PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Real-world outcomes of US employees with type 2 diabetes mellitus treated with insulin glargine or neutral protamine Hagedorn insulin: A comparative retrospective database study
AUTHORS	Wang, Li; Wei, Wenhui; Miao, Raymond; Xie, Lin; Baser, Onur

VERSION 1 - REVIEW

REVIEWER	Carrie McAdam Marx, RPh, PhD Research Assistant Professor Pharmacotherapy Outcomes Research Center University of Utah Salt Lake City, Utah Participating in a research study in diabetes that has received research funding in the past 12 months from Bristol Meyers Squibb. Served as a consultant in the past 12 months for GlaxoSmithKline and NovoNordisk.
REVIEW RETURNED	23-Jan-2013

THE STUDY	The limitation of data to those with productivity data and in reduction of glargine cohort via propensity score matching likely reduced the external validity of the study cohort over and above the limitations in external validity of MarketScan data. Methods for assessing persistence and adjusted MPR lack clarity. Conclusion is generally not supported by the data.
RESULTS & CONCLUSIONS	The conclusions are not supported by the data. The message in the conclusion is clouded by the suggestion that non-signficant differences in utilization and short-term disability are associated with glargine use.
REPORTING & ETHICS	No mention of IRB review.
GENERAL COMMENTS	This study provides a perspective on type 2 diabetes treatment outcomes comparing insulin glargine to NPH insulin in by considering reduction in productivity through assessment of short-term disability costs. However, the study conclusions are not supported by the study data, and need to be revised and several liberal applications of the term “significant” should be removed. Assuming this will be addressed, the study is reasonably well designed, although I have a concern about whether the study was adequately powered given the variability in cost and utilization. I recommend adding a power analyses, a statement of the a priori level of significance, possibly the use of 95% confidence intervals, and the avoidance of the use of the term “marginally significant”. It would be helpful to include a more comprehensive discussion of the differences in all-cause hospitalizations and endocrinology visits in light of the finding that diabetes-related use (except medications) did

	not differ, as this finding could be related to uncontrolled confounding and not insulin treatment. Additional, specific comments are provided below. Abstract: Page 3, line 11 – Please clarify that employees are those with employer sponsored health insurance coverage. Page 3, line 24 – Sentence starting “Adult employees” is incomplete, Page 3, line 30 (and elsewhere). Propensity score matching helps to reduce the risk of confounding by indication, but does not prevent selection bias. Page 4, line 4 – Suggest going beyond reiterating results, but conclusion needs to be supported by the data (see comments specific to the conclusion.) Introduction General – the introduction is long and could be more focused. Page 4, line 11 –United States not Unites States Page 4 line 21 – Several references in this section (e.g., line 35 as well) are in general, quite old. Please consider adding more current references, if available. Page 5, line 10 – another possible option for basal insulin is detemir. Is there a reason it is not mentioned? Please discuss why the analyses was limited to glargine and NPH. This section is somewhat misleading by suggesting glargine is always once a day and NPH is generally twice a day. Please mention that glargine is approved for twice daily dosing, and if available, the proportion of patients on twice daily dosing vs. daily dosing for these 2 insulins. Methods- please clearly state the perspective of this analysis (employee/employer/society Page 6, line 17 – MarketScan data includes information on medical/pharmacy claims not patient medical records. Please add additional detail about the data, such as geographical locations and age, to assess external validity. Page 6, line 37. For clarity, I would add that by including the productivity data, this study includes only employees and not dependents. Page 6, line 49. Please discuss whether 90 days is sufficient to assess prior insulin exposure considering the use of mail-order and/or 90-day retail fills in this population. Page 7, line 6. Please clarify timeframe for cost/utilization and disability data – is it limited to 90 days for all patients or does it include all pre-index data for those with greater than 90 days of pre-index eligibility? Page 7 line 19. Please clarify how persistence was evaluated. The description is not clear – could possibly benefit from a simplified explanation or an example. This feeds into the Kaplan-Meyer curve – in which there is no persistence drop off for 90-days, but from this description, it is not clear as to why this occurs. Page 7, line 31. For readers unfamiliar with the adjusted MPR, please clarify if the adjustment factor is based on only insulin prescriptions, all anti-diabetics, all insulins, etc. Please add more justification as to how this accounts for differences in product size/packaging. Page 7, line 46. Please discus the accuracy of DACON estimating insulin dose. For those on low dose, and who following instructions to discard unused insulin (particularly in vials) after approximately 1 month, dispensing/claims data can lead to an overestimation of

