2.5 Effectiveness of method improvements of OECD GD 75 – Evaluation by the ICP-PR Bee Brood Working Group*

Roland Becker[1], Johannes Lückmann[2], Jens Pistorius[3]

[1] Basf Se, Apd/Ee - LI425, 67117 Limburgerhof, Germany, e-mail: roland.becker@basf.com
[2] Rifcon GmbH, Goldbeckstraße 13, 69493 Hirschberg a. B., Germany, e-mail: johannes.lueckmann@rifcon.de
[3] Julius Kühn-Institute, Institute for Plant Protection in Field Crops and Grassland, Messeweg 11/12, 38104 Braunschweig, Germany, e-mail: jens.pistorius@jki.bund.de
* ICP-PR Bee Brood Working Group and in cooperation with the AG Bienenschutz (Germany)

Abstract

Background: The OECD Guidance Document No. 75 (2007)[1] is a method to investigate potential effects of plant protection products on the brood of honeybees (*Apis mellifera* L.), with the 'brood termination rate' (BTR, failure of individual eggs or larvae to develop) as the key endpoint. As in recent years a number of studies displayed a strong variability in BTRs, Pistorius *et al.* (2012)[2] recommended some measures for improvements. First results in the season 2011 indicated that these measures led to lower BTRs and lower variability. The ICP-PR bee brood working group has evaluated the effectiveness of the recommended measures for improving the reliability of the method and the resulting BTRs and reports in this paper.

Results: To evaluate the effectiveness of these measures a data analysis of a total of 62 studies was performed which were carried out in Germany and Switzerland between 2011 and 2014. Based on this analysis, the mean BTR in the control was 29.2% and this result did not display a distinct improvement compared to the historical data (34.7%) (Pistorius *et al.* 2012[2]) and neither compared to the data of the bee brood ring-test (28.0%) (Schur *et al.* 2003[3]). Moreover, the proportion of replicates (colonies) with BTRs ≤30% amounted to be 61.5% compared to 55.6% in the years before. And every 2nd study displayed BTRs >30% in two or more replicates. Also, the proportion of replicates (colonies) with BTRs ≤40% amounted to 76.9% compared to 68.3% in the years before and just 21.7% of the studies displayed BTRs >40% in two or more replicates.

Conclusion: Overall, these findings highlight that the test method according to the OECD Guidance document in 2007was not be considerably improved with the recommended measures. But although the reliability of the method and a reliable interpretation regarding potential effects of a plant protection product (PPP) on bee brood was not given in all studies, it currently remains the only available test method to address the potential risk of a plant protection product on honeybee larval development in realistic worst-case (semi-field) exposure conditions. Among other factors, it is assumed that the limitations are most likely due to the confined semi-field conditions. Further work should investigate potential additional improvements in semi-field conditions and also brood termination rates in field conditions.

Key words: honeybees, bee brood test, OECD GD 75, brood termination rate

1. Introduction

Based on EU Regulation 1107/2009/EC, the risk to honeybee larvae or honeybee brood (*Apis mellifera* L.) needs to be addressed in the current regulatory risk assessment on bees and, in case of potential concern, appropriate tests must be conducted. Also, in the 'Guidance Document on the risk assessment of plant protection products on bees' (EFSA 2013[4]), it is concluded that concerns on bee brood need to be addressed. EFSA recommends specific tests, e.g. the OECD Guidance Document No. 75 (2007)[1] (hereafter OECD GD 75) next to the Oomen bee brood feeding test (Oomen *et al.* 1992[5]) as possibilities to refine the risk assessement on honeybee brood if there was reason for concern.

Data analysis of Becker & Lückmann (2011)[6] and Pistorius *et al.* (2012)[2] demonstrated that the key endpoint 'Brood Termination Rate' (hereafter called BTR) is subject to a certain degree of variation in confined semi-field conditions, e.g. resulting in replicates with increased rates up to 100% in the

control and reduced rates in the reference item group down to 21%. In addition, sometimes a high variation occurs between the replicates of a respective treatment group. Such high variation complicates the interpretation of results regarding potential brood effects of the test items, with the outcome that studies sometimes are regarded as invalid (Pistorius *et al*. 2012[2]).

