# ONLINE SUPPLEMENTAL METHODS

## Selection of attributes and levels

The aim of the literature review was to determine the key attributes and levels to be included in the DCE. This involved both a targeted literature review and a product label review.

The targeted literature review involved separate searches in major medical literature databases (Embase and MEDLINE) (**Online Supplemental Table 1**); a search for qualitative studies that considered the patient perspective on AD treatments; and a search for patient preference-specific studies, which considered AD treatments. Once key themes within the literature review were identified, the attributes were classified into corresponding categories. All abstracts were double screened, and disagreements were reviewed by a senior member of the research team.

The search strategy for the qualitative studies focused on studies that conducted interviews or focus groups, which mentioned AD or eczema, and their available treatments, as well as quality of life or patient preferences. The search excluded any non-adult studies, animal studies, clinical trials, and editorial notes. The search strategy for patient preference-specific studies sought studies that were explicitly patient preference in design, such as those utilising DCEs. Additionally, the studies had to mention AD or eczema.

The targeted literature search identified 33 potential studies. No duplicates were found, and all 33 were screened for eligibility. The abstracts were screened sequentially by two reviewers, and a third reviewer compared the rationale for inclusion and exclusion of studies to obtain the final list of full texts to screen. Seven studies were excluded because they did not involve adult patients, 13 because they

weren't about AD, six because they did not have the study design of interest, and four because no full text was available. The remaining three studies included one quantitative[1] and two qualitative studies.[2, 3] In the quantitative study, the most important treatment attribute was the appearance of eczema (dryness/flakiness). In the two qualitative studies, itch reduction (symptom control), monitoring of symptoms, flexibility of treatment regimens to control flares, appearance (dryness/flakiness), and skin pain were identified themes.

Additionally, a product label search was conducted. Ten product labels for medications indicated for use in AD were reviewed in detail, including baricitinib (Olumiant®), dupilumab (Dupixent®), clobetasol propionate (Clobex®), tacrolimus (Protopic®), prednisone (Rayos®), cyclosporin (Neoral®), methotrexate, azathioprine (Imuran®), mycophenolate mofetil (CellCept®), and phototherapy. Itch reduction was most commonly reported as the percentage of patients achieving a meaningful (≥4-point reduction in the itch numerical rating scale) reduction in itch at week 16. Skin appearance was most commonly measured by the proportion of patients achieving clear or almost clear skin at week 16 (Investigator's Global Assessment scores of 0 or 1). The review of product labels also identified conjunctivitis as a differentiating and common side-effect of dupilumab that is not associated with other systemic therapies. Risk of serious infections were associated with other treatments, such as baricitinib and cyclosporine. The product label review also highlighted different modes and frequency of administration for systemic treatments, which included daily oral medication or subcutaneous administration every 2 weeks. Monitoring was also required for baricitinib and cyclosporine, but not for dupilumab.

**Model specification**

The analysis of all DCE responses followed random utility theory.[4-6] The model assumes that each respondent (*n*) chooses the alternative (*j*) in every DCE question (*t*) that results in the highest utility (a measure of desirability) of all available alternatives. Utility in a random utility model is defined as:

$$u(\boldsymbol{x}_{jnt}) = v(\boldsymbol{x}_{jnt}) + \varepsilon_{jnt}$$

Here the systematic utility component $v(\boldsymbol{x}_{jnt})$ is a function of the DCE attributes and $\varepsilon_{jnt}$ is a type 1 extreme value distributed random error. Two models are presented: a dummy-coded MNL model and an MNL model with skin appearance coded linearly, which is required to estimate the maximum acceptable decrease (MAD) in the probability of achieving clear or almost clear skin at week 16. For the former, the utility function was defined as:

