Journal of Diabetes Science and Technology
2016, Vol. 10(1) 185–190

DOI: 10.1177/1932296815597918
dst.sagepub.com
SAGE

IQWiG Reanalyzes and Raises Questions About an Article by Ly et al Which Concluded Low Glucose Suspend Is Very Beneficial

Lutz Heinemann, PhD[1] and Norbert Hermanns, PhD[2]

Abstract

In 2013, Ly et al published a study in *JAMA* reporting a massive reduction in the frequency of severe hypoglycemic events when the patients used sensor augmented insulin pump therapy with low glucose suspense. The data of this study were reanalyzed by the IQWiG when this German institute started its evaluation of the evidence for continuous glucose monitoring (CGM) usage. The IQWiG came to a contrary conclusion than the authors of the Ly study. Decisive for this was the statistical analysis of the Ly study that led the IQWiG to evaluate this result as a lack of evidence for the superiority of CGM (plus pump) for preventing hypoglycemia. In this commentary, a direct English translation of the IQWiG analysis is provided to enable the reader to come to his or her own conclusion about this study.

Keywords

CGM, hypoglycemia, low glucose suspend, type 1

The German Institute for Quality and Efficiency in Health Care (IQWiG) assesses the advantages and disadvantages of medical procedures (eg, drugs, medical devices, and surgical procedures) by conducting systematic searches through international scientific literature for studies in which the comparisons in question are described. The IQWiG uses methods of evidence-based medicine for selecting and assessing these studies. Based on this research, IQWiG provides reports to the German Common Federal Council (Gemeinsamer Bundesausschuss), which makes decisions regarding the reimbursement of medical procedures. IQWiG reports about advantages or disadvantages of medical procedures usually have a large impact on reimbursement decisions made by the Common Federal Council.

In May 2015, the IQWiG released a report about potential benefits of continuous glucose monitor (CGM) use in people with diabetes (www.iqwig.de/de/projekte-ergebnisse/projekte/nichtmedikamentoese-verfahren/d12-01-kontinuierliche-interstitielle-glukosemessung-cgm-mit-real-time-messgeraten-bei-insulinpflichtigem-diabetes-mellitus.3258.html). They concluded that while there was evidence that CGM is beneficial with regard to HbA1c improvement in adults with type 1 diabetes, there was less clear evidence for the avoidance of severe hypoglycemia in this patient group. Furthermore, the IQWiG found an indication that CGM might be beneficial with regard to hypoglycemia avoidance and HbA1c improvement in children with type 1 diabetes.

The IQWiG conservatively assesses CGM benefits in their report; they prefer clinical end points such as severe hypoglycemia over biochemical end points, that is, glucose readings obtained with blood glucose meters or CGM systems. This preference of the IQWiG is shared by other regulatory bodies including the Federal Drug Administration and the National Institute for Clinical Excellence.

In 2013, Ly et al from Australia published a study in the highly ranked journal *JAMA* that reported a significant reduction in the frequency of moderate hypoglycemia (defined as the requirement of third-party assistance for recovery from hypoglycemia) as well as severe hypoglycemic events (SH; defined as a hypoglycemic episode accompanied by seizure or coma) in patients with type 1 diabetes who used an insulin pump (continuous subcutaneous insulin infusion; CSII) with low glucose suspend (LGS).[1] The authors concluded that there is a benefit of using LGS in a clinical end point (combined rate of severe and moderate hypoglycemia). The study raised a lot of attention and was widely cited and also found its way into standards of care, in

[1]Science & Co, Düsseldorf, Germany
[2]Forschungsinstitut Diabetes Akademie Bad Mergentheim (FIDAM), Bad Mergentheim, Germany

Corresponding Author:
Lutz Heinemann, PhD, Science & Co, Kehler Str 24, 40468 Düsseldorf, Germany.
Email: l.heinemann@science-co.com

Table 1. *Definition of SH Events and Serious Hypoglycemic Events.*

Definition	*Susceptibility to systematic bias*
Severe[a]: Hypoglycemic event "Requiring assistance from another person"[b] or specific neuroglycopenic symptoms (eg, impaired level of consciousness or confusion), "Requiring assistance from another person"[c]	*Possible, because based on information provided by patient(s) or parents of patient(s)*
Serious[d]: Hypoglycemic events with hypoglycemic seizure or coma	*Unlikely*

