity, local authority) will be the focus of a further article." I believe that this paper would make a much stronger case for publication if these data were included rather than promised in a future possible publication.

2. Flowing from this, the paper needs discuss the observation that the prescription rates in the homes are relatively low at inception at 18-19%. The report referred that I authored, while this was not a point prevalence, worked on the basis of a 33% base rate for antipsychotic prescribing in dementia with the assumption that prescribing in care homes was higher than this this and in the community lower. The authors should consider that these (selected) homes might have been high functioning with respect to antipsychotic use at baseline and so there was little change because their practice was already optimised. The use of these drugs had been an issue that had been raised for many years before as the paper states. If this had been a random sample of all care homes this would not have been a problem. Since it is a group selected by the use of a specific system then it may well be.

3. The response to R2 Q4/A4 is similarly still problematic. The text continues to read in a somewhat polemical fashion with, in my opinion, insufficient attention paid to the limitations in inference inherent in the data obtained. It is something of a 'straw man' argument to present the data obtained in these care homes in effect as a failure of policy.

4. Finally I think that it is problematic that the National Dementia Strategy is so prominently featured in the title and in the text and that this work is presented as an evaluation of the NDS. The NDS was published in February 2009. The goals described and suggested in my report were not in fact part of the NDS but instead a separate enquiry and report to DH published in October 2009. It was also incidentally not specific to care homes and two thirds of people with dementia live in their own households rather than in care homes. It is therefore potentially a little misleading to represent the data as an evaluation of the NDS and a demonstration that it "failed" to effect a specific change in antipsychotic prescribing.

REVIEWER	Hans Wouters
	University of Groningen, Netherlands
REVIEW RETURNED	23-Feb-2016

GENERAL COMMENTS	Thank you for reading this important and very interesting paper. I appreciate the authors' efforts to further improve the manuscript.
	Personally, I still think you cannot talk about the impact of the NDS since you did not conduct an RCT (which was of course impossible to conduct and of course we cannot blame the authors for not conducting an RCT).
	While I liked the authors' strong arguments to my other points, I do not think it is a strong argument to say that you did not rephrase the Abstract because the other reviewers had no comments, because you could have given that answer to all of my other points as well, whereas you didn't.

REVIEWER	Peter Watson
	Medical Research Council
	UK
REVIEW RETURNED	09-Feb-2016

GENERAL COMMENTS	Antipsychotic prescribing in care homes before and after launch of a national dementia strategy: an observational study in English institutions over a 4-year period. Bmjopen-2015-009882.R1
	Having looked at the revised paper there appears to be no change in the analysis used in Table 3 on page 10 to examine care home characteristics in the baseline sample. It still seems to me a more natural way of analysing this data which is hierarchical is to use a model which allows for this hierarchy namely of patients nested within care homes. In the previous review I mentioned that, for example, Barber ND et al. (2009) Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 18 341-346 used a multilevel analysis examining patients. (This can also be used to examine the next level up namely comparing residential home characteristics). Generalised estimating equations could also be used.
	In particular Barber et al. mention in their statistical analysis section: "Generalised estimating equations (library geepack, V.1.0-10) were used to model patient level odds of errors, allowing for clustering in homes and using an independence correlation structure. Multilevel models were also used to model patient level odds of errors, using the MLwiN 2.03 software (multilevel models project, University of Bristol, http://www.cmm.bristol.ac.uk), fitting variance components at the various levels".
	Multilevel software is available in most statistical packages including Stata which was the software used in this paper with, for example, the 'mixed' procedure fitting linear mixed models and the 'meglm' procedure for other distributions such as the poisson. The idea is that one is accounting for correlations between residents in the same care home who might be expected to be more related being in the same care home than residents who are in different care homes. It's a bit like in a repeated measures ANOVA where we have two sources of variation (within subject and between subject) where responses on the same subject are regarded as being more closely related than responses between different subjects.
	As I also suggested previously an omnibus test comparing all levels of a factor is usually performed and further pairwise post-hoc tests as presented here (with the addition of a Bonferroni correction or similar for multiple testing) performed if the omnibus test is statistically significant. The results in the second paragraph on page 9 are based upon non-overlapping confidence intervals however the authors omit the result that dual registered (Table 3) residents have a Q5/Q1 confidence interval that does not overlap with residential and nursing home types suggesting, using their approach, that people are less likely to be in the highest quintile if they are living in dual registered homes. One could also add to the results on page 9 that the 75-84 year olds have a 95% CI in Table 3 (page 10) which does not overlap with either ages 65-74 or 85+ which using the non-overlapping CIs would suggest they are more likely to be in the upper quintile than those aged 85+ but less likely than those aged 65-74.
	There is a greater 'degree' of non-overlap of most of the confidence intervals which are reported on page 9 but it does not follow that you can compare differences in group outcomes by whether or not the confidence intervals overlap which is a reason why we do F tests, for example, in ANOVAs rather than diving straight in to perform post-