$$\begin{aligned} u_{jnt} = {} & \alpha_{\text{Treatment A}} + \alpha_{\text{Old Treatment}} + \beta_1 40\%\_\text{itch\_reduction}_{jnt} \\ & + \beta_2 50\%\_\text{itch\_reduction}_{jnt} + \beta_3 20\%\_\text{skin\_appearance}_{jnt} \\ & + \beta_4 40\%\_\text{skin\_appearance}_{jnt} + \beta_5 10\%\_\text{eye\_inflammation}_{jnt} \\ & + \beta_6 0\%\_\text{eye\_inflammation}_{jnt} + \beta_7 3\%\_\text{serious\_infections}_{jnt} \\ & + \beta_8 0\%\_\text{serious\_infections}_{jnt} + \beta_9 1\_\text{week\_onset}_{jnt} + \beta_{10} 2\_\text{days\_onset}_{jnt} \\ & + \beta_{11} \text{flare\_management}_{jnt} + \beta_{12} \text{long\_term\_no}_{jnt} \\ & + \beta_{13} \text{long\_term\_yes\_pauses}_{jnt} + \beta_{14} \text{oral\_admin}_{jnt} + \beta_{15} \text{no\_check\_ups}_{jnt} \\ & + \beta_{16} \text{occassional\_check\_ups}_{jnt} + \varepsilon_{jnt} \end{aligned}$$

The constants $\alpha_{\text{Treatment A}} + \alpha_{\text{Old Treatment}}$ controlled for potential bias to select the left option (Treatment A), and the Old Treatment, $\beta_1$to $\beta_{16}$ were the estimated marginal utilities (i.e., estimated preference parameters), $\varepsilon_{jnt}$ was an extreme value type I distributed error that allowed the function to be estimated in a logit model.[6] All attributes were dummy-coded. The reference level was the assumed worst-case option. Each of the estimated marginal utilities measured respondents' sensitivity to deviations from the reference level of the corresponding attribute. The sign (+ or –) of

a marginal utility denotes whether patients valued this deviation positively or negatively. Only the initial choices (A vs. B vs. old treatment) were considered for the analysis of preferences. The initial and follow-up choices can be combined to allow for a more precise measurement of preferences. However, it is appropriate to combine these two types of choices only when they generate approximately the same information about participants' preferences. This condition was verified in two ways. Two MNL models were separately estimates for the initial (4,848 observations) and follow-up choices (1,126 observations), and then their preference estimates were compared. The Pearson correlation coefficient between the two sets of estimates was relatively low (0.32) as was the coefficient of determination for the linear regression (0.104), indicating poor agreement between the sets of estimates. A third MNL model was estimated on the combined initial and follow-up choices (5,974 observations), and its statistical performance was compared with the MNL model based on initial choices only. The adjusted McFadden pseudo-$R^2$ was lower for the model based on combined choices (7.3%) than for the initial model (8.3%), indicating that combining the initial and follow-up choices had a detrimental effect on the explanatory power of the model.

The linear coding of skin appearance was required to derive meaningful MAD measures. This measure was obtained by estimating the baseline utility function with skin appearance being coded as linear (i.e., one marginal utility is estimated instead of $\beta_3$ and $\beta_4$ for skin appearance). The utility function was defined as:

$$
\begin{aligned}
u_{jnt} = {} & \alpha_{\text{Treatment A}} + \alpha_{\text{Old Treatment}} + \beta_1 40\%\_\text{itch\_reduction}_{jnt} \\
& + \beta_2 50\%\_\text{itch\_reduction}_{jnt} + \beta_3\_\text{skin\_appearance}_{jnt} \\
& + \beta_4 10\%\_\text{eye\_inflammation}_{jnt} + \beta_5 0\%\_\text{eye\_inflammation}_{jnt} \\
& + \beta_6 3\%\_\text{serious\_infections}_{jnt} + \beta_7 0\%\_\text{serious\_infections}_{jnt} \\
& + \beta_8 1\_\text{week\_onset}_{jnt} + \beta_9 2\_\text{days\_onset}_{jnt} + \beta_{10}\text{flare\_management}_{jnt} \\
& + \beta_{11}\text{long\_term\_no}_{jnt} + \beta_{12}\text{long\_term\_yes\_pauses}_{jnt} + \beta_{13}\text{oral\_admin}_{jnt} \\
& + \beta_{14}\text{no}_{\text{check}_{\text{ups}}\,jnt} + \beta_{15}\text{occassional\_check\_ups}_{jnt} + \varepsilon_{jnt}
\end{aligned}
$$

