PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association between Daily Antiretroviral Pill Burden and Treatment Adherence, Hospitalization Risk, and Other Health Care Utilization and Costs in a United States Medicaid Population with HIV
AUTHORS	Davis, Keith; Cohen, Calvin; Meyers, Juliana

VERSION 1 - REVIEW

REVIEWER	Prepared by Kartika Palar and Sheri Weiser Kartika Palar, Phd Postdoctoral Fellow Department of Health Policy and Management Fielding School of Public Health University of California, Los Angeles, USA Sheri Weiser, MD MPH Assistant Professor of Medicine Division of HIV/AIDS, San Francisco General Hospital University of California, San Francisco, USA There are no conflicts of interest to declare for Drs. Weiser or Palar
REVIEW RETURNED	07-May-2013

THE STUDY	This study is generally well written and examines the association between single tablet regimens on ART adherence and hospitalization risk in a large database. A main strength of this study is its use of a large and detailed administrative database of Medicaid beneficiaries and their health care use. Yet, the paper has some significantly limitations that need to be addressed. Is the research question clearly defined? The paper aims to assess the effect of pill burden on adherence, hospitalizations, health care utilization, and health care costs among Medicaid enrollees. Yet there is inconsistency in how the research question is defined in the abstract, summary statement, and main article, which variably mention “adherence and hospitalizations”, or “health care utilization and costs”, or all four, as key outcomes. Based on the key messages in the Article Summary, and in the Discussion section, the reader also expects that the authors would also present data exploring whether adherence may be a mechanism for how a single tablet regimen

	(STR) affects hospitalizations compared to a multi-pill regimen (defined as two-or-more-pills-per-day, or 2+PPD). However, this data is not presented. It would strengthen the paper if the authors presented this data. Is the overall study design appropriate and adequate to answer the research question? With some revisions, the study design could be appropriate. The authors should clearly describe that they are looking for an association (and not an "effect") of pill burden on adherence, hospitalizations and health care costs. The selection bias issues between the two groups cause concern and have not been fully addressed. Since adherence seems to be a primary outcome, the authors should consider analyzing it using adjusted models. The adjusted analyses of pill count and hospitalizations and health care costs also seem to miss opportunities to control for confounding factors that can likely be measured from this dataset (see later comments in analysis section). Health care utilization (outside of hospitalizations) does not seem to be a primary outcome, and this should be clarified in the summary and introduction. Finally, the authors did not assess as far as I can tell an association between adherence and hospitalizations, but this is mentioned both in the key messages section and later in the paper. The authors should not include this if they did not assess it. Are the study participants adequately described, their conditions defined, and the inclusion and exclusion criteria described? Some descriptive information is missing on the study participants that could affect the study outcomes, including their race/ethnicity, whether they were Medicare dually eligible for all or part of the study period, and their Medicaid eligibility category. Are the patients representative of actual patients the evidence might affect? The paper would benefit from clarification on the sampling approach used to construct the overall MarketScan Medicaid database (is it constructed to be representative of all Medicaid participants, and if not, do we have information on who it might over or under-represent?) Are the methods adequately described? The measurement of adherence and the selection criteria are well described. However, there are some key pieces of information missing. First, there is not enough description of the cost data. Costs are not listed among the variables included in the database in the first paragraph in the Methods, so it is not clear whether costs came directly from the database or were linked using external sources. The authors should also better define what "costs" mean in the context of this paper. Are these Medicaid expenditures? Provider payments from all sources? Do they include patient co-pays?