[a]*In the publication, defined as moderate hypoglycemia, although it better suits the definition of sum of severe and moderate hypoglycemia. It corresponds to the definition of SH events for all other studies included in the IQWiG CGM report, as well as the definition by the ADA 2013.* [b]*Patients >12 years.* [c]*Patients ≤12 years.* [d]*In the publication, defined as SH. The definition in the publication exclusively includes hypoglycemic events that fulfilled at least 1 criterion for serious adverse events. For this reason, these hypoglycemic events are defined as serious hypoglycemic events in this report.*

particular in pediatric standards.[2-4] This publication was also used in a HECON analysis which is used for reimbursement attempts.[5]

The Ly study (as this article is most often referred to) was submitted to IQWiG when this institute started its evaluation of the evidence for CGM usage. The expectation was that this study would strengthen the case for CGM, especially for the combination of CGM with CSII because this study proved the impact of CGM on the clinical outcome of severe hypoglycemia rather than relying on biochemically defined hypoglycemia or symptomatic hypoglycemia. This positive outcome of such an efficacy analysis can be expected to shift the weighing of benefits and disadvantages toward a reimbursement for CGM.

We would like to emphasize that it is not our intention to disrespect the Ly study in any way. We simply wish to present the differing view of the IQWiG on the Ly study to enable the reader to come to his or her own conclusion about this study. This is also likely helpful to get a better understanding of the requirements of outcomes studies in the view of regulatory authorities like the IQWiG. Please also acknowledge the complexity of the definitions for hypoglycemic events used (Table 1) by Ly et al and the IQWiG.

Evaluation of the Evidence for CGM Usage by the IQWiG

The IQWiG approached the study center in Australia to get access to the study data. After receiving the data, the IQWiG ran its own analysis. This procedure is also a reflection of the scrutiny and systematic approach with which the IQWiG performed the complete CGM analysis. It is clearly their aim to come to reliable statements when analyzing published studies. While the authors of the Ly study concluded that "sensor augmented pump therapy reduced the rate of severe and moderate hyperglycemia in patients type 1 diabetes," the IQWiG reached a contrary conclusion.

Decisive for this was the statistical analysis conducted by Ly et al. During the 6-month study phase, 41% of CGM users experienced at least 1 moderate or severe hypoglycemic event, whereas only 31% of control patients were affected. Thus CGM patients had a 57% higher likelihood to be affected by at least 1 moderate or severe hypoglycemic event. Also, the incidence rates of moderate and severe hypoglycemic events in the CGM group were more than 2-fold higher than in the control group (28.4 vs 11.9 events per 100 patient months). There was also no significance test of this huge difference reported by the authors of the Ly study. This picture shifts completely if moderate and severe hypoglycemic events rates were adjusted for prebaseline differences. The incidence rates of hypoglycemic events in the CGM group decreased from 28.4 to 9.5 events per 100 patient months, whereas the event rate in the control group increased from 11.9 to 34.2 events per patient months. This difference was tested for statistical significance and yielded a positive result in favor for CGM. This dramatic difference between observed and adjusted event rates was due to a large baseline difference in the hypoglycemic events rates (129.6 vs 20.7 hypoglycemic events per 100 patient months), which can be perceived as very unfortunate given that this study was randomized.

Another reservation of the IQWiG refers to the use of event rates (in this case of moderate and severe hypoglycemic events per 100 patient months). The IQWiG doubt that likelihood of hypoglycemic events is equally distributed, because it can be assumed that the likelihood of a first hypoglycemic event in a certain patient is different from the likelihood of a repetitive event in another patient. Therefore the use of Poisson models for significance tests of events rates were considered as not appropriate. A more detailed explanation of the problematic use of events rates is provided by Windeler and Lange.[6] This led the IQWiG to evaluate this result to not be evidence of the superiority of CGM for preventing moderate or severe hypoglycemic events.