	DACON. To help add transparency to this issue, I suggest reporting DACON for vial/syringe and Pen patients separately. Page 8, line 41. Again reference to selection bias – PSM helps to address confounding by indication. Page 9, end of methods section. Suggest adding a sample size analyses or power analysis. Also, please add a statement your a priori level of significance. Was the study protocol reviewed by an IRB? Results Page 9, line 18. How did the PSM matched cohort differ from the pre-matched cohort? This is of particular concern given the large reduction in glargine cohort as a result of matching. Page 9/10, Table 1. It would be useful to remind readers of duration of time used to collect baseline characteristics (3 months prior to index date). Suggest adding median values or interquartile range for cost and utilization data, and the time frame for data if it differs from assessment of other baseline characteristics. Please add p-value for pen vs. vial/syringe use. Page 11 Table 2. Consider adding median number of days of persistence and report DACON by vial/syringe vs. pen. Page 11, line 53. Diabetes related costs/used reported intermittently, please discuss as related to hospitalization and endocrinology visits. Because all cause hospitalization and endocrinology visits differ, it is also important to know if they differ specifically for diabetes-related care – which they did not. Page 12, line 27. Please remove suggestion that total costs favored glargine – p-value suggests this difference could be due to chance and not a true difference. Page 12, line 30. Unsure what this correlation analyses adds to the literature – hospitalization association with short-term disability seems obvious. Discussion The discussion section needs to address why statistical difference was seen in all-cause hospitalization and endocrinology visits but not for diabetes specific use. Page 14 line 4. Should clarify that the study findings are relevant to glargine patients when compared to the NPH regimen. The statement could be misinterpreted much more broadly. Page 14, line 7. Recommend that you discuss the limitations of estimating DACON for insulin due to discarded product with open vial/pen expiration. Please also discuss limitations to external validity based on the MarketScan population overall (mainly large, self-funded employers with a heavy mid-west presence) further limited by including only patients whose employer supplies productivity data, and by using PSM which likely limited glargine patients to those most similar to NPH patients and not insulin-using T2DM patients. Also discuss potential of selection bias with sampling and PSM approach, and the implications on study findings. Conclusion Page 14, line 31. The opening conclusion is not supported by the data. Glargine's association with healthcare utilization is more narrow than the statement suggests as the association was only seen with all-cause hospitalization and endocrinology visits, not in ER, LOS or for any diabetes-related use, or costs except diabetes medications. Short term disability differences were also not established per typical statistical practices. Concluding that glargine is a cost effective treatment option is also beyond the scope of this study as treatment effectiveness was not evaluated. Thus, consider

	reworking conclusion to fit the study and the data.

REVIEWER	Katherine Esposito, MD, PhD Associate Professor of Endocrinology and Metabolism, Second University of Naples, Italy I have no conflict of interest with this paper Katherine Esposito, MD, PhD
REVIEW RETURNED	04-Feb-2013

THE STUDY	As stated in limitations, it should not be assumed that the sample is representative of the overall US population. The references quoted are in general relevant but some references are missing, in particular some detailing the use of long-acting insulins. I suggested some reference to include and some to delete.
GENERAL COMMENTS	The authors aimed to compare, in the real-world, persistence and adherence to insulin therapy among employees with type 2 diabetes initiating insulin therapy with glargine versus NPH insulin. They found that employees with T2DM initiating glargine were more persistent and adherent with their treatment than those initiating NPH insulin. My comments: 1. Introduction is too long and somewhat dispersive. There are too many cost details, which are not analyzed individually thereafter in discussion. My suggestion is to cut introduction by nearly 40-50%. I also suggest to shorten the long reference list of costs. On page 5, second para, the authors introduce an important issue, i.e. semplicity of basal insulin regimen which may affect adherence to therapy. However, options for basal insulin supplementation worldwide are much more than those described by authors, and only one reference (ref. 18) is too scanty for this important issue. Please add some recent work (i.e. Diabetes Care 2012;35:2698-705; Diabetes Care 2012;35:1364-1379; Expert Opin Biol Ther. 2012;12:1541-50.). I agree that once daily dosing is the best approach for initiating basal insulin supplementation in type 2 diabetic patients no more responsive to oral drugs. 2. Another important issue of this retrospective study is the lack of difference in hypos between groups. There may be several reasons for this, some addressed by authors. I would like to suggest another one which may be relevant to the issue: as the switch from NPH to glargine is normally favored by evenience of hypo or increasing hypos, the NPH-using population in the present analysis may be that with basically low hypoglycemia with NPH, explaining the lack of difference. 3. Perhaps I missed daily insulin doses. 4. In Discussion, Ref 41 appears too solitary in describing the complex issue of insulin analogs. Please, alslo use the references indicated above. 5. Limitations. The many limitations of this study are rightly addressed by the authors.