Each marginal utility was then divided by the marginal utility for skin appearance:

$$\text{MAD}_k = \frac{\hat{\beta}_k}{\hat{\beta}_3}$$

Such linear encoding is based on the underlying assumption of linearity in preferences, wherein a one-unit change in the attribute has a constant effect on respondents' choices and does not depend on the absolute value of the attribute level (e.g., a one-unit change from 15% to 16% increase in the chance of achieving clear or almost clear skin at week 16 has the same effect as the change from 20% to 21%). The validity of the assumption of linearity in preferences was tested by analysing the trend in risk estimates from the dummy-coded MNL model. Estimates were obtained for every attribute level in the dummy-coded MNL model (i.e., 3 levels for skin appearance). The linearity of skin appearance was tested by fitting a linear regression and evaluating its coefficient of determination. The assumption of linearity in skin appearance was accepted with a coefficient of 0.81, which exceeds the threshold of 0.7 to verify linearity.

**Combination of choice data from different countries**

We estimated a heteroscedastic MNL (HMNL) model allowing for scale differences between countries. The likelihood ratio test (LRT) indicated that this model performed significantly better than the standard MNL model (D=11.45, P=0.003). We also estimated an extended version of the MNL model allowing for interaction effects

between country of residence and the attributes' levels. This interacted MNL (IMNL) model also significantly outperformed the standard MNL model (D=66.44, P=0.001).

Using the scale estimates from the HMNL model, we applied a scale correction to the dataset and then re-estimated the IMNL model (RIMNL) to determine whether the interaction effects found to be significant in the initial IMNL model would remain significant after accounting for potential scale differences between countries. This was the case, indicating that differences in choice behaviours between countries could not be fully explained as the consequence of a change in underlying utility scale (**Online Supplemental Table 4**).

**Accounting for unobserved heterogeneity in preferences**

We estimated an MXL model allowing all parameters to be independently and normally distributed (i.e., diagonal covariance matrix of random effects). The MXL model significantly outperformed its MNL counterpart (LRT: D=678.39, P<0.001), but a comparison of estimates between the two models showed a high level of agreement (**Online Supplemental Figure 1**). We fitted a linear regression line through the set of coordinates (MNL; MXL) and the coefficient of determination was close to 100%. The intercept, which can be interpreted as a measure of bias associated with use of MNL estimates instead of MXL ones, was close to zero (0.012) and non-significant (P=0.462). However, the slope (1.172), which can be interpreted as a measure of scale, was significantly different from 1 (P<0.001), indicating that the MXL model measured the same preference effects but on a higher (more precise) utility scale. Given the research objectives of our study were to quantify trade-offs between attributes, and more specifically the MAD in the probability of achieving clear/almost clear skin at week 16, this change in utility scaling was deemed irrelevant.

**Independence of treatment options relative to old treatment**

A nested logit (NL) model was estimated to allow for a repartition of the choice options in two different nests: treatments A and B in a “New treatment” nest and the opt-out option in an “Old treatment” nest. The inclusive value (IV) parameter, which captures the degree of correlation in unobserved factors over alternatives within the "New treatment" nest, was significant (P=0.003) and implied a weak-to-moderate correlation (1-0.63=0.37). The LRT indicated that the NL model significantly outperformed the MNL model (D=8.09, P=0.004). However, a comparison of estimated effects between the two models showed a high level of agreement ($r^2$>99%) and the intercept of the linear regression line was null (**Online Supplemental Figure 2**).

**References**

1. Hauber AB, Mohamed AF, Gonzalez JM, et al. Benefit-risk tradeoff preferences for chronic hand eczema treatments. *Journal of Dermatological Treatment* 2016;28:40-46.
2. Howells LM, Chalmers JR, Cowdell F, et al. ‘When it goes back to my normal I suppose’: a qualitative study using online focus groups to explore perceptions of ‘control’ among people with eczema and parents of children with eczema in the UK. *BMJ Open* 2017;7:e017731.
3. DiBenedetti D, Baranowski E. Assessing patient and physician experiences with severe chronic hand eczema. *J Am Acad Dermatol* 2015;72:AB77.
4. Marschak J. Binary choice constraints on random utility indicators Mathematical methods in the social sciences. Mathematical methods in the social sciences. Palo Alto, CA, USA: Stanford University Press 1959:312-29.
5. Manski CF. The structure of random utility models. *Theory and Decision* 1977;8:229-54.
6. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, ed. Frontiers in econometrics. New York: Academic Press 1973:105-42.