Also the second finding by Ly et al that the incidence of SH (with coma and seizure) was significantly reduced by CGM (0 events vs 2.2 events per 100 patient months, $P = .02$) was not evaluated as evidence of the ability of CGM to prevent clinical end points like SH by the IQWiG. The reason for this was that an inquiry to the authors of the Ly study revealed that the incidence rate of 2.2 events per 100 patient months was based on only 3 patients. Furthermore, the IQWiG criticized that the use of event rates for severe hypoglycemic events is highly problematic to compensate a low statistical power due to study drop outs for the same reasons as explained above.[6] Given that the significant result

Table 2. *Changes in HbA1c in the Ly study (Mean [95% Confidence Interval]; n/N).*

	Start of the study	End of the study (6 months)	Change in HbA1c	Group difference[a] P value
LGS	7.6 [7.4, 7.9]; 46/46	7.5 [7.3, 7.7]; 46/46	–0.1 [–0.3, 0.03]	0.07 [–0.2; 0.3]; P = 0.55
SMBG	7.4 [7.2, 7.6]; 49/49	7.4 [7.2, 7.7]; 49/49	–0.06 [–0.2, 0.09]	

[a]*Result from mixed-effects model repeated measures, adjusted according to HbA1c value at the start of the study, visit interaction, age group interaction, as well as age group × visit interaction, age group × treatment group interaction, and treatment group × visit interaction. n, number of patients analyzed; N, number of randomized patients.*

was based on only 3 patients, a drop out of 9 patients out of 95 patients without performing an intention to treat analysis was regarded as not convincing evidence for the prevention of SH by the IQWiG.

Furthermore, it is obvious that in the Ly study not only CGM as an intervention was studied, but the combination of CGM with CSII versus self-monitoring of blood glucose (SMBG) with CSII in the control group. The patients in the intervention group used an insulin pump with LGS function which automatically turned off the basal insulin supply for a maximum of 2 hours if the glucose value measured by the CGM device fell <60 mg/dl. From a more general point of view, it is important to note that the IQWiG also consider the extent of the lowering of blood glucose levels in the treatment group as highly relevant to be able to interpret changes in the hypoglycemic event rates observed in a controlled study comparing different therapy options for lowering blood glucose levels appropriately. Therefore, hypoglycemic events in conjunction with long-term lowering of blood glucose levels measured using HbA1c values must also be evaluated in studies comparing LGS versus SMBG.

The IQWiG first analyzed the data with regard to the

1. extent of the long-term lowering of blood glucose levels,
2. occurrence of severe/serious hypoglycemic events
3. combined assessment of severe/serious hypoglycemic events and lowering of blood glucose levels.

The following text is an excerpt of the conclusion section of the IQWiG CGM report (p. 166-174). It is a direct translation (not a summary or the like).
This section is in italics to highlight this point.

Outcome Evaluation for HbA1c

Analyzing the percentage of patients with an HbA1c value below a seemingly appropriate cutoff value appeared most suitable as it allowed for analyzing individual patients having attained glycemic control as recommended by medical associations. Nonetheless, only the result of the difference in mean HbA1c changes was reported in the Ly study. The results of the Ly study showed a high risk of bias with respect to HbA1c, as the difference in the percentage of data not used is >5% (CGM 1/46 patients [2%]; BGSM 4/49 [8%]):

- *Risk of bias at the study level—low*,
- *Blinding of the end point investigator—yes, (statement is based on asking the authors)*
- *ITT principle adequately implemented—no*,
- *Result-independent reporting—yes*,
- *Additional aspects missing—yes*,
- *Risk of bias at the end point level—high.*

The analysis of the HbA1c throughout the course of the study (differences in mean changes)—comparison between LGS versus SMBG (Table 2)—showed no statistically significant differences between the groups. So, it is not possible to determine an effect in favor of either treatment option with regard to HbA1c.

Outcome Evaluation for Hypoglycemic Events

In the Ly study, both severe and serious hypoglycemic events were reported. No details were given on the frequency of nighttime hypoglycemic events. Mild hypoglycemic events were also not reported. The primary end point was the occurrence of SH events. In the Ly study, 3 different categories of hypoglycemic events were used:

- *moderate hypoglycemia*,
- *severe hypoglycemia, and*
- *sum of severe and moderate hypoglycemia.*

All SH events also fulfilled criteria for moderate hypoglycemic events, thus all SH events were recorded as the "sum of severe and moderate hypoglycemia." Serious hypoglycemic events were recorded under "severe hypoglycemia" (defined by coma or seizure).

