Statistical Review and Evaluation, December 1, 2011 - MenHibrix

BLA Number:	125363.0.21
Product Name:	MenHibrix (Meningococcal Groups C and Y and Haemophilus b Tetanus Toxoid Conjugate Vaccine)
Applicant:	GlaxoSmithKline Biologicals SA
Date Submitted:	December 1, 2011
Action Due Date:	June 1, 2012
Statistical Reviewer:	Tsai-Lien Lin Lead Mathematical Statistician Viral and Bioassay Team, VEB/DB/OBE
Through:	A. Dale Horne Branch Chief, VEB /DB/OBE
To:	Joseph Temenak David Staten Kirk Prutzman
CC:	Review Team Estelle Russek-Cohen Henry Hsu Christopher Egelebo DB chron

1. EXECUTIVE SUMMARY

GSK submitted BLA 125363 Amendment 21 to provide responses to the remaining 26 items listed in the second CR letter issued on 9/21/2010. This statistical review covers items 1-3 regarding the meningitides serotype Y (MenY) hSBA assay stability , and items 4 and 16 regarding the CMC issues.

1.1 Serology Issues: Items 1-3

The applicant provided additional Deming Regression and Bland & Altman analyses on all retested samples from all clinical studies tested either for assay maintenance, reagent qualification, or as part of QC panels. The QC charts and various trending plots requested in the second CR are also submitted.

Conclusions and Recommendations:

The –(b)(4)--- of MenY titers for the re-tested samples from study Hib-MenCY-TT-005 remains unexplained. However, this issue may be less critical to the study results of pivotal studies -009 and -010. The more relevant assay stability data are the data for the samples from study -013 added as sentinel samples in routine hSBA testing of samples from pivotal studies -009 and -010. The retest values of these -013 sentinel samples appear to be lower than the original test results after ----(b)(4)-------, which was supposed to adjust for the change of human complement lot that occurred prior to this testing period. However, no further drift is observed during the 4 weeks of testing of -009 and --010 samples. Because the laboratory analysts are blind to the treatment

assignment and the subjects are randomized to treatment groups, it may be reasonable to assume that any change in the assay that occurred over time would affect all treatment groups equally. The statistical reviewer defers the decision on the acceptability of serology assay results to the product reviewers.

1.2 CMC Issues: Items 4 and 16

For item 4, the applicant proposed to remove the previously submitted Comparability Protocol (CP) and instead use a modified CP for ---(b)(4)----- for determination of the polysaccharide content in Hib vaccine. In addition, to address the concerns in items 4a and 4b, the applicant proposed validity criteria for verification of the calibration factor. For item 16, a statistical approach for setting the product specifications was proposed, which takes the process variability and process performance indices into account.

Conclusions and Recommendations:

Item 4: The comparability protocol (CP) for the change of reference material for the Hib identity (b)(4) on conjugated (b)(4) (PS-TT) and Final Container HibMenCY is acceptable. With the use of an absolute quantitative method, drift during replacing standards is no longer an issue. The proposed validity criteria for verification of the calibration factor specified in this CP are thus not needed. However, these verification criteria should not be applied to other future CPs without careful evaluation of each individual situation.

Item 16: The proposed method for setting product specifications is generally acceptable. However, when used by the applicant to set specifications for –(b)(4)---, because of the large amount of <LLOQ values and the assigned value (=LLOQ) for these <LLOQ values, the proposed method yields upper limits that may be unacceptably high. This issue is not present in setting the specifications for purified TT (items 6, 11a, and 11b) and –(b)(4)--- (item 15). The statistical reviewer defers to the product reviewers on whether to make the setting of --(b)(4)-- specifications a post-marketing commitment.

2. Background

GSK submitted the original BLA 125363 on 8/12/2009 to seek licensure of a combination vaccine Hib-MenCY-TT intended for the vaccination of US infants at 2, 4, and 6 months of age, with a fourth dose to be administered at 12 to 15 months of age. The first CR letter issued on 6/11/2010 contains 88 items. On 9/21/2010, CBER issued the second CR letter containing 26 items. Amendment 21 was submitted to provide responses to the remaining 26 items listed in the second CR letter. This statistical review covers items 1 to 3 regarding the meningitides serotype Y (MenY) hSBA assay stability. In addition, statistical issues involved in the CMC items 4 and 16 are reviewed at the product reviewers' request.

