Response from Bernd Richter, written in his personal capacity, to the ...

Essential Medicines List (EML) 2017
Application for the inclusion of long acting insulin analogues – glargine and detemir insulins – in the WHO Model List of Essential Medicines, as treatments used for basal insulin support in diabetes type 1.

I wish to record my objection to the proposed inclusion of long-acting insulin analogues for type 1 diabetes mellitus in the WHO’s Model EML 2017 (1). In my opinion the applicants have not provided a convincing case for the inclusion of long-acting insulin analogues – glargine and detemir insulin – in the WHO Model List of Essential Medicines, as treatments used for basal insulin support in type 1 diabetes mellitus. I base my objections on the arguments as specified below:

(A) General comments
As stated by the applicants they used their own published systematic review and meta-analysis to summarise the evidence on the long-acting insulin analogues glargine and detemir compared with intermediate acting insulin (neutral protamine Hagedorn (NPH), lente) (2) which was previously published as a protocol (3). Taking into account that this systematic review executed its literature search on January 8, 2013 one has to first consider that the evidence base for the current application is approximately four years old. Consequently, the applicants have embarked on another project aiming to improve the evidence base by means of an individual patient data meta-analysis and have published an associated protocol (4).
Overall, the benefits of long-acting insulin analogues as suggested by the applicants, even if theoretically taken as proven, appear marginal. No patient-important outcome measure (e.g. microvascular/macrovacular complications, all-cause mortality, cancer, health-related quality of life, serious/severe hypoglycaemic events) showed a substantial benefit of long-acting insulin analogue therapy. Therefore, the benefit-risk ratio for long-acting insulin analogue treatment is unknown, especially with regard to the short duration of follow-up in the included studies (16 to 26 weeks).
(B) Major clinical comments

The applicants did not report the type of insulin therapy (e.g. whether intensified insulin therapy was used in some study arms) and its effects on the results of their meta-analyses, especially what kind of combination (if any) of prandial insulin and basal insulin or continuous subcutaneous insulin infusion (CSII) was applied in the various arms of the included studies. Because short-acting insulin also has effects on outcome measures such as HbA1c, body weight and hypoglycaemic events it is virtually impossible to firmly associate the use of long-acting insulin analogues in this systematic review with any of the outcome measures investigated. Moreover, the systematic review does not inform us whether some studies used continuous glucose measurements in some of their intervention arms, again with possible effects on metabolic control and hypoglycaemic episodes. With regard to insulin therapy as such one would also need information about whether some included studies evaluated patients who had participated in a patient education and management programme, also with well-known effects on both metabolic control and hypoglycaemia (5-8).

(B) Major methodological comments

General:

For transparency reasons it would have been very helpful in the publication to report which direct comparisons contributed to the estimation of the network meta-analysis (NMA) treatment effects (i.e. the establishment of a contributions matrix to quantify how much information each direct comparison contributed to the estimation of the NMA treatment effect) - (9). Moreover, it would have been helpful to provide a clear evaluation of study limitations for each network estimate for pairwise comparisons and each outcome (instead of just aggregates of Cochrane’s risk of bias results (appendix 2 of the publication) or just some risk of bias domains for HbA1c, body weight and severe hypoglycaemia - (appendix 9, appendix 18, appendix 27 in the publication). A major shortcoming is the way review authors handled risk of bias assessments: almost all included studies employed an open-label design which review authors analysed as low risk of bias (Appendix 2 of the publication) across all outcomes where in fact an evaluation on study level and endpoint level would have been appropriate. With regard to study personnel and especially patients a correct classification would have revealed a high risk of bias, especially for body weight and hypoglycaemic episodes. Appendix 18 (risk of bias for body weight) and appendix 27 of the publication (risk of bias for
hypoglycaemia) demonstrate that review authors incorrectly associated risk of bias for (non)blinding of participants and personnel with low risk of bias in their network geometry. The same probably holds true for blinding of outcome assessment which was analysed in 96% as low risk (appendix 2 in the publication). A thorough analysis of risk of bias per outcome (assessment) would probably have revealed that there was a high risk of bias especially for the endpoints body weight and hypoglycaemia (less so for probably automated assessments of HbA1c). Furtheron, attrition bias because of missing data would need to be addressed per outcome as well (currently specified as low in 63%, appendix 2 in the publication).

Of note, review authors in the EML 2017 application associated HbA1c, body weight and hypoglycaemia with severe risk of bias, but only with regard to unclear risk of allocation concealment and/or selective reporting (Table 1, 3 and 4 in the GRADE tables). Also, though funding bias is mentioned in the EML 2017 application list, an analysis of its effects on results on main outcomes is neither mentioned in the publication nor in the EML 2017 application.

For treatment ranking an integration of risk of bias assessments from each direct comparison to formulate a single overall confidence rating for treatment rankings would have increased confidence in the findings, as well as publication of ranking probabilities (e.g. by means of rankograms) and their overlap to assess the precision of treatment rankings for specific pairwise effects.

Finally, the GRADE approach used in the EML 2017 application could have been improved by employing the published GRADE method for NMA (10).

Specific comments:

**Reductions in HbA1c**

Review authors correctly state that none of the comparisons remained statistically significant if NMA took predictive intervals into account (figure 2 in the publication) which is an adequate way to address heterogeneity. Therefore, there were no relevant differences for HbA1c between intervention and comparator groups. Also, review authors considered NPH twice daily as the “reference standard” but many analyses specified 'NPH[od/bid]', i.e. NPH once or twice daily, which clinically would be assumed to have a different effect compared with NPH twice daily on both metabolic control and hypoglycaemia. Unfortunately, comparisons with the reference standard were not provided in the publication. Moreover, for comparisons with NPH [od] it is difficult to know how many studies contributed data to the analysed comparisons.
**Body weight**

Again, comparisons of different insulin dosage schemes are not helpful, such as Detemir [od] vs NPH [od/bid]. Moreover the clinical message of “detemir once or twice daily resulted in statistically significantly less weight gain than NPH once or twice daily” and “patients experienced statistically significantly more weight gain with NPH once daily and detemir once daily when compared with NPH once or twice daily” (both statements from the EML 2017 application) is of unclear clinical relevance. Also, inclusion of predictive intervals, as stated in the publication, revealed that these effects of detemir once/twice daily versus NPH once/twice daily and detemir once daily versus NPH once daily did not remain statistically significant.

**Severe hypoglycaemia**

Review authors did not investigate the influence of the various definitions of severe-serious hypoglycaemia on results which are prone to bias (e.g. hypoglycaemia leading to a coma or fulfilling criteria for serious adverse events is less susceptible to risk of bias than “patients unable to treat themselves” or “help from third party”). Again, the effects of various insulin dosage schemes were not investigated with regard to hypoglycaemic episodes. In the NMA inconsistency was observed for severe hypoglycaemia. When prediction intervals were used in the NMA, effects did not remain statistically significant. Finally, review authors did not report data on serious adverse events (apart from hypoglycaemia).

Prof. Bernd Richter, MD, PhD  
Coordinating Editor  
Cochrane Metabolic and Endocrine Disorders Review Group  
Institute of General Practice  
Medical Faculty of the Heinrich-Heine University Duesseldorf  
PO Box 101007, 40001 Duesseldorf, Germany  
Office address: Werdener Str. 4, 40227 Duesseldorf, Germany  
Phone +49 211 811 8773 | Fax +49 211 81 015 18773  
richterb@uni-duesseldorf.de  
www.endoc.cochrane.org
References