The definition of moderate hypoglycemia corresponds to the definition of SH events for other CGM studies evaluated by the IQWiG as well as the ADA 2013 definition.[7] *According to this definition, SH must be included in the category moderate hypoglycemic events. However, in the Ly study, moderate hypoglycemia and SH were recorded separately, but the sum of both was also reported. The definition of SH in the Ly study included only hypoglycemic events which fulfilled at least 1 criterion for serious adverse events.*

The results reported for SH events were evaluated as being subject to a high risk of bias:

Table 3. *Results for Severe Hypoglycemic Events.*[a]

Number of patients n/N and group	Number of patients with at least 1 SH event,[a] number (%),[b] OR [95% CI]; P value	Incidence rate of SH events per 100 patient months at the end of the study (start of the study), P value	Incidence rate of SH events (per 100 patient months), adjusted[c], P value	Absolute number of SH events at the end of the study (start of the study), P value
LGS				
41[c]/46	17 (41)[d]	28.4 (129.6)[e]	9.5	35 (175[e])
Control group				
45[f]/49	14 (31)[d]	11.9 (20.7)[e]	34.2	19[d,g] (28[e])
	1.57 [0.65, 3.80] P = .359[h]	P nr	P < .001[i]	P nr

[a]Includes the hypoglycemic events that were reported in the publication as sum of SH and moderate hypoglycemia. The definition of moderate hypoglycemia corresponds to the definition of SH events for all other studies included in the IQWiG report, as well as the definition by the ADA 2013. The definition of SH in the Ly study exclusively included hypoglycemic events which fulfilled at least 1 criterion for serious adverse events. For this reason, these hypoglycemic events are defined as serious hypoglycemic events in this report. [b]The percentage refers to the number of analyzed patients (n). [c]Five patients were excluded from the analysis as they revoked their declaration of consent. One patient moved away, 3 patients discontinued treatment, and 1 patient did not show up to the follow-up appointments. [d]Personal calculation. [e]Information for 45 patients. [f]Four patients were excluded from the analysis as they revoked their declaration of consent. One patient moved away, 2 patients discontinued treatment, and 1 patient was dissatisfied with the randomization. [g]There were 13 moderate and 6 SH events reported; however for the sum there were 13 hypoglycemic events (Table 2 of the publication) or 16 hypoglycemic events (page 1244 in the publication text). [h]Personal calculation, absolute exact test (CSZ method). [i]o-inflated Poisson model, adjusted in accordance with the values at the start of the study. CSZ, test statistics with chi-square statistic as classification criterion; n, number of patients analyzed; N, number of randomized patients; nr, not reported; OR, odds ratio.

- *Risk of bias at the study level—low*,
- *Blinding of the end point investigator—no*,
- *ITT principle adequately implemented—yes*,
- *Result-independent reporting—yes*,
- *Additional aspects missing—no, (see below)*
- *Risk of bias at the end point level—high.*

The definition of SH events is regarded as problematic (see below) as well as—with regard to the evaluation the number of SH events per 100 patient months—the adjustment for the baseline values.

The risk of bias for the results for serious hypoglycemic events was also analyzed:

- *Risk of bias at the study level—low*,
- *Blinding of the end point investigator—unclear*,
- *ITT principle adequately implemented—no*,
- *Result-independent reporting—yes*,
- *Additional aspects missing—yes*,
- *Risk of bias at the end point level—high.*

The 9 patients who discontinued the study were not included in the evaluation of both the SH and serious hypoglycemic events. In view of the rarity of serious hypoglycemic events this leads to a high risk of bias.

For SH events, the missing blinding of the end point investigator as well as the definition itself was problematic—the latter because it included the criterion "Requiring assistance from another person" (Table 1). This criterion is susceptible to a subjective influence as could also be understood as, for example, the administration of dextrose by another person when unspecific symptoms occur.

For serious hypoglycemic events, blinding of the end point investigator was evaluated as unclear as it was rarely reported by only the patients themselves, but also by the attending doctors. Whether the latter was subject to blinding is unclear from the study register entry or the publication.