VERSION 1 – AUTHOR RESPONSE

Dr McAdam's Comments

General:

The effect of PSM on external validity has been further emphasized in the limitations section. The methods for assessing persistence and adjusted MPR have been clarified in the methods section. We have rewritten the conclusion to both the main manuscript and abstract so that it is more firmly based on the findings of the study, reference to non-significant differences have been removed. The instances of the term 'significant' which caused concern have been removed.

Specific Comments
Abstract:
1) Page 3, line 11 – Please clarify that employees are those with employer sponsored health insurance coverage.
Clarification added to objective section of abstract.

2) Page 3, line 24 – Sentence starting "Adult employees" is incomplete, Page 3, line 30 (and elsewhere). Propensity score matching helps to reduce the risk of confounding by indication, but does not prevent selection bias.

Instances of the use of selection bias have been changed to confounding by indication throughout.

3) Page 4, line 4 – Suggest going beyond reiterating results, but conclusion needs to be supported by the data (see comments specific to the conclusion.)

The conclusion has been amended.

Introduction:
1) General – the introduction is long and could be more focused.

The introduction has been shortened significantly.

2) Page 4, line 11 –United States not Unites States Page 4 line 21

Typo amended.

3) Several references in this section (e.g., line 35 as well) are in general, quite old. Please consider adding more current references, if available.

Ref #4 (National US diabetes cost statistics) has the most recent national data – CDC fact sheet of 2011, published May 2012, quotes 2007 data. More recent ref's replace older ones where available. Ref list amended.

4) Page 5, line 10 – another possible option for basal insulin is detemir. Is there a reason it is not mentioned? Please discuss why the analyses was limited to glargine and NPH. This section is somewhat misleading by suggesting glargine is always once a day and NPH is generally twice a day.

N number of detemir was too small as it was not available until 2005, note added to Cohort selection criteria section of methods. Added reference detailing that insulin glargine can be used twice daily.

5) Please mention that glargine is approved for twice daily dosing, and if available, the proportion of

patients on twice daily dosing vs. daily dosing for these 2 insulins.

Added reference detailing that insulin glargine can be used twice daily.

Methods:

1) Please clearly state the perspective of this analysis (employee/employer/society

Added to database section of the methods and end of introduction.

2) Page 6, line 17 – MarketScan data includes information on medical/pharmacy claims not patient medical records.

Copy amended.

3) Please add additional detail about the data, such as geographical locations and age, to assess external validity.

Data added into Table 1.

4) Page 6, line 37. For clarity, I would add that by including the productivity data, this study includes only employees and not dependents.

Added to cohort selection criteria section of methods.

5) Page 6, line 49. Please discuss whether 90 days is sufficient to assess prior insulin exposure considering the use of mail-order and/or 90-day retail fills in this population.

Possibility of missing patients for this reason added to limitations.

6) Page 7, line 6. Please clarify timeframe for cost/utilization and disability data – is it limited to 90 days for all patients or does it include all pre-index data for those with greater than 90 days of pre-index eligibility?

Timeframe was limited to 90 days for all patients, clarification added to baseline characteristics section of methods.

7) Page 7 line 19. Please clarify how persistence was evaluated. The description is not clear – could possibly benefit from a simplified explanation or an example. This feeds into the Kaplan-Meyer curve – in which there is no persistence drop off for 90-days, but from this description, it is not clear as to why this occurs.

Clarified in persistence and adherence section of methods.

8) Page 7, line 31. For readers unfamiliar with the adjusted MPR, please clarify if the adjustment factor is based on only insulin prescriptions, all anti-diabetics, all insulins, etc. Please add more justification as to how this accounts for differences in product size/packaging.

Only insulin glargine/NPH dispensed are included, clarified in persistence and adherence section of methods.

9) Page 7, line 46. Please discus the accuracy of DACON estimating insulin dose. For those on low dose, and who following instructions to discard unused insulin (particularly in vials) after approximately 1 month, dispensing/claims data can lead to an overestimation of DACON. To help add transparency to this issue, I suggest reporting DACON for vial/syringe and Pen patients separately.

