CONFIDENTIAL UNTIL PUBLISHED

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Tolvaptan for treating autosomal dominant polycystic kidney disease

Final response to company's response to ACD and critique of new evidence submitted by the company

02 July 2015

In this document, the Evidence Review Group (ERG) discusses the company's response[1] to the Appraisal Consultation Document (ACD)[2] as well as the new evidence submitted by the company[3]. The discussion will be presented in two parts, a discussion of the comments on the clinical effectiveness followed by a discussion of the comments of the cost effectiveness alongside a discussion of the new evidence submitted by the company.

Clinical effectiveness

The company proposed 'that the Committee focuses its discussions at the 2nd Committee meeting on a revised decision problem'[3]:

"*To appraise the clinical and cost effectiveness of tolvaptan within its licensed indication for treating autosomal dominant polycystic kidney disease,* ***specifically in patients at CKD stage 2 or 3 at initiation with evidence of rapidly progressing disease***."

ERG comment: According to table B6 of the company's submission[4], CKD was classified based on renal function according to the CKD-EPI formula (CKD stage 1: GFR ≥ 90; stage 2: 60 ≤ GFR < 90; stage 3: 30 ≤ GFR < 60). Of the 1,445 participants of the TEMPO 3:4 trial, approx. two thirds were in CKD stages 2 or 3 at the baseline (628 of 961 in the tolvaptan arm, 308 of 484 in the placebo arm).

Should the committee consider to focus on this revised decision problem, two aspects should be considered:

1. It seems that the optimised subgroup proposed by the company can be identified in clinical practice. However, given that the revised decision problem could result in a narrower recommendation, the impact on the patients not included in the proposed optimised subgroup (i.e. patients in CKD stage 1) should be considered.
2. Although the original company's submission presented subgroup results for some of the outcomes, the results were largely based on the whole group (i.e. including patients in CKD stage 1). Therefore, additional information on the optimised subgroup would be required. For example, baseline characteristics for this population should be presented and detailed results for all endpoints (e.g. figures B5 to B10) for the optimised subgroup should be given.

Re 1. Has all of the relevant information been taken into account?

According to the company's response to the ACD[1], "Point 3.19 on page 13 incorrectly states that the results of the subgroup analysis by CKD stage were marked as commercial in confidence and could not be displayed. This is valuable information that should have been made available in the ACD to all commentators".

ERG comment: The original company's submission[4] presented results by CKD stage as *academic in confidence*, e.g. tables B12 and B13.

Re 2. Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?

In the response to the ACD[1], the company criticises the conclusion reached by the committee regarding the generalisability of the TEMPO 3:4 study.

ERG comment: The ERG considers the statement *"The Committee concluded that the generalisability of the trial results may be limited because of differences in the trial population compared with people with ADPKD seen in routine clinical practice."*[2] to be accurate.

The company's response to the ACD[1] also discusses the handling of missing data and presents results following a 'jump to placebo' approach as well as from 'tipping point analyses'.

ERG comment: The ERG has a number of comments on this section:

1. While the company reports the assumption for patients in the tolvaptan arm who withdrew from the study ('100% loss of efficacy of tolvaptan'), it is unclear what assumption, if any, was applied to participants who withdrew from the placebo arm.
2. More information should be included on the 'tipping point analysis', i.e. what software was used, which assumptions were applied and it would be good to include a reference supporting this.
3. The results of the analyses are discussed but only p-values are given. It would be good to include effect estimates and corresponding confidence intervals.

The company presents results for a post-hoc power calculation for rate for decline in renal function.[1]

ERG comment: It should be noted that there is an ongoing discussion of the usefulness of post-hoc calculations.[5]

The post-hoc analysis presented by the company included 1,086 participants "who had an actual month 36 eGFR observation". This seems not to be limited to the optimised subgroup proposed by the company, i.e. "patients at CKD stage 2 or 3 at initiation with evidence of rapidly progressing disease". Looking at table B6 of the original company's submission, 961 participants were included in the tolvaptan arm (with 628 in CKD stages 2 and 3) while 484 participants (308) were included for placebo so the ERG was unable to check the numbers used in the power calculation. Given the very high power reported (99% vs. a usually accepted standard of 80%), it is likely that even the optimised subgroup will be adequately powered in a post-hoc analysis. In order to be certain, the post-analysis should be re-rerun for the optimised subgroup.