# ONLINE SUPPLEMENTAL TABLES

## Online Supplemental Table 1. Targeted literature review search terms

| No. | Query | Results | Date |
|---|---|---|---|
| #6 | #1 AND (#2 AND #3 OR (#4 AND #5)) | 33 | 10-Sep-18 |
| #5 | ((('qualitative research'/exp OR 'nursing methodology research'/exp OR ethnograph*:ti,ab OR lived) AND experience*:ti,ab OR narrative) AND analysis:ti,ab OR grounded) AND interview*:ti,ab OR themes:ab,ti | 80104 | 10-Sep-18 |
| #4 | 'treatment attribute*':ab,ti OR 'attributes':ab,ti OR 'preference*' OR 'trade off':ab,ti OR value:ab,ti OR 'patient decision making':ab,ti OR 'treatment satisfaction':ab,ti OR 'patient experience':ab,ti OR perception*:ab,ti OR attitude*:ab,ti OR 'patient preference':ab,ti | 1743076 | 10-Sep-18 |
| #3 | 'quantitative study'/exp OR 'discrete choice' OR 'dce':ab,ti OR 'discrete choice experiment*':ab,ti OR 'choice experiment*':ab,ti OR 'conjoint':ab,ti OR 'conjoint analysis':ab,ti OR 'bws':ab,ti OR 'benefit risk':ab,ti OR 'thresholding':ab,ti OR 'multiple criteria decision analysis':ab,ti OR 'benefit-risk':ab,ti OR 'tradeoff':ab,ti OR 'best-worst scaling':ab,ti OR 'ahp':ab,ti OR 'analytic hierarchy':ab,ti OR 'swing weighting':ab,ti OR 'threshold technique':ab,ti OR 'risk benefit analysis':ab,ti | 68917 | 10-Sep-18 |
| #2 | 'treatment attribute*':ab,ti OR 'attributes':ab,ti OR 'preference*' OR 'trade off':ab,ti OR value:ab,ti OR 'patient decision making':ab,ti OR 'treatment satisfaction':ab,ti OR 'patient experience':ab,ti OR perception*:ab,ti OR attitude*:ab,ti OR 'patient preference':ab,ti | 1370306 | 10-Sep-18 |
| #1 | 'eczema'/exp OR 'atopic dermatitis'/exp | 61560 | 10-Sep-18 |