	hoc pairwise contrasts. I think the authors could at least mention in the paper why they have not used either omnibus tests to compare factor levels in Table 3 on page 10 or hierarchical multilevel modelling. Page 10. An addition to the footnote of Table 3 could state what the figures in brackets under the columns headed Q1 and Q5 represent. I thought these might be percentages of each factor in each level in each quintile but these don't add up to 100% e.g. for age in Q5 in Table 3 (page 10) the sum of the percentages is 12.6+39.4+37.5 = 89.5 < 100.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1: Professor Victor Molinari, University of South Florida, School of Aging Studies, Tampa, USA

Professor Molinari provides two final comments, both of which we have now addressed:

(Comment 1.1) "I believe it should be clarified that that although the National Dementia Strategy encouraged the more judicious use of anti-psychotic medications for those with dementia, many residents of care homes do not have dementia, and prescribing anti-psychotic medications for those residents with primary diagnoses of psychosis e.g., schizophrenia or bipolar disorder (~ 10% in the U.S.) may well be evidence based practice".

A: Although use of antipsychotics for patients with schizophrenia etc is evidence-based, people with such a primary diagnosis would only very rarely be placed in a UK non-specialist care home. The 2012 census of care homes (reference 5) identified 42.4% of admissions to care homes as being for dementia and only 4.3% for mental illness (other than dementia); among this 4.3%, the most commonly occurring illness was depression.

(Comment 1.2) "One minor point - please review p.13, lines 28-32: it is stated that there is more prescribing of anti-psychotics in other countries versus the USA, but the numbers for Canada appear to contradict this. Please clarify".

A: We thank Professor Molinari for pointing this out, although it is actually the Australian figure (25.1%), not the Canadian (30.3% & 36.9%), which is low & similar to the lowest USA figure (25%).

The non-US figures are generally less reliable because they are based on very small-scale studies compared to our study and those reported figures from the USA (e.g. ref 3 Briesacher analysed 2.5 million Medicare beneficiaries). We have therefore edited the manuscript to make this distinction much clearer (paragraph 2, p13):

"In contrast to these US figures, which are based on 2-3 million Medicare beneficiaries, smaller scale studies in other parts of the world generally report higher levels of antipsychotic use"

We have also re-arranged the non-US rates quoted in order of increased sample size (i.e. number of care homes) to emphasise this.

Reviewer 2: Professor Sube Banerjee, Director of Centre for Dementia Studies, Brighton and Sussex Medical School

We thank Professor Banerjee for his conclusion that "The paper remains a useful contribution" and we are happy to respond to his "four points that I believe continue to be worth addressing":

(Comment 2.1) The response to a point made before (R2 Q2/A2 in the response) is less than satisfactory. Given the small sample size relative to the total, the issue of generalisability is not a minor one. I think that the data are difficult to interpret without a presentation and analysis of the provider type. These data are available, the authors state "Exploration of chain size and other care home characteristics such as type of ownership (commercial, voluntary/charity, local authority) will be the focus of a further article." I believe that this paper would make a much stronger case for publication if these data were included rather than promised in a future possible publication.