The results of the Ly study for SH (Table 3) and serious hypoglycemic (Table 4) events are summarized. With regard to the number of patients with at least 1 SH event, the difference between the groups was not statistically significant. Regarding the absolute number of SH events the numeric difference between the groups was clearly against a benefit of CGM with LGS function. Information provided in the study report shows that the number of events between patients varies massively and is dominated by some outliers. Therefore, analyses that are based on the number of events cannot be interpreted sufficiently.

According to the study report, the results for SH events reported in Ly et al do not refer to the whole study duration (6 months), but only for the last 3 months. The results for this period of time are in line with the number of patients with at least 1 SH event and the total number of SH events reported for the total study duration.

Therefore, using the available data, it is not possible to determine a beneficial effect for either treatment option with regard to the occurrence of SH events.

With regard to the number of patients with at least 1 serious hypoglycemic event, the difference between the groups was not statistically significant. With regard to the incidence rate of serious hypoglycemic events per 100 patients, a statistically significant difference was reported between the groups in favor of CGM with LGS function. However, the acceptance of independence of events used by the Poisson

Table 4. *Results for Serious Hypoglycemic Events*[a].

Number of patients n/N and groups	*Number of patients with at least 1 serious hypoglycemic event*[a]*, number (%)*[b]*, P value*	*Incidence rate of serious hypoglycemic events (per 100 patient months) end of the study (start of the study), P value*	*Absolute number of serious hypoglycemic events at end of the study (start of the study), P value*
LGS			
41[c]*/46*	*0 (0)*	*0 (1.8)*	*0 (5)*
Control group			
45[d]*/49*	*3 (7)*[e] *P = .101*[f]	*2.2 (2.1)* *P = .02*[g]	*6 (6)* *P nr*

[a]*In the publication, defined as SH (added by LH: with seizure or coma). The definition in the publication exclusively includes hypoglycemic events which fulfilled at least 1 criterion for serious adverse events. For this reason, these hypoglycemic events are defined as serious hypoglycemic events in this report.* [b]*The percentage refers to the number of analyzed patients (n).* [c]*Five patients were excluded from the analysis. Four patients revoked their declaration of consent. One patient moved away, 3 patients discontinued the treatment. One additional patient did not show up to the follow-up appointments.* [d]*Four patients were excluded from the analysis as they revoked their declaration of consent. One patient moved away, 2 patients discontinued treatment, and 1 patient was dissatisfied with the randomization.* [e]*Personal calculation.* [f]*Personal calculation, absolute exact test (CSZ method, see above).* [g]*P value for incidence rate using Poisson regression, exact calculation. Underline, result of asking an author.*

model to calculate the P *value is contested, as the 6 hypoglycemic events occurred in 3 patients. Therefore, the result of the reported analysis is questionable. Summarized, using the available data, it is not possible to determine a beneficial effect for either treatment option with regard to the occurrence of serious hypoglycemic events.*

Combined Outcome Evaluation of Hypoglycemic Events and HbA1c

A combined assessment was undertaken with regard to the lowering of blood glucose levels and the occurrence of SH and serious hypoglycemic events. The combined assessment of the HbA1c value and the hypoglycemic events provided no indication for an advantage of CGM with LGS function vs SMBG.

Conclusions

The IQWiG reanalysis of the Ly study reached contrary conclusion than the study authors. In their publication, Ly et al reported a statistically significant difference of moderate or severe hypoglycemic events rates in favor of the group with CGM plus LGS function with almost unchanged glycemic control (HbA1c). However, the IQWiG concluded that results regarding hypoglycemia were based on evaluations which could not be classified as reliable, due to the use of event rates. In addition, the huge baseline difference in the prevalence of moderate hypoglycemia in spite of randomization represents a big problem for the interpretation of the hypoglycemia results. A further issue identified by the IQWiG was the distribution of SH events and the low number of affected patients. The incidence reduction of SH relied only on 3 patients out of 45 patients; underscoring the problems of using events rates instead of patients affected. In a previous publication the current director of the IQWiG classified the use of event rates as a "dubious concept," which might be suited "for a chapter in Methodological Errors in Medical Research."[6]

From the IQWiG regulatory perspective, it seems important to assess both the effect of an intervention or device on defined clinical outcomes as well as the methodological quality of the study. Our conclusion is that the thorough reanalysis by the IQWiG provides a new and different methodological approach than the one used by both Ly et al and the reviewers of *JAMA*. The following aspects should be considered in further planning of CGM studies:

