Continuous glucose monitoring (CGM) is a rapidly evolving field, and is far from maturity. According to a study published in 2015, even between two widely used products, 30% difference in a major performance measure was seen, which was quite contrary to established areas of medicine. On that background, it is necessary for academic researchers to carry out objective evaluation of CGM, because overwhelming proportion of studies was done by the device manufacturers for regulatory purposes. Moreover, results of independent investigation can be a good basis to make recommendation to a patient about the advantages and disadvantages of a specific CGM device. However, it is problematic there’s no widely accepted reference method for assessing sensor accuracy and studies tend to become obsolete at the time of publication due to frequent updates of the products by the manufacturers.

There are multiple areas of clinical researches where CGM is mainly involved. Among them, the most obvious and critical one for a new CGM device or technology would be the investigation upon the sensor accuracy. Reference methods to assess sensor accuracy have been either whole blood laboratory glucose measurement or self-monitoring blood glucose (SMBG). It is noticeable there came more recent studies on clinical application of CGM for managing diabetic patients. Through the studies, roles of CGM in adjustment of insulin dosing and its implication on clinical outcomes are being better understood. Nowadays, clinical trials of new medications and novel therapeutic strategies utilize CGM more proactively, mainly because CGM is deemed to be exceptionally useful for measuring hypoglycemic or hyperglycemic exposure in ambulatory care, and enables comprehensive assessment of glucose variability.

Sensor accuracy of CGM has been widely explored by communities of laboratory professionals. Unfortunately, there still is no universally accepted protocol to compare performance among sensors, unless they are worn at the same time. Several parameters have been used to assess sensor accuracy. Although none of them was endorsed by the professionals unanimously, mean absolute relative difference (MARD) was emerged as the preferred metric. Almost all the latest publications on sensor accuracy used MARD. MARD can be expressed as a single number, and the lower MARD means more accuracy. Despite the intuitiveness of MARD, it is impacted by study design, and its threshold of 10% has little pathophysiological rationale. Median absolute relative difference tends to be lower than MARD, because of the removal of outliers. Mean absolute difference used to be the main parameter of accuracy in hypoglycemic ranges. However it is less commonly used thanks to the improvement of sensor performance in low glucose concentrations. Sometimes, MARD is complemented by paired or precision absolute relative difference calculated by the comparison of readings generated by identical CGMs worn simultaneously by subjects. ISO 2013 guidelines are occasionally used for assessment of sensor accuracy. The guideline was originally intended for SMBG to achieve regulatory approval, but some found it useful also for characterizing system performance of CGM complementing MARD.

Continuous Glucose Clarke’s Error Grid Analysis (CG-EGA) has been recommended by Clinical and Laboratory Standards Institute (CLSI). However, due to its excessive complexity, only one published study claimed to use the procedures in CLSI POCT05-A until 8 years after the guideline came out. Compared to CG-EGA, the earlier Clarke’s Error Grid analysis is more widely used. There are other methods to assess sensor accuracy, such as proportions of values...
that lie within 20 mg/dL or within 20% from the reference method, Bland-Altman plots, and Surveillance Error Grid. They are used either alone or along with better accepted parameters such as MARD.

It is true sensor technology has been markedly improved since the first CGM was introduced 17 years ago. Nevertheless, understanding sensor accuracy is far from complete. There are remaining areas where further studies are needed, such as implication of usability and the human interface, accuracy over sensor lifetime, and interference caused by common factors like acetaminophen ingestion.