To improve the current methodology, the Working Group 'Honeybee brood' of the AG Bienenschutz discussed some aspects of the method, e.g. timing of the experiment, crop area, size and composition of bee colonies, digital comb assessment vs. acetate sheet assessment of brood cells in spring 2011 (Pistorius unpublished[7]; Becker & Lückmann 2011[6]), and proposed concluded recommendations at the ICP-BR meeting in Wageningen, The Netherlands 2011 (Pistorius *et al*. 2012[2]).

There it was recommended:

- to use bigger colonies with 3 to 4 brood combs, containing a high number of capped cells
- to avoid major modifications of the colonies shortly before application
- to use 4 instead of 3 replicates for better interpretation of data
- to start the study early in the season, if possible
- to use large tunnels, which provide effective crop areas of >60 m^2, preferably >80 m^2
- to water the crop if dry conditions reduce nectar flow
- to evaluate termination rate and pupal mortality in the toxic reference item

It was suggested also to use digital brood cell assessment instead of the cell assessment on acetate in order to reduce the time span of brood combs outside the hive and consequently the stress for the colonies, and to increase the number of observed cells to 200 to 400. Additionally the data analysis had shown that colonies with more than 7,000 bees displayed higher probabilities to achieve BTRs ≤30 % in the control.

In the season 2011 these measures seemed to indicate a distinct improvement as mean BTR decreased from 34.7% to 21.7% and the proportion of replicates with BTR ≤30% increased from 55.6% to 78.0%.

The current paper evaluates the effectiveness of these measures for studies carried out in Germany and Switzerland between 2011 and 2014 and re-investigates BTR driving factors.

2. Material and Methods

To obtain a reliable database, contract labs and plant protection product producing companies were requested in summer 2013 and 2014 to submit control and reference item data from bee brood studies performed according to OECD GD 75 and Pistorius *et al.* (2012)[2].

The following parameters were requested for each replicate (colony):

- brood termination rate (BTR) at 'Brood area Fixing Day' (hereafter BFD) 21/22
- day of the year at BFD 0 (calculated from the date of BFD 0)
- colony strength at BFD 0
- number of days in the tunnel before application
- total number of cells with brood, pollen or nectar/honey in a colony at BFD 0
- number of cells with eggs, pollen or nectar on marked comb side(s) at BFD 0
- number of cells with pollen or nectar/honey on comb side(s) adjacent to marked comb side(s) at BFD 0
- mean number of dead pupae/day during post application period
- application rate of the reference item fenoxycarb

In total, data of 75 honeybee brood studies were provided from Germany/Switzerland, France, Spain and the US. The studies were mainly carried out under GLP and were provided by BASF SE, Bayer CropScience, BioChem agrar, Dow AgroSciences, DuPont, Eurofins Agroscience, Ibacon, Ies, Rifcon, Syngenta and Testapi. A summary of the available number of studies and the number of replicates (tunnels) for the control and reference item for the respective countries is given in Table

1. This summary contains also data of six terminated studies and data of three studies which were initiated in 2014, but which were not finalised yet.

For the evaluation of the 'effectiveness of the measures proposed by Pistorius *et al.* (2012)[2]' (chapter 0) and the 'analysis of additional potential BTR driving factors' (chapter 0) the studies from Germany/Switzerland were used. To be in line with the data analysis of Pistorius *et al.* (2012)[2] which contained only finalised studies a total of 8 out of a total 62 studies from Germany/Switzerland were not considered as they were terminated early due to high BTRs (6 studies) or were carried out at a very late growth stage of the crop (BBCH code) (1 study) or because of daily rain during the exposure period (1 study). Out of this data set only 13 studies included all requested information which limited the analysis of some parameters. The descriptive statistics, e.g. calculation of medians, means, standard deviations, minima, maxima were performed with the reduced and the complete data set.

The 54 analysed studies from Germany/Switzerland were evenly distributed over several years: 13 studies from 2011, 16 from 2012, 15 from 2013 and 10 from 2014.

Table 1 Number of bee brood studies performed since 2011 and provided for data analysis

Country	Number of studies [n]	Number of replicates (tunnels) [n]	
		Control	Reference item
Germany/Swiss	54°(62*)	208° (239*)	192° (207*)
France	4	12	12
Spain	5	18	14
US	4	16	13

* all studies, including 6 terminated studies and 3 studies started in 2014 but not finalised; containing at least BTR data but not necessarily complete data sets; ° 8 studies were excluded due to high BTR, late BBCH at application or daily rain during exposure period

For comparing the current data to those derived from brood studies performed before 2011 (i.e. before the recommendations were formulated by Pistorius *et al.* 2012[2]), these last will be described as 'historical data'.