- Given these discrepancies in the evaluation of the Ly study, a broader discussion appears to be called for on the different views and interpretation of the data by Ly et al and the IQWiG. It can be expected that Ly et al will share their views on the IQWiG position in a commentary or rebuttal letter.
- A discussion in the diabetes community at diabetes meetings about the methodological standards of regulatory authorities also seems necessary, given the disappointing evaluation of the IQWiG report toward the Ly study. It may also be worth discussing, although not at our discretion, whether each regulatory authority should define its own methodological standards and the implication for the conduct of evaluation trials in the field of new medical devices. This is a special problem in cases where multinational studies are performed.
- For the use of CGM and LGS, more evidence is needed to evaluate the efficacy and effectiveness of the product. These studies should take into account the additional evaluation criteria applied by IQWiG for the Ly study. An open discussion at future diabetes meetings is needed about the benefits of LGS and the quality of the evidence that supports the use of this product.

- For further studies evaluating potential benefits of CGM with regard to clinical end points such as moderate and severe hypoglycemia, it is important to ensure a comparable distribution of moderate and severe hypoglycemic events between both groups at baseline. For avoiding important baseline differences in spite of randomization, a stratification strategy might be appropriate to ensure a comparable distribution.
- To demonstrate that clinical end points such as severe hypoglycemic episodes (accompanied by coma or seizure) could be avoided, a sufficient number of affected individuals should be recruited, since the use of events rates seems problematic to compensate for low statistical power which is frequently associated with the study of rare events.
- A meta-analysis of all appropriate studies taking into account the above mentioned methodological issues to see how the conclusions by Ly et al fit with the rest of the literature would be welcomed.

Abbreviations

CGM, continuous glucose monitoring; CSII, continuous subcutaneous insulin infusion; IQWiG, German Institute for Quality and Efficiency in Health Care; ITT, intention to treat analysis; LGS, low glucose suspend; SH, severe hypoglycemic event; SMBG, self-monitoring of blood glucose.

Acknowledgments

We would like to thank a number of colleagues very much for their quite helpful comments to this manuscript, especially Hans DeVries and David Klonoff.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: LH holds shares in the Profil Institute for Metabolic Research, Neuss, Germany, and the Profil Institute for Clinical Research, San Diego, USA. He is a consultant for a range of companies that develop new diagnostic and therapeutic options for the treatment of diabetes. NH is member of the Global Diabetes Educator Advisory Board of Eli Lilly and the Global DAWN 2 Study International publication committee. He is also member of the national advisory board of DAWN 2 (Novo Germany), Abbott, Germany, and the HypoDE study (Dexcom). He receives speaker honoraria from Berlin Chemie, Astra Zeneca, Novo Nordisk, Germany.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Ly TT, Nicholas JA, Retterath A, Lim EM, Davis EA, Jones TW. Effect of sensor-augmented insulin pump therapy and automated insulin suspension vs standard insulin pump therapy on hypoglycemia in patients with type 1 diabetes: a randomized clinical trial. *JAMA*. 2013;310:1240-1247.
2. Ly TT, Maahs DM, Rewers A, Dunger D, Oduwole A, Jones TW. Assessment and management of hypoglycemia in children and adolescents with diabetes. *Pediatr Diabetes*. 2014;15(suppl):180-192.
3. Rewers MJ, Pillay K, de Beaufort C, et al. Assessment and monitoring of glycemic control in children and adolescents with diabetes. *Pediatr Diabetes*. 2014;15(suppl):102-114.
4. Phillip M, Danne T, Shalitin S, et al. Consensus statement: use of continuous glucose monitoring in children and adolescents. *Pediatr Diabetes*. 2012;13:215-228.
5. Ly TT, Brnabic AJ, Eggleston A, et al. A cost-effectiveness analysis of sensor-augmented insulin pump therapy and automated insulin suspension versus standard pump therapy for hypoglycemic unaware patients with type 1 diabetes. *Value Health*. 2014;17:561-569.
6. Windeler J, Lange S. Events per person year—a dubious concept. *BMJ*. 1995;310:454-456.
7. American Diabetes Association. Standards of medical care in diabetes—2013. *Diabetes Care*. 2013;36:S11-S66.