3. Statistical Evaluation

3.1 Serology Issues

Item 1

In response to Item 1a, you indicated that study 005 sera were not handled according to the SOP. You indicated that the study 005 sera ----------

----------(b)(4)----------. Thus, these data do not support your hypothesis that the observed –b(4)----- in Men Y titers is attributable to ----(b)(4)----------.
In addition, you have also suggested that the age of the study 005 sera may have played a role in the ----------(b)(4)----------. We conclude that the reasons for the –(b)(4)--- in study 005 hSBA titers remain unknown. Please provide any additional information you may have that would explain the –(b)(4)--- in study 005 hSBA results.
Applicant's Responses :
Additional data and analysis are presented, including:

- **Bridging analysis –**

After release of clinical results, samples from clinical trials are repeatedly tested either for assay maintenance, reagent qualification, or as part of QC panels (sentinel samples). Deming regressions and Bland & Altman plots of all retested samples from all clinical studies (004, 005, 006, 007, 008, 009, 010, 013, and 014) were performed.

Except for study 005, the data from the analyses display no consistent trends. The results of Deming regression analyses show that all r values are >0.8; accuracy, CCC, and slope of the Deming regressions close to 1; and y-intercepts close to 0. For the Bland & Altman analyses, y-intercepts are close to 1 and slopes close to 0. In addition, geometric mean ratios between the original and retest values are close to 1, and the % concordance values are all above 94%. Though significant McNemar tests are noted for studies 010 and 013, their relevance is limited in view of the high concordance. Overall, these results indicate good agreement between the original and retest values.
For study 005, GSK acknowledges that the lower results from the retesting remain unexplained. The applicant has previously evaluated several factors that were investigated as potential causes of the -005 retesting results, including: number of ---(b)(4)------ cycles, MenY ----(b)(4)------ stability and change in critical reagents (-----(b)(4)-------- lots). None of these factors could explain the lower retest results. GSK believes that the observation of lower results with retesting 005 samples is an isolated finding, and since 005 retesting samples were

not handled according to the normal procedure, these retest data should not be considered representative of the long-term performance of the assay.

- **Trending plots –**

  The trending analyses were performed on the same datasets used for Deming Regression and Bland & Altman analyses. The ratios of retest values relative to original values are plotted over time (from August 2006 to August 2011) for each MenCY-TT study (002, 004, 006, 007, 008, 009, 010, 013, and 014) and MenACWY-YY study (012, 013, 028, 052, 055, 059, 062, 071) and for all these studies pooled.
  A total of 1826 samples were retested with the majority of ratios clustering around 1. Ninety four percent (94%) of the samples fall within the range of 0.25-4.0 (+/- 2 dilutions) which is considered an acceptable range by the applicant. During the period covering the testing of Phase 3 studies (Jan-March 2009), the average geometric mean ratio (GMR) was less than 1, but still within the range (0.25-4.0). Therefore, the applicant concluded that the trending plots showed no significant drift from August 2006 to August 2011.

- **QC charts with change of critical reagents –**

  The QC charts indicating reagent changes that occurred during all clinical testing (December 2005-March 2009) are provided. Prior to August, 2006 QC charts were not normalized. The QC charts showed that the variability of the assay was reduced with the use of the ----(b)(4)-------.

Reviewer's Comments:

- *The lower retest results observed for studies 007/008 previously submitted were no longer evident in the Deming regression analyses presented in this amendment, which include many more retested samples. For study 005 no explanations can be found for the –b(4)------ retest titer values, except that the 005 retesting did not follow the normal operating procedure.*
- *The applicant defined the acceptable range of re-test/initial titer ratio as--(b)(4)-- (+/- (b)(4) dilutions). This is a very wide range, but in the responses to item 2b.i, the applicant justified this acceptable range based on the assay intermediate precision (%CV=(b)(4),) assuming the re-test and initial titers are independent. The trending plots and the QC charts exhibit large variability of the assay.*

- *During the time of phase 3 pivotal studies testing, the geometric mean ratio is lower, though still within the wide acceptable range defined by the applicant. Several individual re-test/initial ratios are below (b)(4) and some are even lower than(b)(4).*

