Taken from:
http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm341718.htm#s4

4.1.1 *In vivo* testing for SpO_2 accuracy under laboratory conditions

We recommend you follow Clause 201.12.1.101.2 and Annex EE.2 of ISO 80601-2-61:2011 *Procedure for invasive laboratory testing on healthy volunteers*, or equivalent method to validate the SpO_2 accuracy specifications of your pulse oximeter system by comparing each value from your system and a simultaneous value from co-oximetry of an arterial blood sample. We recommend you submit a detailed clinical report for this testing. Your report should describe the device configuration tested and include the following:

- test apparatus used, including means for arterial catheterization and blood sampling, means for recording SpO_2 values, and means for delivering medical grade oxygen-nitrogen mixtures of varying fractional inspired oxygen (FiO_2) levels;
- inclusion and exclusion criteria;
- number of subjects;
- number of samples taken per subject;
- specific conditions of testing, including laboratory conditions, subject motion, low pulse amplitude;
- type and frequency of motion for testing under motion conditions, if applicable;
- criteria and methods for determining stability of reference arterial blood oxygen saturation (SaO_2) at the pulse oximeter sensor site;
- desaturation profile, including target saturation plateaus and ranges; and
- formula used for determination of root mean square difference (A_{rms}). (See Clause 201.12.1.101.2.2 of ISO 80601-2-61:2011 for recommended formula.)

We recommend you conduct the study described in Clause 201.12.1.101.2 and Annex EE.2 of ISO 80601-2-61:2011 on 10 or more healthy subjects that vary in age and gender. Your data should include 200 or more data points (paired observations: pulse oximeter, co-oximeter). These data should be distributed as described in Annex EE.2.3.4(g).

Your study should have subjects with a range of skin pigmentations, including at least 2 darkly pigmented subjects or 15% of your subject pool, whichever is larger.

We recommend you provide a line listing, a Bland-Altman plot, error plots (i.e., SaO_2 versus (SpO_2-SaO_2) for both individual test subjects and all subjects pooled), and rationale for any points excluded from analysis. Please include population mean bias ($\mu 0$), between-subject variance ($\sigma\mu^2$), within-subject variance (σ^2), and upper 95% and lower 95% limits of agreement. Please provide this information as outlined in Section 3 of “Agreement Between Methods Of Measurement With Multiple

Observations Per Individual" by Bland and Altman.[2] If you note that the plots show noticeable outliers, please provide the following:

- a discussion of the state of health, subject characteristics, test setup, test procedure, and any other factors that may have affected these data points; and
- a discussion of how the outlier(s) do not raise safety and performance concerns regarding the accuracy of the device.

We recommend an A_{rms} specification in conformance with Clause 201.12.1.101.1 of ISO 80601-2-61:2011. We recognize that accuracy is, among other things, a function of patient characteristics, application site and sensor geometry. The table below outlines the typical A_{rms} between measured values (SpO_2) and reference values (SaO_2) under normal conditions ranging from 70% to 100% SpO_2.