	Second, the rationale is unclear for censoring STR patients when they switched regimens but not censoring patients on 2+PPD regimens when they switched . This can introduce bias because people who have switched regimens due to side effects and/or treatment failure may be likely to do worse than people who have not needed to switch regimens irrespective of the type of dosing that they use. Third, there is no rationale given for the covariates selected. Finally, there are some inconsistencies in how the methods are presented across the paper. On page 24 it says the regressions controlled for type of ART received and year the ART was received but these were not noted in the Methods, nor presented in the tables or results. Is the main outcome measure clear? The main outcome measures can be clarified as above. In the text, adherence is alternately presented as a main outcome, as a descriptive characteristic, and as a mechanism linking pill counts with hospitalizations. I would suggest that the role of adherence in the analysis be clarified. In addition, hospitalizations are inconsistently presented either as a stand-alone outcome, or as one outcome under health care utilization. The latter makes more sense – I would suggest presenting hospitalizations as the primary health care utilization outcome of interest; other data on health care utilization is primarily descriptive and can be presented as such. Are the abstract/summary/key messages/limitations accurate? Please see response above regarding research question, which is inconsistent across several parts of the paper. Other issues: Abstract: Authors should define “utilization and costs” which is too vague. Article Summary: Under “key messages”, the last two bullets are inaccurate. The authors never assessed whether adherence was statistically associated with hospitalization risk in this analysis (as noted in 2nd bullet). It is unclear why this point is emphasized in key messages, and also later in the paper. The third bullet says that hospitalization rates were higher in patients with lower adherence, but again, there is no data presented to support this. Are the statistical methods described and are they appropriate? In terms of choice of model, using poisson and negative binomial models with the hospitalization and cost data, respectively, is appropriate and preferable to an OLS approach. However, the hospitalization data has a very high number of zeros (Table 4), so one question is whether they considered using a zero-inflated or 2-part model to take this into account. Second, although the authors do acknowledge that unmeasured confounding could affect the results, and that they didn’t have access to laboratory data (e.g. CD4, viral load, etc), HIV-related health status is a potential confounder which could affect the

	outcomes. This should be addressed in the paper if possible. They controlled for general co-morbidities (such as cancer, heart failure, etc…) but they did not assess HIV-related opportunistic infections or common comorbidities like TB or hepatitis. Third, it is unclear why the adherence variable was excluded from the analysis on hospitalizations, especially since the authors make the point several times that adherence was associated with hospitalizations and suggest that this is one pathway whereby the STR may influence outcomes. It seems it would be simple to include adherence in the model and show the results with and without adherence, in order to provide evidence on these claims. Fourth, it is unclear why length of follow-up was included in the hospitalization model but not included in the cost model. Since the STR group was followed for less time, and costs naturally accrue over time, it seems it should be a covariate in the cost model as well. Fifth, given the structure of this data (Medicaid data from 11 states over 3 years), it was not clear whether they should have controlled for state and year effects in their multivariable models. If there was no reason to take these into account (i.e. if STR vs. 2+PPD were evenly distributed across states and years, and if there was no reason to think outcomes would vary by state or year), then this could simply be stated. However, it seems plausible that there may be differences in health care costs and utilization across states and years due to differential policies and contextual changes. Finally, given the statements in the abstract/intro, the authors should consider creating a multivariable model for the adherence outcome.
RESULTS & CONCLUSIONS	Do the results answer the research question? As above, the authors should better clarify the research question, and should be careful not to use causal language (ex: effect). Are they credible? The concerns in methods need to be addressed to add credibility. These results are generally in line with previous research. Are they well presented? The tables could be presented more clearly, and p-values would be helpful in comparing the two groups in Table 1 and Table 3. It is unclear why results from the cost regression were not presented in a table. Table 1 can be abbreviated, and they can cut down some of the detail on the age categories and general comorbidities. Table 3 can also be abbreviated. The purpose of providing so much data on each measure of health care utilization is not clear, especially when most of them were not significantly different between groups (and no hypotheses were presented to believe they would be significantly different). Another problem is the presentation of new results in the Discussion section. The authors present a sensitivity analysis (lines 392-402) – I would suggest that, following the convention for most