Item 2

2a. In response to Item 3a, you provided data relevant to the reliability of the hSBA for Men Y. Specifically, you presented a table of hSBA values from the Y assay for the sentinel samples included in study HIB-MENCY-TT-013 (Table 5). We note that seven out of 50 samples show a greater than four-fold discrepancy between the highest and lowest reported values. Four samples show results both above and below a titer of (b)(4), including one sample with a titer in the -----(b)(4)----------. The samples with only one replicate provided are not included in the totals. In addition, a substantial amount of data is missing from the table which precludes a complete assessment of assay stability. Together, the---(b)(4)---- titers seen in the repeat analyses for samples from Study Hib-MenCY-005 (refer to item 1 above), in conjunction with the ---(b)(4)---- titers and the discrepancies in the data submitted in response to Items 1b and 3a (refer to item 2 above), raise concerns with regard to the ability of the hSBA assay for the Y strain to produce reliable and consistent data over time. While it is acknowledged that sample storage may have been one factor leading to ---(b)(4)------ hSBA titers in the Men Y retest, adequate control of the assay during the sample analysis of the pivotal studies is critical. In this regard, we request the following additional information:

a. To evaluate whether small changes in the assay over time would have affected all groups from a given study equally, please provide the blinding and randomization scheme for analysis of the samples from the pivotal studies.

Applicant's Responses:

Samples from studies 009 and 010 were tested in parallel by a team of (b)(4) analysts from ----------(b)(4)----------. The analysts were blinded to the treatment groups. In the initial testing of the samples, each assay run mostly consisted of samples from a single study and samples were assigned to assay runs according to receipt date order. On repeat testing of samples for which valid titers could not be obtained during the initial runs, sera from multiple time points and both studies were more frequently analyzed in the same assay run. Because the treatment groups are randomized and because the laboratory personnel are blinded to this assignment, any change in the assay over time would affect all treatment groups equally.

Reviewer's Comments:

The applicant's response appears to be acceptable.

2b. i. Please provide data that demonstrate that the ----(b)(4)-------- algorithm maintains consistent assay performance across changes in control and complement lots. Please provide a trending analysis for the ----(b)(4)-------- values that demonstrates consistent assay performance within control and complement lots. Please show that the ----(b)(4)-------- algorithm is independent of sample titer, i.e., that the variance of the ----(b)(4)------- ratio is constant relative to titer. (CBER Oct-13-11 email correspondence)

Applicant's Responses:

QC charts of two controls contained in each assay -(b)(4)-- and trending plots of the **----(b)(4)--------** factor and ratio of retesting versus historical results are provided. These QC charts and trending plots showed that within a single control or hC' lot, no apparent drift can be observed. The **----(b)(4)--------** factors vary over time, but all fall within the range ---(b)(4)---- considered acceptable by the applicant.
Impact of the **----(b)(4)-----** factor on assay precision was examined by plotting the %CV of retested samples over different ranges of titers based on all repeat data, with or without –b(4)--- (Figures 5 and 6 on page 26 of m1.11.3 in the submission) between 2006 and 2009. Precision appears to improve slightly with **----(b)(4)--------** for titers (b)(4), but worsen slightly for titers (b)(4). Overall, the impact of the **---(b)(4)-----** factor on the assay precision is not much.
Impact of the **----(b)(4)--------** factor on clinical studies results is also minimal. **----(b)(4)-----** increases the GMTs, but has little impact on the % responders. The criterion for % responders is met for both **----(b)(4)-------- ----------** .
Reviewer's Comments:
Although the conclusions of clinical studies used to support this BLA remain the same with or without **----(b)(4)--------** *, whether* **----(b)(4)--------** *is justified needs to be further investigated if the assay is to be used in future studies. The acceptable range for the* **----(b)(4)----** *factor defined by the company--(b)(4)-- is quite wide. If the qualification of a new lot of critical reagent requires the comparability criteria (the geometric mean ratio of positive samples must be within --(b)(4)-- and the agreement between results with the two reagents must be --(b)(4)- to be satisfied, it seems not reasonable for the* **----(b)(4)-----** *factor to fluctuate so much.*
2b.ii. Please present the analysis that demonstrates that the four-parameter model can be appropriately fitted to the bacterial count data generated in the assay. Please describe how the a and d parameters for each sample are determined and controlled. Please comment on whether the curve fitting is constrained, and if so, please explain how it is constrained. Please provide the basis for the criterion that each sample has an R2 greater than (b)(4).
Applicant's Responses:
GSK re-analyzed the **----------(b)(4)----------** titer data of pivotal studies -009 and -010 and the results were compared to the results based on **----(b)(4)----------** titers. Although the difference in GMT is about 3 fold (contributed by both **----(b)(4)-------** and **---(b)(4)--------**), there is very little difference in % responders.
Deming regression analyses were performed on the **----(b)(4)----------** titers versus **----(b)(4)------** titers for studies -009 and -010. The agreement is quite good, with slopes of approximately 1.0 and a shift due to the fact that the **----(b)(4)--------** method always underestimates the titer associated with –b(4)------
The a and d parameters are not constrained. The applicant cites a paper by Wang (2008) to support that there is no need for constraining a and d for the estimation of titer at –b(4)------- The criterion of –b(4)--, one of the system suitability criteria, was previously communicated to CBER during the IND stage. About 0.33% standard curves in study 009 and 0.29% standard curves in study 010 were rejected due to R 2 values.
Reviewer's Comments:
Theoretically, the **----(b)(4)----------** *method should generate titer estimates with better precision and less bias, if the (b)(4) model fits well. The criterion of –b(4)-- may appear*