Added to limitation section, unlikely to confound due to similar level of pen use between groups.

10) Page 8, line 41. Again reference to selection bias – PSM helps to address confounding by indication.

Copy amended.

11) Page 9, end of methods section. Suggest adding a sample size analyses or power analysis. Also, please add a statement your a priori level of significance. Was the study protocol reviewed by an IRB?

The power analysis suggested a sample size of 566 patients in GLA group and 283 patients in NPH group would be needed to detect a difference of 10% in persistence rates, with 80% power at two-tailed α of 0.05.

As power analysis is typically included in the statistical analysis plan, instead of adding it to the manuscript, we added a statement of a priori level of significance, and also in the limitation section, acknowledging the small sample size in this study may make it unable to detect significant differences.

This study data is fully HIPPA-compliant and IRB was waived. A statement was added at the end of data source section

Results:

1) Page 9, line 18. How did the PSM matched cohort differ from the pre-matched cohort? This is of particular concern given the large reduction in glargine cohort as a result of matching.

Added to beginning of results section.

2) Page 9/10, Table 1. It would be useful to remind readers of duration of time used to collect baseline characteristics (3 months prior to index date).

Added to title.

3) Suggest adding median values or interquartile range for cost and utilization data, and the time frame for data if it differs from assessment of other baseline characteristics. Please add p-value for pen vs. vial/syringe use.

Data added.

4) Page 11 Table 2. Consider adding median number of days of persistence and report DACON by vial/syringe vs. pen.

Data not available.

5) Page 11, line 53. Diabetes related costs/used reported intermittently, please discuss as related to hospitalization and endocrinology visits. Because all cause hospitalization and endocrinology visits differ, it is also important to know if they differ specifically for diabetes-related care – which they did not.

Discussion relates costs to utilization, amended to include endocrinologist visits.

6) Page 12, line 27. Please remove suggestion that total costs favored glargine – p-value suggests this difference could be due to chance and not a true difference.

Copy amended

7) Page 12, line 30. Unsure what this correlation analyses adds to the literature – hospitalization association with short-term disability seems obvious.

We would like to keep this section as confirms that disability payments are directly related to hospitalization and not other factors. Also, added text on correlation of persistence.

Discussion:

1) The discussion section needs to address why statistical difference was seen in all-cause hospitalization and endocrinology visits but not for diabetes specific use.

Added to first paragraph of discussion.

2) Page 14 line 4. Should clarify that the study findings are relevant to glargine patients when compared to the NPH regimen. The statement could be misinterpreted much more broadly.

Paragraph 3 of discussion amended.

3) Page 14, line 7. Recommend that you discuss the limitations of estimating DACON for insulin due to discarded product with open vial/pen expiration.

Added as mentioned above.

4) Please also discuss limitations to external validity based on the MarketScan population overall (mainly large, self-funded employers with a heavy mid-west presence) further limited by including only patients whose employer supplies productivity data, and by using PSM which likely limited glargine patients to those most similar to NPH patients and not insulin-using T2DM patients.

Issues addressed in limitations section.

5) Also discuss potential of selection bias with sampling and PSM approach, and the implications on study findings.

Addressed in limitations section.

Conclusion:

1) Page 14, line 31. The opening conclusion is not supported by the data. Glargine's association with

healthcare utilization is more narrow than the statement suggests as the association was only seen with all-cause hospitalization and endocrinology visits, not in ER, LOS or for any diabetes-related use, or costs except diabetes medications. Short term disability differences were also not established per typical statistical practices. Concluding that glargine is a cost effective treatment option is also beyond the scope of this study as treatment effectiveness was not evaluated. Thus, consider reworking conclusion to fit the study and the data.

Conclusion has been reworded as stated above.

Dr Esposito's Comments:

1) Introduction is too long and somewhat dispersive. There are too many cost details, which are not analyzed individually thereafter in discussion. My suggestion is to cut introduction by nearly 40-50%. I also suggest to shorten the long reference list of costs.

Introduction has been shortened significantly.

2) On page 5, second para, the authors introduce an important issue, i.e. simplicity of basal insulin regimen which may affect adherence to therapy. However, options for basal insulin supplementation worldwide are much more than those described by authors, and only one reference (ref. 18) is too scanty for this important issue. Please add some recent work (i.e. Diabetes Care 2012;35:2698-705; Diabetes Care 2012;35:1364-1379; Expert Opin Biol Ther. 2012;12:1541-50.). I agree that once daily dosing is the best approach for initiating basal insulin supplementation in type 2 diabetic patients no more responsive to oral drugs.