The response to the ACD[1] states that "it is incorrect that the TEMPO 3:4 study measured TKV by the ellipsoid method".

ERG comment: It should be noted that the statement in the ERG report (cited below) does not directly refer to the TEMPO 3:4 study:

> Page 40 of the CS states that "*TKV measured by ellipsoid method correlates well with TKV measured by the stereology method, and this has been validated using data from the CRISP cohort.*[6]*"*.
>
> The ERG doubts whether TKV measurement by the ellipsoid method is a reliable approximation of the gold standard: the stereology method for ADPKD patients. The formula given in the CS would yield the exact volume if the kidney were an ellipsoid (=elongated/flattened sphere). It would still give a fair correlation if the kidney had a fixed shape (however elongated or flattened). However, in ADPKD the kidneys lose their predictable shape and become distorted. The cited work by Breau et al 2013[6] validated the method in 28 patients *"without congenital, cystic or neoplastic abnormalities"*, i.e. patients with predictable kidney shapes. If TKV change would be used as a measure of progression of disease, this potential unreliability might have serious consequences, e.g. a single cyst rapidly growing at the pole of a kidney (or rupturing and disappearing) would change the estimated kidney volume considerably.

ERG comment: The Company also made a number of comments regarding the committee's conclusion on hepatotoxicity. These are discussed in the following section on cost-effectiveness.

Cost effectiveness

This section will critically assess the changes the Company made to their original economic model, respond to the issues raised by the Company and present additional analyses including the committee's preferred scenario based on the appraisal consultation document.[2]

Analyses submitted by the Company

The Company presented an updated analysis.[3] The following adjustments were applied by the Company to the originally submitted model:

- Model code correction
- PAS discount of [redacted] on the Tolvaptan NHS list price
- Setting baseline patient profile to represent ADPKD patients with CKD stage 2 to 3
- Setting a subgroup specific relative reduction in renal function decline (29.7%)
- Estimate the "annual change in renal function" and "percentage TKV change" using the regression equations instead of using data observed over the trial period (for the first three year).
- CKD-EPI as approximation for eGFR
- Equal CKD-stage costs for CKD-stage 3 as for CKD-stage 4
- Disutility of 0.0025 for Tolvaptan treatment
- Disutility of 0.02 for haemodialysis and peritoneal dialysis complications
- Assuming treatment discontinuation of 2.9% after year 3
- Assuming decreased post-transplant costs
- Assuming increased mortality (hazard ratio: 1.6)
- Assuming increased monitoring costs for the first two years while assuming the same monitoring costs as in their base case for subsequent years

The analysis presented by the Company deviates from the scope, i.e. the Company proposes to change the decision problem, focusing on CKD stage 2 or 3 at initiation. To reflect this subgroup of patients in the model, the Company adjusted the baseline patient profiles (mean age increased from 39 to 44, proportion of females remained 48%, mean eGFR (ml/min/1.73m^2) decreased from 82 to 60 and mean TKV (ml) increased from 1692 to 2300) and adjusted the percentage relative reduction in renal function decline from 26.4% to 29.7%. The results of this analysis are presented in Table 1.

ERG comment: The Company did not adjust the "annual change in renal function" and "percentage TKV change" to reflect this subgroup. However, as the reduction in renal function will probably be higher and the percentage TKV change will probably be lower for this subgroup[7] (both decreasing the ICER), using data for the whole population will probably overestimate the ICER and can thus be speculated to be conservative. However, this speculation should ideally be verified by the Company. In addition, the ERG would have preferred the "annual change in renal function" and "percentage TKV change" to be informed by the actual trial data for the first three year instead of the regression equations (as in the

initial model submitted by the Company). Using the regression equations in the first three years without further justification is not conservative (given this decreases the ICER).