a little loose. It is not known whether the R 2 criterion is stringent enough to ensure adequate model fit. Nevertheless, the **----(b)(4)--------** *does not affect the conclusions drawn from the pivotal studies.*

2b.iii. You presented quality control charts for positive controls in the Men Y hSBA assay (QC1 and QC2) for the testing period from July 2009 to June 2010 to demonstrate assay stability. We notice that in the QC chart for Control 1 ((b)(4)/C1/04/08) for the period from July 2009 to January 2010 (Section 4.3.8, Figure 2, page 30), many data points are below the lower limit. For the period from February 2010 to June 2010 (Section 4.3.8, Figure 3, page 30), the target value for Control 1 ((b)(4)--/C1/05/09) is changed to a higher level. Although all data points are within the control limits, the range between the lower and upper control limits becomes much wider. In light of these observations, please explain why you conclude that the hSBA Men Y assay is stable.

Applicant's Responses:

The applicant explained that the control limits around control -(b)(4)- /C1/04/08 were defined on the basis of 81 tests resulting in 124 values. The variability for the control during this time period (July 2009 to June 2010) is lower (%CV=15%) than is normally seen for this assay (%CV=45%). With this tighter range, the number of samples outside the limit was 11% of the total number of samples.

The range for control -(b)(4)-/C1/05/09 was calculated initially on 41 values with a %CV of 37%, though much higher than the variability for control -(b)(4)-/C1/04/08, was more in line with the normal assay variability. The range has been re-calculated on a greater number of values (124). The re-calculated %CV is 28%.

Reviewer's Comments:

The applicant's response is acceptable, although it may be a little unusual to observe a %CV much smaller than normal with a sample size of 124.

2b.iv. The time period covered by these QC charts (July 2009 to June 2010) began several months after the testing of samples from studies 009 and 010 (Jan 2009 to February 2009) was completed. Thus, these QC charts do not provide information regarding the assay stability at the time the testing of samples from the clinical studies supporting this BLA was performed. Please provide data that support the stability of the assay covering the actual testing period from study -005 to study -010. Data that would be supportive include all QC charts form controls with trending analyses, reagent qualification data for any new controls or complement introduced during the analysis of samples from a given study, and all sentinel data. A detailed and continuous time line depicting the changes in controls and complement lots during the entire testing period should also be included.

Applicant's Responses:

The requested information including QC charts and trending plots for the time frame between -005 and -010 is provided by the applicant. Each change of reagent is indicated in the charts and plots.

Reviewer's Comments:

The applicant's response is acceptable.