A: In terms of generalisability, we would ideally have liked to be able to present data to address this question. Unfortunately, we were unable to identify any statistics for all care homes in England to compare with the characteristics presented in Table 1 i.e. mean/median [Inter-quartile range] size of homes, number of GP practices serving a home, geographical location of homes (i.e. urban/rural & deprivation). However, in order to make this limitation clearer, we have inserted a final paragraph in the Discussion highlighting this, and other, limitations (p15).

Finally, we apologise if our mention of a further article has raised expectations. Further analyses are still at an early stage and will have an entirely different focus, namely exploring the use of operations research methods to examine productive efficiency of care homes and links to organisational & funding type.

(Comment 2.2) Flowing from this, the paper needs discuss the observation that the prescription rates in the homes are relatively low at inception at 18-19%. The report referred that I authored, while this was not a point prevalence, worked on the basis of a 33% base rate for antipsychotic prescribing in dementia with the assumption that prescribing in care homes was higher than this this and in the community lower. The authors should consider that these (selected) homes might have been high functioning with respect to antipsychotic use at baseline and so there was little change because their practice was already optimised. The use of these drugs had been an issue that had been raised for many years before as the paper states. If this had been a random sample of all care homes this would not have been a problem. Since it is a group selected by the use of a specific system then it may well be.

A: Early in the Discussion we do make it clear that the rates we report are relatively low compared to international data (final sentence, 2nd paragraph p13). It is much more difficult to place our figures in a UK context where published studies are generally much smaller. However, we do reference three studies which have reported similar rates in a cross-section of nursing and residential homes during the same period, although with far fewer sites (second sentence, 2nd paragraph p13). We also reference a much higher rate i.e. 27% Ballard (2008) and a much lower rate of 12% (Backhouse 2013). The former was recorded in a small number of specialist nursing homes for people with dementia, and the latter was based on unconfirmed, self-reported figures in a survey (60% non-response rate). The issue of rates is further confused by some studies reporting prescribing rates for residents with dementia, as opposed to all residents. For example, a rate of 33% in residents with dementia, combined with ~45.6% dementia prevalence in UK care homes would equate to a prescribing rate of 15.0% for the whole care home population.

Professor Banerjee raises another important question when he asks whether the care homes in our sample might be 'high functioning' at baseline and 'their practice already optimised'. If this were the case, it would mean there is less incremental improvement possible in prescribing rates. However, the range of prescribing rates we observed (see median [Inter-quartile range] in Table 2) would seem to

refute this hypothesis. Furthermore, as well as no significant downward trend in prescribing rates, other measures of sub-optimal prescribing also showed no evidence of improvement (see first sentence, 3rd paragraph p13) i.e. no shift towards SGAs, off-label prescribing remaining high, and length of treatment continuing to exceed the recommended 6-12 weeks.

(Comment 2.3). The response to R2 Q4/A4 is similarly still problematic. The text continues to read in a somewhat polemical fashion with, in my opinion, insufficient attention paid to the limitations in inference inherent in the data obtained. It is something of a 'straw man' argument to present the data obtained in these care homes in effect as a failure of policy.