We have added information of alternative insulins available and used references suggested.

3) Another important issue of this retrospective study is the lack of difference in hypos between groups. There may be several reasons for this, some addressed by authors. I would like to suggest another one which may be relevant to the issue: as the switch from NPH to glargine is normally favored by evidence of hypo or increasing hypos, the NPH-using population in the present analysis may be that with basically low hypoglycemia with NPH, explaining the lack of difference.

We have addressed this in the limitations section.

4) Perhaps I missed daily insulin doses.

Doses are estimated by DACON and reported in tables and text.

5) In Discussion, Ref 41 appears too solitary in describing the complex issue of insulin analogs. Please, also use the references indicated above.

References added.

VERSION 2 – REVIEW

REVIEWER	Carrie McAdam Marx, PhD, RPh. Research Assistant Professor, University of Utah Pharmacotherapy Outcomes Research Center, USA. In the past 12 months, I have received consulting honoraria from GlaxoSmithKline and NovoNordisk.
REVIEW RETURNED	11-Mar-2013

THE STUDY	While the "content" necessary to establish the background and rationale for the study is in the background section, it continues to be lack focus and flow. Rearranging text would help set up the research question better (i.e., burden of illness, benefit of good glycemic control, role of adherence, current treatment options, specific role of basal insulins, benefits of basal insults per the literature, and hypothesis about glargine vs. NPH in regards to adherence and outcomes. While mentioned in the methods, it needs to be explained why only NPH and glargine are included (i.e., too few patients on other basal insulins,namely detemir.) Also, the use of terms Selection Bias and Confounding by Indication is still not sufficiently clear. These are two distinct issues/sources of bias in observational studies. Selection bias is defined as " a distorted estimate of the effect that results from the way in which subjects are ascertained or selected for the study population and includes factors such as differential surveillance, diagnosis, and referral of persons into the study". PSM serves to balance differences between two groups included in the study (which can lead to other bias, such as confounding by indication) and not address issues between those who were and who were not included in the study. Please see: Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. Am J Epidemiol. 1999 Jun 1;149(11):981-3. PubMed PMID: 10355372.
GENERAL COMMENTS	This is an interesting paper. I have few concerns remaining about the study itself and have only recommended a few areas where edits could help flow and clarity.

REVIEWER	Katherine Esposito, MD, PhD Second University of Naples Italy Conflict of interest: NONE
REVIEW RETURNED	05-Mar-2013

RESULTS & CONCLUSIONS	The results section still appears too long. I suggest reduction by say 30%
GENERAL COMMENTS	Manuscript improved since last version. My suggestion is to reduce further the Results Section in order to improve readability. Please, check the included new references for spelling of Authors

VERSION 2 – AUTHOR RESPONSE

Dr McAdam's Comments

While the "content" necessary to establish the background and rationale for the study is in the background section, it continues to be lack focus and flow. Rearranging text would help set up the research question better (i.e., burden of illness, benefit of good glycemic control, role of adherence, current treatment options, specific role of basal insulins, benefits of basal insults per the literature, and hypothesis about glargine vs. NPH in regards to adherence and outcomes. While mentioned in the methods, it needs to be explained why only NPH and glargine are included (i.e., too few patients on other basal insulins, namely detemir.)

We agree with your comments and have re-ordered the introduction to follow the suggested flow. The reason for excluding detemir patients has been further highlighted in both the introduction and the methods section.

Also, the use of terms Selection Bias and Confounding by Indication is still not sufficiently clear. These are two distinct issues/sources of bias in observational studies. Selection bias is defined as " a distorted estimate of the effect that results from the way in which subjects are ascertained or selected for the study population and includes factors such as differential surveillance, diagnosis, and referral of persons into the study". PSM serves to balance differences between two groups included in the study (which can lead to other bias, such as confounding by indication) and not address issues between those who were and who were not included in the study. Please see: Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. Am J Epidemiol. 1999 Jun 1;149(11):981-3. PubMed PMID: 10355372.

We have clarified this issue making it clear that these are distinct forms of bias and pointed this out in the article summary as well as the main body of text.

Dr Esposito's Comments:

The results section still appears too long. I suggest reduction by say 30%

Manuscript improved since last version. My suggestion is to reduce further the Results Section in order to improve readability.

We have substantially reduced the length of the results section.

Please, check the included new references for spelling of Authors

Checked and corrected.