3. Results

3.1. Results of bee brood studies from Germany/Switzerland

3.1.1. Descriptive statistics

A summary of the descriptive statistics of the bee brood studies performed before 2011 (=historical) and during or after 2011 is given in Table 2. It shows that the values of the current studies were only slightly better compared to the historical data and thus the suggested improvements had not led to distinctly lower BTRs and much lower variability.

Table 2 Summary of descriptive statistics of bee brood studies performed before 2011 and in or after 2011

Parameter	Brood termination rate [%]			
	Historical data		Data ≥ 2011	
	Control n=63	Reference n=54	Control n=208° (n=239``)	Reference n=192° (n=207``)
Median*	25.9	83.4	23.4 (26.5)	77.4 (75.0)
Mean	34.7	76.8	29.2 (32.9)	70.7 (70.4)
Standard Deviation*	24.8	24.2	21.6 (24.4)	27.4 (27.3)
Minimum	4.9	20.9	2.0 (2.0)	2.6 (2.6)
Maximum	100	100	100 (100)	100 (100)

n=number of replicates (colonies), * calculated from all replicates; ° 8 studies excluded; `` all studies

3.1.2. Reliability of the test method: control

To evaluate the reliability of the test system in the control, it is assumed that relative low levels of BTRs in the controls indicate good reliability of the test system, and reversely that relative high levels indicate bad reliability of the test system. We analysed the distribution of the BTRs according to magnitude (size) categories and the numbers of replicates with BTR below a certain threshold. And we studied the distribution of replicates with BTR´s of >30% or >40%.

The BTRs of colonies follow a normal distribution when arranged according to magnitude (size), with a shifted maximum of approximately 26% at BTRs between 10 and 20% (see Figure 1). The total of colonies below BTRs ≤30% and ≤40% summed up to 61.5% and 76.9%, respectively. Considering all studies together without studies excluded, these totals were 55.6% and 70.7% for BTRs ≤30% and ≤40%, respectively. Comparing these values to the historical data (Table 3) it indicates that a high proportion of replicates had BTRs distinctly higher than 30% and 40%.

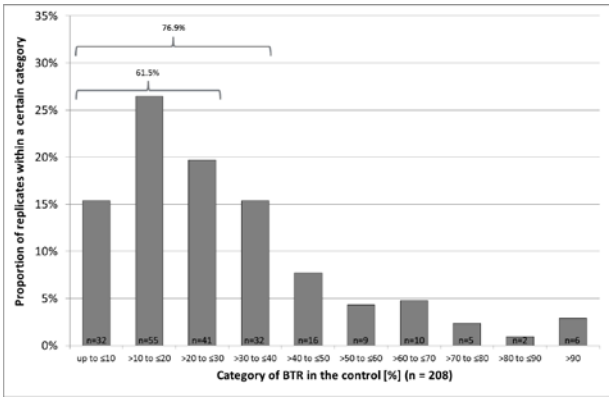

Figure 1 Distribution of BTRs in the control according to size categories

Table 3 Reliability of the test method in the control, according to the BTR-level

Replicates	Historical data (n=63) % of replicates	Data ≥ 2011 (n=208°, n=239*) % of replicates
- with BTRs ≤30%	55.6	61.5 (55.6)
-with BTRs ≤40%	68.3	76.9 (70.7)

n=total number of replicates (colonies), ° 8 studies excluded; * all studies

For the analysis of studies according to their BTR-levels, only studies with four replicates (colonies) were considered. The results show that 50.0% of the performed studies exhibit no or one replicate with a low BTR (>30%) whereas this was 78.3% for an intermediate BTR (>40%, see Figure 2).

Taking into account all studies, these levels were 45.3% and 66.0% for BTRs >30% and >40%, respectively.

Therefore overall the data show that few studies in the controls have low BTR-levels as an indicator or reliability of the study. It means that the reliability of the test method is limited. Because such high variability of BTRs as in the controls must be assumed for the test item groups as well, several studies could not be interpreted for effect of the PPP tested. And the question remains unanswered whether the obtained results indicate the real impact of a PPP on bee brood or whether data showed chance results.