Item 3

We are concerned that missing data for the samples from Study -013 added as sentinel samples in routine hSBA testing of samples from Studies -009 and -010

may have biased the results of the Men Y assay stability evaluation, especially for week 1.
Out of the 38 samples tested in week 1, only 20 samples have valid titer results. Eight samples have a missing value code "TC", meaning that they were supposed to be retested at the lower dilution because less than 2 dilution points of the curve have ≥ 50% killing. Since these missing TCs are not missing at random, excluding these samples could make the GMR at week 1, relative to the initial reference, higher than the true ratio had those TC samples been re-tested (based on their titers at weeks 2-
4). Overall, the GMRs during the four weeks clearly suggest that a –b(4)------ in MenY titers from the initial reference values is also present for these sentinel samples from study -013. Also, the concordance analysis may not be useful for evaluating this unidirectional (decrease) assay stability issue and its potential impact on the clinical studies results, because there are many samples with titers (b)(4) initially and few samples near the cutoff point. Please comment.
Applicant's Responses:
GSK explained that in order to conserve sample volume, and in line with the purpose of the sentinel samples, no retests for those 8 "TC" samples were performed during week 1. If those 8 samples were replaced with their week 2 results, the GMR at week 1 would be approximately (b)(4), which is in alignment with the GMR for weeks 2 (b)(4) , week 3 (b)(4) and week 4 (b)(4) ; all significantly lower than (b)(4) . The GMR from May-June 2011 re-testing of study 013 samples is higher than the GMR from the phase 3 testing (Jan-Feb 2009). Because GMR among weeks 1-4 ranges from - (b)(4)-- , the applicant concluded that the hSBA-MenY assay was stable across the testing of the phase 3 samples.
Regarding the concordance analyses, GSK acknowledges that the concordance analyses should be supplemented with additional analyses, i.e., the results of concordance analyses have limited utilities.
The applicant attributes the observed lower re-testing results for the -013 sentinel samples to the change of b(4)'lot during the -009 and -010 testing and concluded that the lower re-testing results does not indicate a drift in the assay but rather the normal variability of the assay when different b(4)lots having differing levels of activity are used.
Reviewer's Comments:
The applicant's responses actually confirm that the retest titer values for the -013 sentinel samples during the -009 and -010 testing period are lower than the original values, but no further –b(4)----- is observed during this time period. GSK explains the -b(4)--- in titers by the hC' lot change which occurred between the initial testing and re-testing of these sentinel samples and the higher activity level of the b(4) lot used in the initial testing. The reviewer does not think this explanation is convincing, because the -----(b)(4)--------- factors should have adjusted for the difference caused by the change of this critical reagent. The-----(b)(4)--------- factors during the -009/-010 testing period are lower, i.e., the titers during this period of time are adjusted higher, yet a –b(4)---- in the re-test values for the sentinel samples is observed.
Conclusions: Serology – Items 1, 2, and 3
The –b(4)---of MenY titers for the re-tested samples from study Hib-MenCY-TT-005 remains unexplained. However, this issue may be less critical to the study results of

pivotal studies -009 and -010. The more relevant assay stability data are the data for the samples from study -013 added as sentinel samples in routine hSBA testing of samples from pivotal studies -009 and -010. The retest values of these -013 sentinel samples appear to be lower than the original test results after -----*(b)(4)*----- , which was supposed to adjust for the change of human complement lot that occurred prior to this testing period. However, no further drift is observed during the 4 weeks of testing of -009 and -010 samples. Because the laboratory analysts are blind to the treatment assignment and the subjects are randomized to treatment groups, it may be reasonable to assume that any change in the assay that occurred over time would affect all treatment groups equally. The statistical reviewer defers the decision on the acceptability of serology assay results to the product reviewers.

3.2. CMC Issues

Item 4

The Comparability Protocols (CPs) provided in response to Item 82 for changes in reference standards are inadequate for the purposes of reporting such changes in your annual Report. Please address the following deficiencies in the CPs:

a. Most of the CPs has acceptance criteria of less than 10% differences in results generated with new and old reference standards. However, the CPS for the Free TT Content and Identity assays contain qualification criteria stating that comparability between the old and new standard is demonstrated if the results are ± (b)(4). Such a large variability in the calibration of new standards is not acceptable. Please revise the criteria for calibrating new standards for the Free TT Content and Identity assays.

b. Even 10% differences between new and old reference standards can cause problems, particularly when qualifying a new reference standard against the current standard multiple times over the life of the product. A new standard should be calibrated against the original or primary standard to avoid drift away from original value. Please develop a primary reference standard for each assay to avoid drift in calibration of reference standards over the life of the product.