Besides the deviation from the scope, the analyses presented by the Company deviated from the Committee's preferred scenario (according to the ACD) regarding the following assumptions:

- Equal kidney pain probability for both arms
- Disutility of 0.0123 for Tolvaptan treatment
- Estimate the "annual change in renal function" and "percentage TKV change" using the data observed over the trial period instead of using regression equations (for the first three year). This was changed by the Company without justification.
- Assuming that 3 out of 10 Hy's Law cases would have a liver transplant at year 1 and would die immediately thereafter – exploratory analysis
- Assuming increased mortality (hazard ratio: 2.0) – exploratory analysis
- Assuming increased monitoring costs – exploratory analysis

Equal kidney pain probability for both arms
The Company does not agree with assuming equal kidney pain probability for both arms since in their view this contradicts the findings of the TEMPO 3:4 trial.[1] However, as argued in the ERG report,[8] the ERG believes this may have introduced a downward bias to the ICER, as it is assumed by the Company that the difference in kidney pain as observed in TEMPO 3:4 is independent from the effect of tolvaptan on disease progression (i.e. assuming the kidney pain probability is CKD-stage independent / constant for CKD-stages 1-4). This is questionable, as pain is a known symptom of chronic kidney disease, increasing with disease progression.[9] The separate modelling of pain may have led to a double counting (i.e. the higher utility for lower CKD-stages might already capture the effect of kidney pain). Therefore, the ERG would prefer a conservative approach which incorporates an equal kidney pain probability for both arms.

Disutility of 0.0123 for tolvaptan treatment
The Company does not agree with the disutility of 0.0123 for tolvaptan treatment applied by the ERG. The ERG conservatively incorporated this disutility since no other adverse events than kidney pain were incorporated in the model (ideally all relevant adverse events would be incorporated, making this disutility obsolete). Moreover, it should be noted that this disutility was part of the ERG preferred base case (not part of the exploratory analyses as mentioned by the Company in their response[1]). The Company argues that this disutility value is an overestimation and proposes a substantial lower disutility of 0.0025. This is the disutility value for tolvaptan treatment where the QALYs gained due to pain reduction are equal to the QALYs lost due to the negative effect of being on tolvaptan. It is however unclear to the ERG why the disutility due to tolvaptan should 'cancel out' that due to pain reduction (as argued by the Company[1]). Moreover, the ERG is not convinced that the conclusion by the Company *"that the net impact on patient HRQoL of tolvaptan AEs, relative to reduction in*

ADPKD complications, is no worse than neutral"[1] is justified given the evidence provided. Hence the ERG would prefer to use the conservative disutility of 0.0123 for being on tolvaptan treatment.

Incorporating Hy's Law cases

The Company stated that they believe it is not appropriate to incorporate Hy's Law cases in the model. The ERG disagrees with this since, as mentioned by the Company,[1] the possibility of future Hy's law cases cannot be eliminated. Therefore, the ERG believes this should, at least, be assessed in an exploratory (worst case scenario) analysis. Moreover, the Company raises potential issues concerning the methods that were used by the ERG for this analysis (though these were not raised as factual inaccuracies by the Company before the first appraisal committee meeting). Considering these issues, the ERG believes that the Company misinterprets this worst case analysis. Regarding the first and second issues raised by the Company, the Hy's law case analysis was an explorative worst case scenario analysis, using worst case assumptions as clearly described in the ERG report:

"a worst case scenario was adopted assuming that all Hy's Law cases would need a liver transplant at the end of year 1 and would die immediately thereafter (severe liver injury may expected in a frequency 1 out of 10 Hy's Law cases). In total, 958 patients were exposed in the Tempo 3:4 and Tempo 4:4 studies and three patients have been identified as Hy's Law cases. Therefore, it was assumed that 0.3% (=3/958) of the patients would need a liver transplant at the end of year 1 and would die directly thereafter. Moreover, it was assumed that these patients would have a utility of 0.77 (= total QALYs / total LYs), an annual costs of £17,714 (= total costs / total LYs) and total transplantation costs of £34,425.67."[8]

Similarly, considering the third issue raised by the Company, there was no coding error as the ERG purposely selected the total costs without PAS to calculate the annual costs for the patients with liver injury (0.3%) to ensure that these costs were not underestimated (the annual costs for patients with liver injury might be higher than for the average ADPKD patient). As mentioned above the ERG believes that the potential consequences of Hy's law cases should be explored, however the approach used by the ERG[8] should be considered as a worst case scenario and it is debatable whether this should be preferred for the base case analysis.