	medical and health publications, this should be noted in the Methods, and results presented in the Results section. Are the interpretation and conclusions warranted by and sufficiently derived from/focused on the data? I was confused about the claims regarding the associations between adherence and hospitalizations and costs in the Discussion section. For example, starting at line 359, the authors state "Our other finding was that higher rates of adherence were associated with similar or lower rates of hospitalization, regardless of the regimen; less than complete adherence was associated with higher rates of hospitalization and overall costs." I could not find any place in the paper that presented these results. They make similar statements elsewhere in the discussion, and highlight them in the "Key Messages" of the paper. Is the message clear? The authors should address the above limitations, and discrepancies in the objectives, methods and results of the paper to clarify the message. The point that non-pharmacy costs also seem to improve under STR could be further explored – this seemed like an interesting result that didn't get much attention in the Discussion.
GENERAL COMMENTS	This study is generally well written and examines the association between single tablet regimens on ART adherence and hospitalization risk in a large database. A main strength of this study is its use of a large and detailed administrative database of Medicaid beneficiaries and their health care use. Yet, the paper has some significantly limitations that need to be addressed. **Is the research question clearly defined?** The paper aims to assess the effect of pill burden on adherence, hospitalizations, health care utilization, and health care costs among Medicaid enrollees. Yet there is inconsistency in how the research question is defined in the abstract, summary statement, and main article, which variably mention "adherence and hospitalizations", or "health care utilization and costs", or all four, as key outcomes. Based on the key messages in the Article Summary, and in the Discussion section, the reader also expects that the authors would also present data exploring whether adherence may be a mechanism for how a single tablet regimen (STR) affects hospitalizations compared to a multi-pill regimen (defined as two-or-more-pills-per-day, or 2+PPD). However, this data is not presented. It would strengthen the paper if the authors presented this data. **Is the overall study design appropriate and adequate to answer**

	the research question? With some revisions, the study design could be appropriate. The authors should clearly describe that they are looking for an association (and not an "effect") of pill burden on adherence, hospitalizations and health care costs. The selection bias issues between the two groups cause concern and have not been fully addressed. Since adherence seems to be a primary outcome, the authors should consider analyzing it using adjusted models. The adjusted analyses of pill count and hospitalizations and health care costs also seem to miss opportunities to control for confounding factors that can likely be measured from this dataset (see later comments in analysis section). Health care utilization (outside of hospitalizations) does not seem to be a primary outcome, and this should be clarified in the summary and introduction. Finally, the authors did not assess as far as I can tell an association between adherence and hospitalizations, but this is mentioned both in the key messages section and later in the paper. The authors should not include this if they did not assess it. **Are the study participants adequately described, their conditions defined, and the inclusion and exclusion criteria described?** Some descriptive information is missing on the study participants that could affect the study outcomes, including their race/ethnicity, whether they were Medicare dually eligible for all or part of the study period, and their Medicaid eligibility category. **Are the patients representative of actual patients the evidence might affect?** The paper would benefit from clarification on the sampling approach used to construct the overall MarketScan Medicaid database (is it constructed to be representative of all Medicaid participants, and if not, do we have information on who it might over or under-represent?) **Are the methods adequately described?** The measurement of adherence and the selection criteria are well described. However, there are some key pieces of information missing. First, there is not enough description of the cost data. Costs are not listed among the variables included in the database in the first paragraph in the Methods, so it is not clear whether costs came directly from the database or were linked using external sources.

	The authors should also better define what "costs" mean in the context of this paper. Are these Medicaid expenditures? Provider payments from all sources? Do they include patient co-pays? Second, the rationale is unclear for censoring STR patients when they switched regimens but not censoring patients on 2+PPD regimens when they switched . This can introduce bias because people who have switched regimens due to side effects and/or treatment failure may be likely to do worse than people who have not needed to switch regimens irrespective of the type of dosing that they use. Third, there is no rationale given for the covariates selected. Finally, there are some inconsistencies in how the methods are presented across the paper. On page 24 it says the regressions controlled for type of ART received and year the ART was received but these were not noted in the Methods, nor presented in the tables or results. **Is the main outcome measure clear?** The main outcome measures can be clarified as above. In the text, adherence is alternately presented as a main outcome, as a descriptive characteristic, and as a mechanism linking pill counts with hospitalizations. I would suggest that the role of adherence in the analysis be clarified. In addition, hospitalizations are inconsistently presented either as a stand-alone outcome, or as one outcome under health care utilization. The latter makes more sense – I would suggest presenting hospitalizations as the primary health care utilization outcome of interest; other data on health care utilization is primarily descriptive and can be presented as such. **Are the abstract/summary/key messages/limitations accurate?** Please see response above regarding research question, which is inconsistent across several parts of the paper. Other issues: Abstract: Authors should define "utilization and costs" which is too vague. Article Summary: Under "key messages", the last two bullets are inaccurate. The authors never assessed whether adherence was statistically associated with hospitalization risk in this analysis (as noted in 2nd bullet). It is unclear why this point is emphasized in key messages, and also later in the paper. The third bullet says that hospitalization rates were higher in patients with lower adherence,