A: We appreciate that the language in some places might be construed as 'polemic', especially without full acknowledgement of limitations in the data. We do now acknowledge more clearly the limitations in the data, and we have added a final paragraph to the Discussion to this effect (see response to Comment 2.1 above). We have also introduced some minor changes to the language in 'Policy and research implications' section (p15). Finally, we have removed any use of the word 'failed' and replaced this as follows:

'failed' replaced by 'was not associated with' (Abstract Conclusions)
'failed to effect' replaced by 'reductions in the prescribing of antipsychotic agents driven by the National Dementia Strategy have not been sustained' (First sentence Discussion p13)
'has failed to achieve change' replaced by 'was not associated with sustained change' (Policy and research implications p15)

(Comment 2.4). Finally I think that it is problematic that the National Dementia Strategy is so prominently featured in the title and in the text and that this work is presented as an evaluation of the NDS. The NDS was published in February 2009. The goals described and suggested in my report were not in fact part of the NDS but instead a separate enquiry and report to DH published in October 2009. It was also incidentally not specific to care homes and two thirds of people with dementia live in their own households rather than in care homes. It is therefore potentially a little misleading to represent the data as an evaluation of the NDS and a demonstration that it "failed" to effect a specific change in antipsychotic prescribing.

A: We did not wish to imply that this is an 'evaluation' of NDS. We have checked and we do not use the word 'evaluation' anywhere in the text. However, to ensure that this is not even implied, in the Abstract under Objectives, we have also replaced:
'To assess the impact of a National Dementia Strategy (NDS)' with 'To assess associations between the launch of the NDS …' (as suggested by reviewer 3).

We have also removed any use of the word failure etc (see response to Comment 2.3 above).

We are aware of the various publication dates; NDS (03.02.2009) followed by Professor Banerjee's 'Time for action' report (03.10.2009). However, the two were interconnected since Prof Banerjee was Joint-Lead of the NDS as well as author of the report. Also, the Department of Health response (12.11.2009) to Time for action stated that "reducing current levels of prescriptions, will from now on form part of the programme for implementing the Strategy" (point 9).

We also realise that the Time for action report was not specific to care homes. Although we would agree that two thirds of people with dementia in England do live in their own homes, the prevalence of dementia is much lower in the community (1.1% Knapp 2007 cited in ref 5) and most people are living with early stage disease which is less likely to exhibit BPSD. In contrast, the prevalence in care homes is much higher (~45.6% ref 5) and these individuals also have much later stage disease, with more likelihood of BPSD. As we point out, one artefact of earlier diagnosis in the community will be to

reduce the overall rate of antipsychotic prescribing because of changes in the case-mix (see penultimate sentence, para 1 in Discussion).

A national policy to reduce prescribing continues with the Prime Minister's challenge on dementia 2020 (added to reference list) which includes the aim of 'a continued significant reduction in inappropriate prescribing of antipsychotic medication for people with dementia and less variation across the country in prescribing levels' on page 36. On page 10 of Appendix 2, NHSE is identified as the lead organisation for this starting in March 2016, with impact to be measured against a target of 67% national reduction in antipsychotic prescribing by March 2019. Also, to emphasise the important role of regulatory agencies in successful implementation of the national strategy, we have added to following to 'Policy and research implications' (p15):

"the National Dementia Strategy did not include long term monitoring mechanisms, let alone enforcement mechanisms. Antipsychotic prescribing patterns in UK care homes are not open to public scrutiny nor routinely reported by regulatory inspection."

Reviewer 3: Dr Hans Wouters, Faculty Mathematics & Natural Sciences, University of Groningen, The Netherlands

Dr Wouters provides two final comments which are inter-linked & which we have now addressed:

(Comment 3.1) Personally, I still think you cannot talk about the impact of the NDS since you did not conduct an RCT (which was of course impossible to conduct and of course we cannot blame the authors for not conducting an RCT).

(Comment 3.2) While I liked the authors' strong arguments to my other points, I do not think it is a strong argument to say that you did not rephrase the Abstract because the other reviewers had no comments, because you could have given that answer to all of my other points as well, whereas you didn't.

A: We agree that an RCT would not be feasible in this case. In our previous response we were not ignoring Dr Wouters' comment but simply trying to make the point that, although he was sensitive to use of the term 'impact' and would have preferred 'association', the general reader (as exemplified by the other three reviewers) appeared not to be so sensitive.