Applicant's Responses:

The applicant proposes to remove the CPs provided in response to item 82. Instead, CP 9000006115CPR001 ---(b)(4)------- for determination of the polysaccharide content in Hib vaccine by (b)(4) has been updated in order to address concerns highlighted in item 4a, b, c, d, and e. In addition, a validation report 9000001118RVR006 is submitted, providing data related to the change in reference material for the "Hib identity by (b)(4) on conjugated (b)(4) (PS-TT) and Final Container HibMenCY".

To address items 4a and 4b, the applicant proposes the followings:

1. The titer of the candidate reference standard will be determined using the current reference standard on a minimum of 30 values. A calibration factor will be calculated as the ratio between the titer of the candidate and the current reference standard. With such a sample size, the expected precision about the mean change that is addressed through the calibration factor will be approximately half of the variability of the method.
2. This calibration factor will be verified through a minimum of 12 determinations of a validity criterion. This validity criterion will be method-specific and will be described

in each comparability protocol.
The verification of the calibration factor will be positive providing that:

- The difference between the mean value of the validity criterion obtained using the candidate reference standard and the current reference standard is within one standard deviation of the a
- AND the sum of the successive differences between the mean values of the validity criterion calculated in all previous bridges is within one standard deviation of the assay.

In the validation report 9000001118RVR006, batch ------(b)(4)----- was tested using 31 independent (b)(4) assays using the current standard ------(b)(4)----- (with titer (b)(4) µg/dose). One value was identified as an outlier by Grubb's test and was excluded from the analysis. The mean antigen content of ------(b)(4)---- based on 30 independent assays is ------(b)(4)----. Thus, the calibration factor for the new standard is (b)(4).
The calibration factor for the new standard is verified by examining the criterion that the difference between the mean values of the cut-offs (determined as OD at 20% of the standard curves) used for identity determination using the current and the new standard is within one standard deviation of the assay. Based on 30 historical data obtained from the current standard, this standard deviation was calculated as (b)(4) . The verification was done in (b)(4) independent sessions. In each session, the new standard was tested --(b)(4)- giving a total of (b)(4) determinations for the new standard. The difference in mean cut-offs between the new vs current standard is (b)(4) and is within one standard deviation of the assay (b)(4) .

Reviewer's Comments:

- *The calculation of the calibration factor for the new standard ----(b)(4)--- appears to be acceptable.*
- *It is not clear what is the statistical rationale for the method and criteria for the verification of the calibration factor. The statistical properties of the sum of the successive differences between the mean values of the validity criterion calculated in all previous bridges are not clear. Nevertheless, according to the OCBQ product reviewer, verification of the calibration factor is really not needed here due to the fact that the new standard is now calibrated by an absolute quantitation method and the comparability with the old standard will be assessed by (b)(4), such that drift during replacing standards is no longer an issue. Therefore, the comparability protocol submitted in this amendment is acceptable in this case. Any future CPs submitted post approval will need to be carefully reviewed in their individual context.*

Item 16

We note that you calculated the ---(b)(4)--- specification based on pooled data from *Neisseria meningitidis* polysaccharides (A, C, W, and Y). However, the calculation of ---(b)(4)--- specifications should be serotype specific. In addition, the ---(b)(4)--- specification for Drug Product should be process capability driven and should reflect actual process data. Please re-calculate your ---(b)(4)--- specification to be reflective of actual process data for each serotype individually.