	but again, there is no data presented to support this. **Are the statistical methods appropriate?** In terms of choice of model, using poisson and negative binomial models with the hospitalization and cost data, respectively, is appropriate and preferable to an OLS approach. However, the hospitalization data has a very high number of zeros (Table 4), so it isn't quite clear why the authors didn't consider using a zero-inflated or 2-part model to take this into account. Second, although the authors do acknowledge that unmeasured confounding could affect the results, and that they didn't have access to laboratory data (e.g. CD4, viral load, etc), HIV-related health status is a potential confounder which could affect the outcomes. This should be addressed in the paper if possible. They controlled for general co-morbidities (such as cancer, heart failure, etc…) but they did not assess HIV-related opportunistic infections or common comorbidities like TB or hepatitis. Third, it is unclear why the adherence variable was excluded from the analysis on hospitalizations, especially since the authors make the point several times that adherence was associated with hospitalizations and suggest that this is one pathway whereby the STR may influence outcomes. It seems it would be simple to include adherence in the model and show the results with and without adherence, in order to provide evidence on these claims. Fourth, it is unclear why length of follow-up was included in the hospitalization model but not included in the cost model. Since the STR group was followed for less time, and costs naturally accrue over time, it seems it should be a covariate in the cost model as well. Fifth, given the structure of this data (Medicaid data from 11 states over 3 years), it was not clear whether they should have controlled for state and year effects in their multivariable models. If there was no reason to take these into account (i.e. if STR vs. 2+PPD were evenly distributed across states and years, and if there was no reason to think outcomes would vary by state or year), then this could simply be stated. However, it seems plausible that there may be differences in health care costs and utilization across states and years due to differential policies and contextual changes. Finally, given the statements in the abstract/intro, the authors should consider creating a multivariable model for the adherence

	outcome. **Do the results answer the research question?** As above, the authors should better clarify the research question, and should be careful not to use causal language (ex: effect). **Are they credible?** The concerns in methods need to be addressed to add credibility. These results are generally in line with previous research. **Are they well presented?** The tables could be presented more clearly, and p-values would be helpful in comparing the two groups in Table 1 and Table 3. It is unclear why results from the cost regression were not presented in a table. Table 1 can be abbreviated, and they can cut down some of the detail on the age categories and general comorbidities. Table 3 can also be abbreviated. The purpose of providing so much data on each measure of health care utilization is not clear, especially when most of them were not significantly different between groups (and no hypotheses were presented to believe they would be significantly different). Another problem is the presentation of new results in the Discussion section. The authors present a sensitivity analysis (lines 392-402) – I would suggest that, following the convention for most medical and health publications, this should be noted in the Methods, and results presented in the Results section. **Are the interpretation and conclusions warranted by and sufficiently derived from/focused on the data?** I was confused about the claims regarding the associations between adherence and hospitalizations and costs in the Discussion section. For example, starting at line 359, the authors state "Our other finding was that higher rates of adherence were associated with similar or lower rates of hospitalization, regardless of the regimen; less than complete adherence was associated with higher rates of hospitalization and overall costs." I could not find any place in the paper that presented these results. They make similar statements elsewhere in the discussion, and highlight them in the "Key Messages" of the paper.

	Is the message clear? The authors should address the above limitations, and discrepancies in the objectives, methods and results of the paper to clarify the message. The point that non-pharmacy costs also seem to improve under STR could be further explored – this seemed like an interesting result that didn't get much attention in the Discussion.

REVIEWER	Deborah Konkle-Parker, PhD, FNP Associate Professor Department of Medicine/Division of Infectious Diseases University of Mississippi Medical Center USA I have no competing interests
REVIEW RETURNED	14-May-2013

RESULTS & CONCLUSIONS	I think that the fact that the model was adjusted for ART experience is critically important, since prior ART experience is likely to be associated with drug resistance and pre-existing adherence difficulties. I felt that this was not emphasized enough to make it clear to the reader. I missed it my first time through and was ready to suggest a re-do of the analyses, but in my second read-through it became clear to me that it was already done. A stronger emphasis would make it easier for the reader and will make the results more credible.