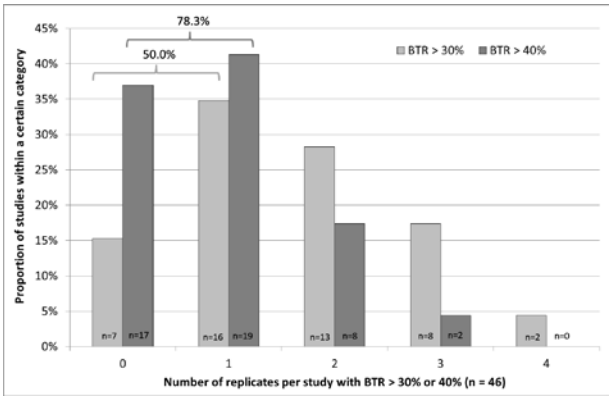

Figure 2 Distribution of studies comprising replicates with >30% or >40% BTR in the control

3.1.3. Reliability of the test method: reference item

For the evaluation of the reliability of the reference item group (i.e. where the PPP is applied for assessing the effect on bee brood) the number of replicates was assessed with a BTR **≥** 70%.

The data analysis showed that 56.6% of all replicates were above this value of BTR ≥70% (53.3% if no colonies were excluded) and this is somewhat lower than in the historical data (Table 4). For those replicates with a lower BTR, 86% of the colonies had a pupal mortality of ≥80 dead pupae during the post-application period (83% for all colonies). Thus a total of 93.8% (92.3%) of the colonies displayed either an increased BTR or an increased pupal mortality, confirming the exposure of the bees.

Table 4 Reliability of the test method in the reference item according to the BTR-level

Replicates	**Historical data, n=54**	**Data ≥ 2011 (n=192°, n=207*)**
- with BTRs ≥70%	70.4	56.6 (53.3)

n=number of replicates (colonies), ° 8 studies excluded; * all studies

3.1.4. BTRs in relation to data of the bee brood ring-test (Schur *et al.* 2003[3])

In the honeybee brood ring-test in 2002 (Schur *et al.* 2003[3]), the mean BTR (n=5 studies with one replicate, each) was 28.0 ± 14.7% (minimum: 8%, maximum: 43%) for the control and 98.8 ± 2.7% for the reference item (minimum: 94%, maximum: 100%) (calculated from the published data). Thus the current mean BTRs in the controls were at the same level whereas those of the reference item were insignificantly lower.

3.1.5. Effectiveness of recommendations on BTRs given by Pistorius *et al.* (2012)[2]

First results from the season 2011 (Pistorius *et al.* 2012[2]) indicated that the proposed measures led to an improvement of BTRs. In fact after the application of the recommendations the mean BTR

decreased from 34.7% to 21.7% and the number of replicates with BTRs ≤30% increased from 55.6% to 78.0%.

The evaluations of the effectiveness of the measures are summarized in Table 5. They indicate that most of the recommendations worked but, overall, they did not confirm the preliminary trend from 2011. E.g. the crop areas increased to more than 65 m^2 in the controls did not result in a further improvement of the BTRs. The same was true for the recommendations about 'colony strength' and 'number of marked cells'. But in contrast to Pistorius *et al.* (2012)[2] no correlation was found between 'day of the year at BFD 0' and the BTR. A correlation may have been hidden by other effects, e.g. weather conditions. The influence of 'watering the crop at dry conditions' could not be evaluated due to the lack of corresponding data.

Table 5 Summary of recommendations of Pistorius *et al.* (2012)[2] and their success

Parameter	Recommendation	Result
Effective crop area	> 60 m^2, preferably 80 m^2	No effect, **but** if crop area is ≥ 65 m^2 in controls, no further improvement of BTR by increase up to 95 m^2
Day of the year at BFD 0	early start in the season, if possible	No effect, **but** influence of weather conditions unclear
Colony strength at BFD 0	approximately 7,000 bees	Colonies with 6,000 to 8,000 bees displayed a higher probability to obtain BTRs ≤ 30% (chi^2-test, p=0.019) (not for BTRs ≤ 40%)
Number of cells to be marked	200 to 400	Studies with 300 to 400 marked cells provided good results
Endpoints in toxic reference	evaluation of BTR **and** pupal mortality	In the case of BTRs ≤70% increased pupal mortality proved exposure (86% of replicates with BTR ≤70% displayed ≥ 80 dead pupae during post-application period)
Application rate in reference item	single (150 g a.s./ha) or double rate (300 g a.s./ha)	Double rate displayed higher reliability: at single rate 73% of replicates with BTR <70% displayed >80 dead pupae during post-application period; at double rate it was 92.5%
Watering of crop	Should be done if dry conditions reduce nectar flow	Cannot be evaluated due to lack of data