However, Professor Banerjee has now raised a similar point about implied 'evaluation' (Comment 2.4 above) so we have now revised the Abstract using the wording originally suggested by Dr Wouters as follows:

Objectives: To assess associations between the launch of the National Dementia Strategy (NDS) and antipsychotic prescribing in long-term residential care (LTC) in England".

Reviewer 4: Dr Peter Watson, Medical Research Council, UK

Dr Watson provides a number of final comments, which we have hopefully now addressed satisfactorily:

(Comment 4.1) Having looked at the revised paper there appears to be no change in the analysis used in Table 3 on page 10 to examine care home characteristics in the baseline sample. It still

seems to me a more natural way of analysing this data which is hierarchical is to use a model which allows for this hierarchy namely of patients nested within care homes. In the previous review I mentioned that, for example, Barber ND et al. (2009) Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 18 341-346 used a multilevel analysis examining patients

A: We thank the reviewer for his suggestion to use multi-level modelling. We have not chosen to use this analysis technique because we are comparing characteristics at a care home level (i.e. the number of homes in Q1 and Q5).

Also, we do not have the individual level resident data in order to undertake multi-level modelling to examine the separate influences for both the patient and the home. Looking back at our original responses we may not have made this explicit. It is now made clear in the paragraph added to Discussion which outlines limitations.

Dr Watson raises an interesting point with respect to the Barber study (ref 36) which has relevance to the present discussion. This can be compared to our earlier study (ref 35). Barber studied a random sample of 256 residents in 55 care homes, collecting information by observing medication rounds and examining paper-based records e.g. MARS. Two medication rounds were observed per resident (morning & tea-time rounds only) on 1 or 2 (max) days per resident (a total of 512 resident medication rounds). Because of the manner in which data were collected, however, patient-level data were available and multi-level modelling was feasible and used.

Our study of medication administration was undertaken at a care home level (similar to the present study). It included all residents in receipt of medication in 13 care homes (345 residents). We also examined all administrations in every medication round (morning, lunch-time, tea-time & night time) on 84 days for each resident (~336 medication rounds per resident and a total of 188,249 medication administrations). However, as in the present study, because we did not have access to patient characteristics, other than gender and age, we could not include 'a multilevel analysis examining patients' as did Barber et al.

(Comment 4.2) As I also suggested previously an omnibus test comparing all levels of a factor is usually performed and further pairwise post-hoc tests as presented here (with the addition of a Bonferroni correction or similar for multiple testing) performed if the omnibus test is statistically significant. The results in the second paragraph on page 9 are based upon non-overlapping confidence intervals however the authors omit the result that dual registered (Table 3) residents have a Q5/Q1 confidence interval that does not overlap with residential and nursing home types suggesting, using their approach, that people are less likely to be in the highest quintile if they are living in dual registered homes.

A: We have amended Table 3 and added omnibus P-values for each factor variable as Dr Watson suggests.

(Comment 4.3) One could also add to the results on page 9 that the 75-84 year olds have a 95% CI in Table 3 (page 10) which does not overlap with either ages 65-74 or 85+ which using the non-overlapping CIs would suggest they are more likely to be in the upper quintile than those aged 85+ but less likely than those aged 65-74.

A: We have inserted additional text as recommended on p9:

"75-84 year olds have a 95% CI in which does not overlap with the other two groups suggesting they are more likely to be in the upper quintile than those aged 85+ but less likely than those aged 65-74."

(Comment 4.4) There is a greater 'degree' of non-overlap of most of the confidence intervals which are reported on page 9 but it does not follow that you can compare differences in group outcomes by whether or not the confidence intervals overlap which is a reason why we do F tests, for example, in ANOVAs rather than diving straight in to perform post-hoc pairwise contrasts.
I think the authors could at least mention in the paper why they have not used either omnibus tests to compare factor levels in Table 3 on page 10 or hierarchical multilevel modelling.