VERSION 1 – AUTHOR RESPONSE

Reviewer: Prepared by Kartika Palar and Sheri Weiser

Kartika Palar, Phd
Postdoctoral Fellow
Department of Health Policy and Management Fielding School of Public Health University of California, Los Angeles, USA

Sheri Weiser, MD MPH
Assistant Professor of Medicine
Division of HIV/AIDS, San Francisco General Hospital University of California, San Francisco, USA

There are no conflicts of interest to declare for Drs. Weiser or Palar

This study is generally well written and examines the association between single tablet regimens on ART adherence and hospitalization risk in a large database. A main strength of this study is its use of a large and detailed administrative database of Medicaid beneficiaries and their health care use. Yet, the paper has some significantly limitations that need to be addressed.

Is the research question clearly defined?
The paper aims to assess the effect of pill burden on adherence, hospitalizations, health care utilization, and health care costs among Medicaid enrollees. Yet there is inconsistency in how the research question is defined in the abstract, summary statement, and main article, which variably mention "adherence and hospitalizations", or "health care utilization and costs", or all four, as key outcomes. Based on the key messages in the Article Summary, and in the Discussion section, the reader also expects that the authors would also present data exploring whether adherence may be a mechanism for how a single tablet regimen (STR) affects hospitalizations compared to a multi-pill regimen (defined as two-or-more-pills-per-day, or 2+PPD). However, this data is not presented. It would strengthen the paper if the authors presented this data.

AUTHORS' RESPONSE: We have made edits in the study title, abstract, summary statement, and in the introduction (last paragraph stating study objectives) to more consistently present the study objectives as suggested by the reviewer. Also, the reviewer is correct that the reader should expect to see data exploring the relationship between adherence status and hospitalization risk. We argue that improved adherence is a key mediating factor that confers lower hospitalization risk among STR patients, and we have now added a new figure (Figure 3; previous Figure 3 is now Figure 4) showing these data. Unfortunately this data presentation was inadvertently omitted from the original submission and we are grateful to the reviewer for pointing this out. We believe that the addition of this figure helps address many of the reviewer's subsequent comments.

Is the overall study design appropriate and adequate to answer the research question?
With some revisions, the study design could be appropriate. The authors should clearly describe that they are looking for an association (and not an "effect") of pill burden on adherence, hospitalizations and health care costs.

AUTHORS' RESPONSE: We have edited language throughout the manuscript (including the manuscript title) to de-emphasize "effects" and make more clear that our findings present "associations" as the reviewer correctly points out. Please also note that this limitation (i.e., concern regarding causal relationships) was already raised in our discussion section.

The selection bias issues between the two groups cause concern and have not been fully addressed. Since adherence seems to be a primary outcome, the authors should consider analyzing it using adjusted models.

AUTHORS' RESPONSE: The adherence benefit of STR vs. multi-tablet regimens is well known and extensively documented in the literature, including in other claims-based studies cited in our current manuscript (including Sax et al., 2012). Because assessing predictive factors of adherence (rather than confirming whether STR confers an adherence benefit, which was the key rationale for our analyses of adherence) was not an objective, we chose not to analyze adherence using regression

models.

The adjusted analyses of pill count and hospitalizations and health care costs also seem to miss opportunities to control for confounding factors that can likely be measured from this dataset (see later comments in analysis section).

AUTHORS' RESPONSE: Addressed in later comments.

Health care utilization (outside of hospitalizations) does not seem to be a primary outcome, and this should be clarified in the summary and introduction.

AUTHORS' RESPONSE: We have clarified in the introduction (lines 145-147) that non-hospital-related resource use and costs are a secondary objective.

Finally, the authors did not assess as far as I can tell an association between adherence and hospitalizations, but this is mentioned both in the key messages section and later in the paper. The authors should not include this if they did not assess it.

AUTHORS' RESPONSE: See response to reviewer's first comment.

Are the study participants adequately described, their conditions defined, and the inclusion and exclusion criteria described?
Some descriptive information is missing on the study participants that could affect the study outcomes, including their race/ethnicity, whether they were Medicare dually eligible for all or part of the study period, and their Medicaid eligibility category.