For the toxic reference item the endpoints BTR and pupal mortality proved to be a reliable endpoint as an indicator of a sufficient exposure (see chapter 0), while the double rate gave a higher confidence than the single rate (see also Hecht-Rost *et al.* 2014[8]).

3.1.6. Analysis of additional potential BTR driving factors

Neither for the number of days in the tunnel before application, for the amount of brood in the colonies and for the number of eggs on the marked comb sides nor for the availability of pollen or honey/nectar in the colonies on the marked or adjacent comb sides a correlation with the BTR was found, taking into consideration that the analysis is limited by the lack of information about the weather conditions or of complete data sets (see chapter 0).

Table 6 Summary of effects of potentially BTR driving factors

Factor	Correlation with BTR
Number of days in the tunnel before application	No correlation, but influence of weather conditions in respective years unclear
Number of brood cells in a colony, or number of eggs on marked comb side(s) at BFD 0	No correlation found; **but more complete data sets necessary for a reliable evaluation**
Number of cells with *pollen* in a colony, on marked comb side(s) or on adjacent comb side(s) at BFD 0	Colonies with a lot of pollen in total or on marked/adjacent comb side(s) did not perform better than colonies without pollen; **but limited availability of data**
Number of cells with *nectar/honey* in a colony, on marked comb side(s) or adjacent comb side(s) at BFD 0	dito

As an example the correlation between the amounts of pollen on the marked comb side(s) at BFD 0 vs. the BTR at the end of the study is showing that colonies without pollen did, interestingly, not perform worse than colonies with pollen (Figure 3).

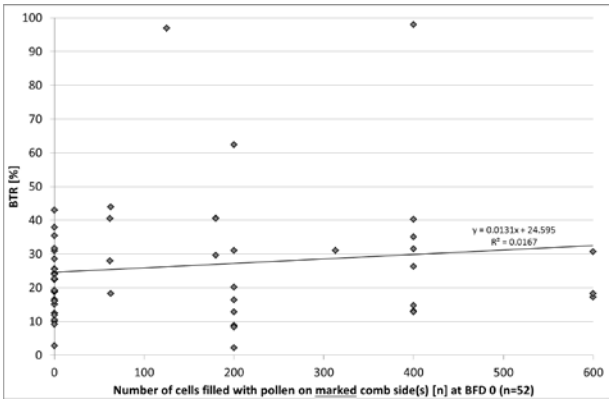

Figure 3 Influence of pollen amount on marked comb side(s) at BFD on BTR

But what are the driving factors then? There are some potential factors which can be influenced to a certain extend by the study set-up but others not. The first category might comprise factors like the growth stage of the crop at the start of the study or the time and extend between the preparation of the colonies and the start of the study. The latter point was already identified by the brood group of the AG Bienenschutz in spring 2011 (Pistorius unpublished[7]) to be most likely another driving factor which led to the recommendation to avoid major modifications of the colonies shortly before application (Pistorius *et al.* 2012[2]). But this recommendation was not specified later on. An analysis of its effect on BTR is very difficult as the timing, kind and degree of colony preparation during the time of study arrangement is normally not documented.

However, if colony modifications are needed and it is not possible to use naturally grown nuclei it seems advisable for optimal results that the colonies are adjusted early to adapt before trial initiation. As a recommendation, assess colony strength 21 days after the modifications and then allot comparable colonies to the control, reference item and test item treatment group(s). Nevertheless, some factors which cannot be influenced by study set-up are individual brood behaviour of the colonies and weather conditions.

Overall it may be assumed that factors may superordinate others, e.g. weather conditions in a respective period of the year which may superpose other factors, e.g. those described above. An analysis of all studies on the relation of colonies with BTR below and above 30% indicated that the

relations between the years completely changed in the course of the last four years (Table 7). Whereas in 2011 there were three times more replicates with BTRs ≤30% there were statistically more colonies with BTRs >30% in 2013 and 2014 (chi²-test, df = 1, p=0.032 and p<0.001, respectively).