A: We have now included omnibus tests (see response to comment 4.2 above) and can confirm that all the omnibus tests were highly significant.

(Comment 4.5) Page 10. An addition to the footnote of Table 3 could state what the figures in brackets under the columns headed Q1 and Q5 represent. I thought these might be percentages of each factor in each level in each quintile but these don't add up to 100% e.g. for age in Q5 in Table 3 (page 10) the sum of the percentages is 12.6+39.4+37.5 = 89.5 < 100.

A: The percentages do not add up to 100% because there is a fourth age category <65 yrs (less than 4% of residents) which is not shown in the Table.

VERSION 3 – REVIEW

REVIEWER	Peter Watson Medical Research Council UK
REVIEW RETURNED	18-Apr-2016

GENERAL COMMENTS	Antipsychotic prescribing in care homes before and after launch of a national dementia strategy: an observational study in English institutions over a 4-year period. Bmjopen-2015-009882.R2 There are now p-values in Table 3 performing omnibus tests which is good. I have two outstanding queries. In the previous submissions I commented that given the hierarchical nature of the data with patients (residents) nested within care homes an appropriate analysis for formally comparing resident and care home characteristics would be the increasingly well used and prevalent technique of multilevel modelling. Multilevel models (or generalised estimating equations) have been used in previously published studies looking at care homes and residents (for example, Barber ND et al. (2009) Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 18 341-346) and strike me as appropriate for the statistical inference such as that presented in Table 3 on page 10. The multilevel analysis allows for correlations between patients within the same care home. I think you could at least mention in the statistical analysis section on page 5 if you have used these techniques and, if not, explain why you haven't used multilevel modelling or generalised estimating equations. Page 5, lines 36-40. On a related point does it now follow that since the categorisation of the quintiles is made at the care home level that comparing quintiles in Table 3 is a comparison only between care homes since all the residents within the care homes have the same quintile value so all the residents within each care home are

	all in the same quintile so the comparisons in Table 3 relating quintiles across patients (for gender and age) do not make sense?

VERSION 3 – AUTHOR RESPONSE

Thank you for your further report on responses of the 4 academics selected to review this manuscript. We are pleased that reviewers 1-3 (Professor Molinari, Professor Banerjee and Dr Wouters) are now happy with the manuscript and have no further suggestions to make. We thank them for their comments.

Below we provide a further response to the suggestion repeated by Reviewer 4 (Dr Peter Watson) that we undertake multilevel analysis of our baseline data, citing Barber at al as an example:

Multilevel models (or generalised estimating equations) have been used in previously published studies looking at care homes and residents (for example, Barber ND et al. (2009) Care homes' use of medicines study: prevalence, causes and potential harm of medication errors in care homes for older people. Qual Saf Health Care 18 341-346) and strike me as appropriate for the statistical inference such as that presented in Table 3 on page 10.

We have previously explained why multilevel analysis is not appropriate for our dataset i.e. due to insufficient patient-level data. We know the Barber et al study well & have similarly undertaken analyses of medication errors. However, the methods adopted in the Barber study for recruiting residents & gathering data are very different from our approach. As a result of using a mixed-methods approach, Barber et al were able to extract detailed patient-level information through "review of GP & care home notes and consultation with resident and/or staff" so that "multilevel models were used to measure patient-level odds of errors", including "examination of administration errors". In our case, the database we use contains no patient-level data other than age & gender (and we could not retrospectively access patient-level information from care homes). In order to further clarify differences between the two approaches to data capture, we have added Supplementary file 2 for your consideration, which should hopefully help to illustrate these.

Dr Watson also mentions "the comparisons in Table 3 relating quintiles across patients (for gender and age)". In fact we are still relating quintiles across care homes. To make this clearer we have moved the comparison of residents' age/gender profiles in care home quintiles to the bottom of Table 3.

All changes in the manuscript are highlighted in red text.