Increased mortality for ADPKD patients

The Company argues that the increased mortality for ADPKD patients applied by the ERG in an exploratory analysis (hazard ratio: 2.0) is too conservative. The ERG agrees that the hazard ratio proposed by the Company of 1.6 is probably a more plausible estimate. In addition, the hazard ratio of 2.0 should be considered in its context (i.e. a worst case analysis; the hazard ratio of 2.0 is equal to the upper bound of the 95% confidence interval of the hazard ratio reported in the paper by Florijn et al.[10]). Moreover, given the uncertainty of this hazard ratio (i.e. absence of reliable data concerning the mortality risk in ADPKD) it is

debatable whether a conservative hazard ratio (i.e. a hazard ratio of 2.0) might be preferred for the base case.

Increased monitoring costs

The Company agreed with the increased monitoring costs applied by the ERG, except for the additional monitoring costs after the second year. Given the ALT elevation after 18 months for tolvaptan versus placebo (Figure 1 in the Company's response[1]), the ERG considers it likely that the scenario proposed by the Company is more plausible than the increased monitoring costs scenario proposed by the ERG in its exploratory analysis.

Conclusion

As described in the previous section, the ERG would regard incorporating equal kidney pain probabilities for both arms and a disutility of 0.0123 for Tolvaptan treatment more conservative and hence preferable. Moreover, the ERG would have preferred the "annual change in renal function" and "percentage TKV change" to be informed by the actual trial data for the first three year instead of the regression equations (which was selected by the Company without justification). Regarding the increased mortality and increased monitoring costs scenario analyses, the scenarios proposed by the Company might be more plausible than the exploratory (worst case) analyses presented by the ERG. Similarly, it can be debated whether the Hy's Law cases exploratory (worst case) analysis presented by the ERG should be preferred in the base case.

Additional analyses performed by the Company and the ERG

The following section reports on the scenario preferred by the Committee (according the ACD). Secondly, the ERG adapted this Committee preferred scenario, taking account of the new evidence submitted by the Company (i.e. excluding the Hy's law case scenario and using the increased mortality and increased monitoring costs scenario analyses proposed by the Company). Thirdly, the subgroup analysis presented by the Company is shown. Fourthly, the ERG adapted this subgroup analysis to include the assumptions preferred by the Committee. Finally, an adapted version of this latter scenario is presented taking account of the new evidence submitted by the Company. It should be noted that this last scenario is similar to the subgroup analysis presented by the Company. These analyses differ with respect to:

- Equal kidney pain probability for both arms
- Disutility of 0.0123 for tolvaptan treatment
- Estimate the "annual change in renal function" and "percentage TKV change" using the data observed over the trial period instead of using regression equations (for the first three year).

Table 1: Additional analyses performed by the Company and the ERG

	Total			Incremental			
	Costs (£)	LY	QALY	Costs (£)	LY	QALY	ICER (€/QALY)
Population in initial submission (CKD stages 1 to 3 at initiation)							
1. ERG's analysis of Committee's preferred scenario[a]							
Standard treatment	[redacted]	16.41	12.45				
Tolvaptan	[redacted]	16.87	12.88	[redacted]	0.46	0.43	50,010
2.ERG's adaptation of Committee's preferred scenario[b]							
Standard treatment	[redacted]	16.56	12.55				
Tolvaptan	[redacted]	17.10	13.03	[redacted]	0.54	0.49	43,514
Subgroup proposed by the Company (CKD stages 2 and 3 at initiation)							
3. Company's analysis[c]							
Standard treatment	[redacted]	14.13	9.83				
Tolvaptan	[redacted]	14.71	10.42	[redacted]	0.58	0.59	23,503
4. ERG's analysis of Committee's preferred scenario[d]							
Standard treatment	[redacted]	14.18	9.93				
Tolvaptan	[redacted]	14.66	10.39	[redacted]	0.48	0.45	31,438
5. ERG's adaptation of Committee's preferred scenario[e]							
Standard treatment	[redacted]	14.33	10.03				
Tolvaptan	[redacted]	14.87	10.52	[redacted]	0.54	0.49	30,025

Abbreviations: ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; LY = life year(s); QALYs = quality-adjusted life years

[a] The following adjustments were made to the initial model submitted by the Company:

1. Model code correction
2. PAS discount of [redacted] on the Tolvaptan NHS list price
3. CKD-EPI as approximation for eGFR
4. Equal kidney pain probability for both arms
5. Equal CKD-stage costs for CKD-stage 3 as for CKD-stage 4
6. Disutility of 0.0123 for Tolvaptan treatment
7. Disutility of 0.02 for haemodialysis and peritoneal dialysis complications
8. Assuming treatment discontinuation of 2.9% after year 3
9. Assuming decreased post-transplant costs
10. Assuming increased mortality (hazard ratio: 2.0)
11. Assuming that 3 out of 10 Hy's Law cases would have a liver transplant at year 1 and would die immediately thereafter
12. Assuming increased monitoring costs

[b] In addition to adjustments 1 to 9 mentioned under [a], the increased mortality and increased monitoring costs scenario analyses proposed by the Company were applied

[c] In addition to adjustments 1 to 3, 5 and 7 to 9 mentioned under [a], the baseline patient profile and relative reduction in renal function decline were adjusted to represent ADPKD patients with CKD stage 2 to 3 and the increased mortality and increased monitoring costs scenario analyses proposed by the Company were applied. Moreover, the "annual change in renal function" and "percentage TKV change" were estimated using regression equations (for the first three year) instead of the data observed over the trial period.

[d] In addition to all adjustments mentioned under [a], the baseline patient profile and relative reduction in renal function decline were adjusted to represent ADPKD patients with CKD stage 2 to 3.

[e] In addition to adjustments 1 to 9 mentioned under [a], the baseline patient profile and relative reduction in renal function decline were adjusted to represent ADPKD patients with CKD stage 2 to 3 and the increased mortality and increased monitoring costs scenario analyses proposed by the Company were applied.

References

[1] Otsuka Pharmaceuticals (UK) Ltd. Response to the Appraisal Consultation Document relating to tolvaptan for the treatment of autosomal dominant polycystic kidney disease [ID652]. Otsuka Pharmaceuticals (UK) Ltd., 2015: 21.

[2] National Institute for Health and Care Excellence. *Appraisal consultation document - Tolvaptan for treating autosomal dominant polycystic kidney disease*. London: National Institute for Health and Care Excellence, 2015. 50p. Available from: http://www.nice.org.uk/guidance/gid-tag447/documents/kidney-disease-autosomal-dominant-polycystic-tolvaptan-id652-final-scope2

[3] Otsuka Pharmaceuticals (UK) Ltd. New evidence in response to the preliminary recommendations in the ACD of Tolvaptan for treating autosomal dominant polycystic kidney disease [ID652]. Otsuka Pharmaceuticals (UK) Ltd., 2015: 17.

[4] Otsuka Pharmaceuticals (UK) Ltd. Tolvaptan for the treatment of autosomnal dominant polycystic kidney disease: Submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA). Otsuka Pharmaceuticals UK Ltd., 2014: 293.

[5] UCLA: Institute for Digital Research and Education. Statistical Computing Seminars Introduction to Power Analysis [Internet]. [accessed 2.7.15]. Available from: http://www.ats.ucla.edu/stat/seminars/Intro_power/

[6] Breau RH, Clark E, Bruner B, Cervini P, Atwell T, Knoll G, et al. A simple method to estimate renal volume from computed tomography. *Can Urol Assoc J* 2013;7(5-6):189-92.

[7] Torres VE, Chapman AB, Devuyst O, Gansevoort RT, Grantham JJ, Higashihara E, et al. Tolvaptan in patients with autosomal dominant polycystic kidney disease. *N Engl J Med* 2012;367(25):2407-2418.

[8] Wolff R, Joore MA, Ramaekers B, Tomini F, Pouwels X, Deshpande S, et al. *Tolvaptan for treating autosomal dominant polycystic kidney disease: a Single Technology Appraisal*. York: Kleijnen Systematic Reviews Ltd., 2015

[9] Pham PC, Dewar K, Hashmi S, Toscano E, Pham PM, Pham PA, et al. Pain prevalence in patients with chronic kidney disease. *Clin Nephrol* 2010;73(4):294-9.

[10] Florijn KW, Noteboom WM, van Saase JL, Chang PC, Breuning MH, Vandenbroucke JP. A century of mortality in five large families with polycystic kidney disease. *Am J Kidney Dis* 1995;25(3):370-4.