AUTHORS' RESPONSE: As requested, descriptive data on these additional variables has been added to Table 1. However, to conserve space and maintain an easily "readable" Table 1, we have eliminated the rows showing the detailed distribution of age categories. We feel that it is sufficient to show mean age, especially considering that mean age and distribution of specific categories were similar between the two study cohorts.

Are the patients representative of actual patients the evidence might affect?
The paper would benefit from clarification on the sampling approach used to construct the overall MarketScan Medicaid database (is it constructed to be representative of all Medicaid participants, and if not, do we have information on who it might over or under-represent?)

AUTHORS' RESPONSE: Unfortunately we do not have data to inform these questions. Although we do know that the database includes information from the Medicaid programs of 11 states, we are blinded (as per data privacy rules) as to which states are included. The vendor providing these data, however, does reveal that the states are "geographically dispersed" (see line 151 under Methods). Nonetheless, we still cannot make assessments regarding the generalizability of our data to the overall Medicaid population in the US. We have now added additional material to our discussion of limitations (see lines 522-526) to highlight this point.

Are the methods adequately described?
The measurement of adherence and the selection criteria are well described. However, there are some key pieces of information missing.
First, there is not enough description of the cost data. Costs are not listed among the variables included in the database in the first paragraph in the Methods, so it is not clear whether costs came directly from the database or were linked using external sources.

AUTHORS' RESPONSE: We thank the reviewer for noting this omission. All cost data represent payments from Medicaid to providers as captured directly on each medical or prescription claim. No cost data are linked from external sources. We have clarified this point in the Methods section (see lines 154-156). We also note, additionally, in lines 248-250 that all cost data were standardized at the claim level to 2010 US dollars using the medical care component of the US Consumer Price Index.

The authors should also better define what "costs" mean in the context of this paper. Are these Medicaid expenditures? Provider payments from all sources? Do they include patient co-pays?

AUTHORS' RESPONSE: We have added text to line 248 clarifying that costs in our study represent the Medicaid perspective as per our response to the comment above. The database does include, in some instances, patient copayments for certain claims. However, because these data are not consistently available for all types of claims and considering that our focus was costs borne by the payer (Medicaid), we did not include copayments in our cost analyses (this is a widely adopted practice in claims-based cost studies).

Second, the rationale is unclear for censoring STR patients when they switched regimens but not censoring patients on 2+PPD regimens when they switched . This can introduce bias because people who have switched regimens due to side effects and/or treatment failure may be likely to do worse than people who have not needed to switch regimens irrespective of the type of dosing that they use.

AUTHORS' RESPONSE: We believe that the reviewer has misunderstood the description of methods on line 187. Patients on a 2+PPD regimen may switch agents comprising the regimen, as long as the combination of agents they are prescribed can continue to be classified as a 2+PPD regimen (i.e., as long as the patient does not leave this "arm" of the study). If the 2+PPD patients switch to STR or cease receiving a complete 2+PPD regimen (regardless of specific medications interchanged within the regimen), then they are no longer in the study arm and are thus censored. We have added text to lines 187-188 for additional clarity. For STR patients, any switch to another regimen (or discontinuation of STR) means that they have left that study "arm" and are therefore censored at that point.

Third, there is no rationale given for the covariates selected.

AUTHORS' RESPONSE: We have added a note to line 263 that the covariate list was selected based on the previous paper by Sax et al. The authors feel that these covariates (particularly demographics) are commonly accepted in all areas of health services research and do not need a case-by-case

rationale.

Finally, there are some inconsistencies in how the methods are presented across the paper. On page 24 it says the regressions controlled for type of ART received and year the ART was received but these were not noted in the Methods, nor presented in the tables or results.

AUTHORS' RESPONSE: Type of ART (STR or 2+PPD) is included in the Poisson and GLM models for hospitalization incidence and costs, respectively, and is presented in the corresponding results table. So, our description is accurate on this covariate. The mention of the year that ART was received was an error and has been removed from the manuscript.