Table 7 Level of BTRs in different years

BTR	Replicates° [n] in 2011	2012	2013	2014
>30%	14	26	30	32
≤30%	39	38	36	24
Relation ≤30 / >30	2.8	1.5 n.s.	1.2*	0.75*

° including all 62 studies; n.s. = not statistically significant different from distribution in 2011
* statistically significant different from distribution in 2011, chi²-test, df=1, p=0.032 (2013), p<0.001 (2014)

Another superordinate factor may be the housing of the bees in tunnels during the pre-exposure and exposure period. In contrast to the data of the OECD GD 75[1] studies, BTRs of detailed brood assessments at Oomen tests with free-flying bee colonies (Lückmann & Schmitzer 2014[9]) were lower and more reliable indicating a 'caging effect' in the tunnel studies (Table 8).

Table 8 Comparison of OECD GD 75[1] and acute/chronic Oomen feeding studies (Lückmann & Schmitzer 2014[9])

	OECD GD 75 (data ≥ 2011) n=208° (n=239``)	Oomen, acute feeding n=65	Oomen, chronic feeding n=27
Mean BTR	29.2 (32.9)	21.3	14.7
SD*	21.6 (24.4)	17.7	13.4
% of replicates with BTR ≤30%	61.5 (55.6)	75.4	85.2

n=number of replicates (colonies), * calculated from all replicates; ° 8 studies excluded; `` all studies

3.1.7. Results of bee brood studies from other EU countries and US

The number of available studies from other EU countries and the US was very low, i.e. 62 studies from Germany and Switzerland were available compared to a total of 13 studies from France, Spain and the US. The results from these countries displayed higher BTRs in the control compared to the data from Germany/Switzerland. Although affirming the limits of the test method, the low number of studies do not allow more than a very limited interpretation.

Therefore more data sets are needed to draw sound conclusions about the suitability and limitations of the test method in these countries.

4. Discussion and conclusions

The evaluation of bee brood studies performed between 2011 and 2014 shows that the BTRs in the controls improved only very little compared to the older 'historical' data (Pistorius *et al.* 2012[2]) and to data of the bee brood ring-test in 2002 (Schur et al. 2003[3]). Thus the suggested recommendations did not result in distinctly lower BTRs and reduced variability, as it was expected from the results in 2011. The improved results in 2011 might be due to better weather conditions during the testing season compared with later years.

On the one hand, approximately 38% of the replicates in the controls had BTRs >30% and every 2nd study had two or more replicates with BTRs >30%. On the other hand, the proportion of replicates (colonies) with BTRs ≤40% went up to 76.9% compared to 68.3% in the years before. And only 21.7% of the studies had BTRs >40% in two or more replicates. Consequently, these high BTR levels confound the interpretation of results of the PPP test items regarding potential brood effects with the outcome that several studies have to be regarded as invalid or are terminated before study finalization. From a regulatory perspective, such trials need to be repeated until

sufficient interpretability is achieved. Moreover, the reliability of the test method should be questioned. The envisaged quality criterion of BTRs <30% might be too stringent for a semi-field test system, considering the multiple influences and the discussion about an overall failure rate of 30% in the in vitro larvae trial. On the other hand, it is questionable if the data with BTRs ≤40% are reliable enough for a test system.

The reasons for the variability of this test method remain unclear now and further research is needed to overcome this variability in confined semi-field conditions. Superordinating factors may be weather conditions and a 'caging effect' superposing other (unknown, not yet considered factors, e.g. timing, kind and preparation of the colonies) factors which make it necessary:

a) to complement the existing data

b) to provide not submitted data of studies ≥ 2011 (incl. of terminated studies)

c) to compile information about the preparation of the colonies (e.g. time between preparation of colonies and BFD0, kind and extent of modifications of colonies)

d) to evaluate the additional data, and

e) to analyse the data in more detail, e.g. with multifactorial analyses.

Moreover, it is necessary to broaden the data base for studies outside Germany/Switzerland and therefore companies are asked to provide their full data sets for evaluation. Based on a more comprehensive data base further clarification might be possible.