Applicant's Responses:

The applicant proposes to use a statistical approach to calculate the **---(b)(4)---** specifications based on *process performance indices* . The basic idea used by the applicant for setting specifications is:

----------(b)(4)----------

Assuming normal distribution, this control levels will cover 99.73% of the values generated by the process. GSK further applies a lower/upper process performance index (Ppl and Ppu) of (b)(4) to the equation above, the lower and upper specifications are then calculated as follows:

----------(b)(4)----------

----------(b)(4)----------

With an index of (b)(4), the expected number of out of specifications results is 318 out of 1 million (two-sided range of specification), provided an estimate of long term variability of the process. The applicant calculated the upper specifications for the MenC purified PS bulk, MenY purified PS bulk, and HibMenCY final product. When the estimated process SD is unreasonably small (due to large amount of <LLOQ values), the assay analytical SD, determined during the method validation, is used instead. Otherwise, the process SD is used in the calculation. For calculations of mean and process SD, the <LLOQ values are set at LLOQ.

The company plans to apply another level of control, an alert level, to generate an out of consistency (OOC) warning when it is reached. The calculated upper specifications, the proposed specifications, and the proposed alert level are:

	Calculated upper specification	**Proposed upper specification**	**Proposed upper alert level**
MenC purified PS (b)(4)	-------(b)(4)-----	-------(b)(4)----	-------(b)(4)-----
MenY purified PS (b)(4)	-------(b)(4)-----	-------(b)(4)-----	-------(b)(4)-----
HibMenCY final product	-------(b)(4)---	-------(b)(4)-----	-------(b)(4)-----

Reviewer's Comments:

- *On page 2 of m1.11.1 Quality Information Amendment –CRL September 2011, CMC Question 16, the equations for calculating the lower and upper specifications are listed as:*

  ----------(b)(4)----------

  ----------(b)(4)----------
- *Since ----(b)(4)-----, the equations in the submission would have the process performance index multiplied twice. These could be just typos, as the calculated specifications are based on the correct equations, i.e., Mean –(b)(4)- (verified by the reviewer.)*
- *The proposed method for setting specifications considers only the process capability (i.e., protects/minimizes the manufacturer's risk). The risk of falsely accepting a bad lot (consumer's risk) is not controlled. However, mean –(b)(4)-- is currently a widely used approach. The applicant further widens the acceptable range by a process performance index to accommodate the long term variability. Unless there are other clinical safety/efficacy data available, it is difficult to formally take the consumer's risk into consideration for setting product specifications.*

- *Though the method of setting specifications is generally acceptable, when applied to the calculations of the ---(b)(4)-- specifications, there is a problem caused by the large amount of <LLOQ values in the data and how <LLOQ values are treated in the calculations of mean and standard deviation. Setting <LLOQ values at LLOQ makes the estimated mean shift to a higher level and thus makes the upper specification higher. For example, for MenC purified PS (b)(4), only one of the 103 lots included in the analysis has a value (22.2) above the LLOQ of (b)(4) IU/mg PS. Based on the estimated mean of 20 and the analytical SD of 10.01, the calculated upper specification is –b(4)-, which is considerably higher than the data range and does not appear reasonable. When the majority of the data is below LLOQ, as displayed in the endotoxin data, this method of setting specification may not be appropriate.*
- *The same method for setting specifications is used for several other quality attributes, e.g., purified TT (items 6, 11a, 11b) and ---(b)(4)--- of MenHibrix (item 15). There are no <LLOQ values for those quality attributes; therefore, this particular problem is not an issue for those situations.*
- *For the final product, based on the calculated upper specification of (b)(4) IU/dose, the applicant proposes an upper specification of (b)(4) IU/dose. This degree of inflation is not justified.*

Conclusions: CMC – Items 4 and 16

The comparability protocol (CP) for the change of reference material for the Hib identity (b)(4) on conjugated (b)(4) (PS-TT) and Final Container HibMenCY is acceptable. With the use of an absolute quantitative method, drift during replacing standards is no longer an issue. The proposed validity criteria for verification of the calibration factor specified in this CP are thus not needed. However, these verification criteria should not be applied to other future CPs without careful evaluation of each individual situation.

The proposed method for setting product specifications is generally acceptable. However, when used by the applicant to set specifications for –(b)(4)----, because of the large amount of <LLOQ values and the assigned value (=LLOQ) for these <LLOQ values, the proposed method yields upper limits that may be unacceptably high. This issue is not present in setting the specifications for purified TT (items 6, 11a, and 11b) and ---(b)(4)------- (item 15). The statistical reviewer defers to the product reviewers on whether to make the setting of –(b)(4)---- specifications a post-marketing commitment.