Is the main outcome measure clear?
The main outcome measures can be clarified as above. In the text, adherence is alternately presented as a main outcome, as a descriptive characteristic, and as a mechanism linking pill counts with hospitalizations. I would suggest that the role of adherence in the analysis be clarified. In addition, hospitalizations are inconsistently presented either as a stand-alone outcome, or as one outcome under health care utilization. The latter makes more sense – I would suggest presenting hospitalizations as the primary health care utilization outcome of interest; other data on health care utilization is primarily descriptive and can be presented as such.

AUTHORS' RESPONSE: See response to first comment in this document.

Are the abstract/summary/key messages/limitations accurate?
Please see response above regarding research question, which is inconsistent across several parts of the paper. Other issues:

Abstract: Authors should define "utilization and costs" which is too vague.

AUTHORS' RESPONSE: We have clarified in the abstract that utilization refers to "health care" utilization. We do not have space in the abstract to elaborate further. We believe that this will be understood by most readers.

Article Summary: Under "key messages", the last two bullets are inaccurate. The authors never assessed whether adherence was statistically associated with hospitalization risk in this analysis (as noted in 2nd bullet). It is unclear why this point is emphasized in key messages, and also later in the paper. The third bullet says that hospitalization rates were higher in patients with lower adherence, but again, there is no data presented to support this.

AUTHORS' RESPONSE: See response to first comment in this document.

Are the statistical methods described and are they appropriate?

In terms of choice of model, using poisson and negative binomial models with the hospitalization and cost data, respectively, is appropriate and preferable to an OLS approach. However, the hospitalization data has a very high number of zeros (Table 4), so one question is whether they

considered using a zero-inflated or 2-part model to take this into account.

AUTHORS' RESPONSE: Poisson models are precisely intended to account for severe skewness in count data that feature a large concentration at zero. Although 2-part or zero-inflated models may also be appropriate, we don't believe (at the sample sizes available) that the results will differ. More importantly, we believe that the Poisson results will be more accessible and intuitive to the general clinical audience of BMJ Open HIV/AIDS. For these reasons, we insist on maintaining the Poisson model specification.

Second, although the authors do acknowledge that unmeasured confounding could affect the results, and that they didn't have access to laboratory data (e.g. CD4, viral load, etc), HIV-related health status is a potential confounder which could affect the outcomes. This should be addressed in the paper if possible. They controlled for general co-morbidities (such as cancer, heart failure, etc…) but they did not assess HIV-related opportunistic infections or common comorbidities like TB or hepatitis.

AUTHORS' RESPONSE: The CCI was developed specifically to be a predictive measure of hospitalization and mortality, and we believe that the CCI adequately controls for patients general health state at baseline, including health state effects due to opportunistic infections. Also, our sensitivity analysis restricting the sample to patients with a CCI of 0 showed no effect on the overall results.

Third, it is unclear why the adherence variable was excluded from the analysis on hospitalizations, especially since the authors make the point several times that adherence was associated with hospitalizations and suggest that this is one pathway whereby the STR may influence outcomes. It seems it would be simple to include adherence in the model and show the results with and without adherence, in order to provide evidence on these claims.

AUTHORS' RESPONSE: As noted in the earlier comments, we show that improved adherence is a key mediating factor for lower hospitalization risk among STR patients. In this respect, adherence and STR may be considered proxies for each other. As a result, the two variables (as we have shown) are highly collinear. For these reasons, and based on the study question sought to be addressed, there is no suitable rationale for including both variables in the model.

Fourth, it is unclear why length of follow-up was included in the hospitalization model but not included in the cost model. Since the STR group was followed for less time, and costs naturally accrue over time, it seems it should be a covariate in the cost model as well.

AUTHORS' RESPONSE: Costs were normalized at the level of per patient per month before the regressions were estimated, so this removes the incongruity the reviewer mentions, as well as the need for a follow-up time covariate in the cost model.

Fifth, given the structure of this data (Medicaid data from 11 states over 3 years), it was not clear whether they should have controlled for state and year effects in their multivariable models. If there was no reason to take these into account (i.e. if STR vs. 2+PPD were evenly distributed across states

and years, and if there was no reason to think outcomes would vary by state or year), then this could simply be stated. However, it seems plausible that there may be differences in health care costs and utilization across states and years due to differential policies and contextual changes.