These limitations are acknowledged. Nevertheless, the method is currently the only possibility to investigate potential effects of a plant protection product on larval development of honeybee brood in semi-field conditions covering exposure to pollen and nectar. It is assumed that problems are not related to the method *per se*, but to confined conditions. In contrast the Oomen method provides an artificial and worst case acute or chronic oral exposure scenario with feeding sugar solution inside the hive (see Lückmann & Schmitzer 2014[9]), which may be considered suitable to address certain risks of a test item; however, as bees are free flying, pollen foraged by bees is not contaminated.

Based on the currently available data there will be currently no attempt to develop the OECD GD 75[1] to an OECD Guideline. Moreover it has to be discussed in the near future:

a) whether it is reasonable to conduct the detailed brood investigation according to the acute/chronic Oomen feeding method (Lückmann & Schmitzer 2014[9]) or the OECD GD 75[1] under field conditions (e.g. Giffard & Huart 2014[10])

b) the need for trigger values resp. validity criteria for BTRs (e.g. < 30%) as discussed earlier.

Overall the results discussed here underline that the test method as described an OECD Guidance document in 2007 cannot be considerably improved now. But although the reliability of the method and a reliable interpretation regarding potential effects of a PPP on bee brood appears to be limited, it currently remains the only available test method using small bee colonies to address the potential risk of a plant protection products on honeybee larval development under realistic worst case (semi-field) conditions of exposure to pollen and nectar.

Acknowledgements

Many thanks to all the contract labs (BioChem agrar, Eurofins Agroscience, IBACON, IES & RIFCON) and companies (BASF SE, BayerCropscience, Dow AgroSciences, E. I. duPont de Nemours and Company, Syngenta) for providing their data on OECD GD 75 testing.

References

1. OECD Guidance Document No. 75, Guidance document on the honey bee (*Apis mellifera* L.) brood test under semi-field conditions. *Series of testing and assessment, Number 75. ENV/JM/MONO*: 223-27 (2007).

2. Pistorius J, Becker R, Lückmann J, Schur A, Barth M. Jeker L, Schmitzer S and Von der Ohe W, Effectiveness of method improvements to reduce variability of brood termination rate in honey bee brood studies under semi-field conditions. In

Hazards of pesticides to bees, 11th Internat Symp ICP-BR, Wageningen, Netherlands 2011, ed. by Oomen PA and Thompson H, *Julius-Kühn-Archiv*, Berlin 437, pp 115-120 (2012).

3. Schur A, Tornier I, Brasse D, Mühlen W, von der Ohe W, Wallner K and Wehling M, Honey bee brood ring-test in 2002: method for the assessment of side-effects of plant protection products on the honey bee brood under semi-field conditions. *B Insectol* **56**(1): 91–96 (2003).
4. EFSA, EFSA Guidance Document on the risk assessment of plant production products on bees (*Apis mellifera*, *Bombus* spp. and solitary bees) (published on July 04, 2013). *EFSA Journal* **11**(7): 3295 (2013).
5. Oomen PA, De Ruijter A and Van der Steen J, Method for honeybee brood feeding tests with insect growth-regulating insecticides. *Bull OEPP/EPPO Bull* **22**: 613–616 (1992).
6. Becker R & Lückmann J, Variability of brood termination rate in honeybee brood studies under semi-field conditions. Poster on 58th Jahrestagung der Arbeitsgemeinschaft der Institute für Bienenforschung e.V., Berlin-Dahlem 29.- 31.3.2011.
7. Pistorius J, unpubl., AG Bienenschutz: Protocol of the meeting at Julius Kuehn- Institute (JKI) on February 23rd, 2011.
8. Hecht-Rost S, Alscher A, Claßen C, Klockner A, Schlotz T, Staffel J and Lückmann J, Single versus double field rate: Do different rates of fenoxycarb in chronic *Oomen* bee brood feeding tests cause different effects sizes? - Poster presentation at the 12th ICPPR conference in Ghent, Belgium, 15-17 September, 2014.
9. Lückmann J & Schmitzer S, The effects of fenoxycarb in a chronic Oomen feeding test – Results of a ring-test. Oral presentation at the 12th ICPPR conference in Ghent, Belgium, 15-17 September, 2014.
10. Giffard H & Huart G, Assessment of brood development in field tests. Oral presentation at the 12th ICPPR conference in Ghent, Belgium, 15-17 September, 2014.