

Determining the Artifact Sensitivity of Recent Pulse Oximeters during Laboratory Benchmarking.

Kästle S.W., Konecny E. *J Clin Monit Comput.* 2000;16(7):509-22.

This study aimed to investigate and compare the performance of the algorithms contained in the newest generation of pulse oximeters (Masimo SET in IVY2000, Nellcor Oxismart N-3000, Agilent M3 rev. B) against a traditional pulse oximeter (Agilent CMS rel. A.0).

The benchmark was performed in an efficient and reproducible way in the laboratory environment using patient signal recordings complemented by a two-hand volunteer motion test. A novel method of creating artifact/reference signal pairs from a clinical database, the noise-mix-composition (NMC), was developed. The new method enabled the simulation of critical clinical situations in a more realistic way than the usual two hand volunteer studies. An advantage of the laboratory tests over live clinical studies was that a continuous saturation reference was available, allowing accurate on-going determination of the SpO₂ error. A new quantitative performance measure, the non-performance index (NPI), was developed and applied to the benchmark results. It covers the 3 performance aspects of a pulse oximeter: (1) SpO₂ accuracy, (2) pulse rate accuracy and (3) drop out times. These factors were weighted according to clinical importance determined by a survey.

During the restricted conditions of steady state and forced motion test on healthy volunteers Masimo/Ivy's pulse oximeter performed best with a 2.6 fold improvement over the conventional technology. Clear improvements were also found for Agilent's M3 (1.6 fold) and Nellcor's N-3000 (1.6 fold). In contrast, the clinically oriented NMC study yielded the best performance improvement--as measured in NPI numbers--for Agilent's M3 rev. B (1.6 fold) and due to more frequent SpO₂ errors only 1.5 for Masimo and 1.3 for N-3000. A large difference was found for the dropout rate: the lowest was achieved by Masimo (3.0% of total time), the largest by Nellcor N-3000 (24.1% of total time), a factor which was rated high by clinicians. Very pronounced improvements (between 2.3 and 3.4 fold) on all of the newer devices were found for the pulse rate.

The NMC turned out to be a very useful tool for generating a standard signal set for algorithm development and benchmarking purposes that eliminates repetitive clinical testing in early stages. The applicability of its results needs confirmation by clinical live studies.