AUTHORS' RESPONSE: See previous comments. Unfortunately we do not know (per data privacy rules) which states are included in the database. We only know that it includes 11 states. We have added text, as previously noted, to the limitations discussion to highlight this point.

Finally, given the statements in the abstract/intro, the authors should consider creating a multivariable model for the adherence outcome.

AUTHORS' RESPONSE: See previous comments on rationale for not analyzing adherence using a multivariate model.

Do the results answer the research question?
As above, the authors should better clarify the research question, and should be careful not to use causal language (ex: effect).

AUTHORS' RESPONSE: See response to previous comments regarding causal language.

Are they credible?
The concerns in methods need to be addressed to add credibility. These results are generally in line with previous research.

AUTHORS' RESPONSE: We agree that our results being in line with previous research indicates credibility of the current findings.

Are they well presented?
The tables could be presented more clearly, and p-values would be helpful in comparing the two groups in Table 1 and Table 3.

AUTHORS' RESPONSE: P-values are now included in Table 1 and Table 3.

It is unclear why results from the cost regression were not presented in a table.

AUTHORS' RESPONSE: Separate GLM models were estimated for every cost endpoint displayed in the table, and considering both space limitations and that the adjusted, predicted value of cost was of most interest (and is what is typically displayed from these models), we believe it is most useful to present only the adjusted estimates derived from the cost models.

Table 1 can be abbreviated, and they can cut down some of the detail on the age categories and general comorbidities.

AUTHORS' RESPONSE: This has been done – substantial cuts have been made.

Table 3 can also be abbreviated. The purpose of providing so much data on each measure of health care utilization is not clear, especially when most of them were not significantly different between groups (and no hypotheses were presented to believe they would be significantly different).

AUTHORS' RESPONSE: We believe the reviewer was referring to Table 4, not Table 3. We agree with the reviewer on this point and have abbreviated the table substantially.

Another problem is the presentation of new results in the Discussion section. The authors present a sensitivity analysis (lines 392-402) – I would suggest that, following the convention for most medical and health publications, this should be noted in the Methods, and results presented in the Results section.

AUTHORS' RESPONSE: The reviewer's point is well taken. However, because the sensitivity analysis was not planned a priori, and instead was motivated by results obtained from the original analysis, we feel that the conception and discussion of the sensitivity analysis more appropriately belongs in the Discussion section of the manuscript. We hope that this is acceptable to the reviewer.

Are the interpretation and conclusions warranted by and sufficiently derived from/focused on the data?
I was confused about the claims regarding the associations between adherence and hospitalizations and costs in the Discussion section. For example, starting at line 359, the authors state "Our other finding was that higher rates of adherence were associated with similar or lower rates of hospitalization, regardless of the regimen; less than complete adherence was associated with higher rates of hospitalization and overall costs." I could not find any place in the paper that presented these results. They make similar statements elsewhere in the discussion, and highlight them in the "Key Messages" of the paper.

AUTHORS' RESPONSE: See previous responses to first comment in this document.

Is the message clear?
The authors should address the above limitations, and discrepancies in the objectives, methods and results of the paper to clarify the message.

AUTHORS' RESPONSE: See previous responses.

The point that non-pharmacy costs also seem to improve under STR could be further explored – this seemed like an interesting result that didn't get much attention in the Discussion.

AUTHORS' RESPONSE: We do note, in fact, that hospitalization costs (which are non-pharmacy) are the key cost driver and, as such, feel that this is sufficient for purposes of the discussion.

Reviewer: Deborah Konkle-Parker, PhD, FNP Associate Professor Department of Medicine/Division of Infectious Diseases University of Mississippi Medical Center USA

I have no competing interests

I think that the fact that the model was adjusted for ART experience is critically important, since prior ART experience is likely to be associated with drug resistance and pre-existing adherence difficulties. I felt that this was not emphasized enough to make it clear to the reader. I missed it my first time through and was ready to suggest a re-do of the analyses, but in my second read-through it became clear to me that it was already done. A stronger emphasis would make it easier for the reader and will make the results more credible.

AUTHORS' RESPONSE: Thank you for the comment. We have added additional text to the discussion section regarding the adjustment for ART experience (treatment naïve status).