**Online Supplemental Table 2. Comparison of results across models**

| | | MLE (SE) | | | |
|---|---|---|---|---|---|
| Attributes and levels | Sample | MNL | HMNL | IMNL | RIMNL |
| 1. Preferences | | | | | |
| Alternative Specific Constant | | | | | |
| Old treatment | Overall | 1.458 (0.115)*** | 1.643 (0.139)*** | 1.392 (0.200)*** | 1.392 (0.182)*** |
| Option A | Overall | -0.038 (0.037) | -0.042 (0.042) | -0.007 (0.061) | -0.007 (0.062) |
| Itch Reduction | | | | | |
| 2 out of 10 (20%) | Overall | Reference | - | - | - |
| 4 out of 10 (40%) | Overall | 0.590 (0.060)*** | 0.671 (0.073)*** | 0.651 (0.101)*** | 0.651 (0.098)*** |
| 5 out of 10 (50%) | Overall | 0.760 (0.058)*** | 0.858 (0.072)*** | 0.733 (0.100)*** | 0.733 (0.095)*** |
| Skin Appearance | | | | | |
| 2 out of 10 (20%) | Overall | 0.214 (0.058)*** | 0.246 (0.066)*** | 0.243 (0.098)* | 0.243 (0.096)* |
| 4 out of 10 (40%) | Overall | 0.481 (0.061)*** | 0.554 (0.072)*** | 0.606 (0.105)*** | 0.607 (0.100)*** |
| 1 out of 10 (10%) | Overall | Reference | - | - | - |
| Eye inflammation | | | | | |
| 20 out of 100 (20%) | Overall | Reference | - | - | - |
| 10 out of 100 (10%) | Overall | 0.273 (0.048)*** | 0.317 (0.056)*** | 0.398 (0.080)*** | 0.398 (0.079)*** |
| 0 out of 100 (0%) | Overall | 0.637 (0.056)*** | 0.723 (0.068)*** | 0.676 (0.092)*** | 0.677 (0.092)*** |
| Serious Infections | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 0 out of 100 (0%) | Overall | 0.722 (0.056)*** | 0.800 (0.067)*** | 0.522 (0.093)*** | 0.523 (0.093)*** |
| 6 out of 100 (6%) | Overall | Reference | - | - | - |
| 3 out of 100 (3%) | Overall | 0.306 (0.050)*** | 0.339 (0.057)*** | 0.197 (0.083)* | 0.197 (0.082)* |
| Speed of Onset | | | | | |
| 2 weeks | Overall | Reference | - | - | - |
| 1 week | Overall | 0.010 (0.052) | 0.011 (0.059) | 0.019 (0.088) | 0.019 (0.086) |
| 2 days | Overall | 0.178 (0.049)*** | 0.205 (0.057)*** | 0.217 (0.083)** | 0.217 (0.082)** |
| Flare Management | | | | | |
| No | Overall | Reference | - | - | - |
| Yes | Overall | 0.090 (0.039)* | 0.109 (0.045)* | 0.161 (0.065)* | 0.161 (0.064)* |
| Long-term Disease Management | | | | | |
| Yes, without the possibility for pauses | Overall | Reference | - | - | - |
| Should not be used long-term | Overall | 0.057 (0.054) | 0.056 (0.062) | -0.012 (0.093) | -0.012 (0.091) |
| Yes, with the possibility for pauses | Overall | 0.360 (0.048)*** | 0.399 (0.056)*** | 0.297 (0.080)*** | 0.297 (0.079)*** |
| Administration | | | | | |
| Injection under the skin, every two weeks | Overall | Reference | - | - | - |
| Oral pill, once or twice daily | Overall | 0.253 (0.047)*** | 0.294 (0.055)*** | 0.322 (0.078)*** | 0.322 (0.079)*** |
| Check-ups | | | | | |
| Frequent check-ups required | Overall | Reference | - | - | - |

| | | | | | |
|---|---|---|---|---|---|
| Occasional check-ups required | Overall | 0.242 (0.054)*** | 0.286 (0.063)*** | 0.328 (0.090)*** | 0.328 (0.091)*** |
| No check-ups required | Overall | 0.312 (0.052)*** | 0.366 (0.061)*** | 0.417 (0.086)*** | 0.417 (0.086)*** |
| 2. Interaction effects | | | | | |
| Alternative Specific Constant | | | | | |
| Old treatment | France | - | - | 0.118 (0.311) | 0.358 (0.257) |
| Old treatment | Spain | - | - | 0.104 (0.336) | 0.586 (0.298)* |
| Option A | France | - | - | -0.066 (0.094) | -0.077 (0.103) |
| Option A | Spain | - | - | -0.035 (0.089) | -0.048 (0.105) |
| Itch Reduction | | | | | |
| 4 out of 10 (40%) | France | - | - | -0.150 (0.156) | -0.069 (0.154) |
| 4 out of 10 (40%) | Spain | - | - | -0.057 (0.153) | 0.134 (0.163) |
| 5 out of 10 (50%) | France | - | - | 0.066 (0.155) | 0.194 (0.151) |
| 5 out of 10 (50%) | Spain | - | - | 0.024 (0.151) | 0.268 (0.159) |
| Skin Appearance | | | | | |
| 2 out of 10 (20%) | France | - | - | 0.029 (0.149) | 0.072 (0.155) |
| 2 out of 10 (20%) | Spain | - | - | -0.099 (0.143) | -0.053 (0.156) |
| 4 out of 10 (40%) | France | - | - | -0.200 (0.162) | -0.135 (0.157) |
| 4 out of 10 (40%) | Spain | - | - | -0.194 (0.162) | -0.062 (0.165) |
| Eye inflammation | | | | | |
| 10 out of 100 (10%) | France | - | - | -0.272 (0.121)* | -0.252 (0.132) |

| | | | | | |
|---|---|---|---|---|---|
| 10 out of 100 (10%) | Spain | - | - | -0.127 (0.114) | -0.040 (0.133) |
| 0 out of 100 (0%) | France | - | - | -0.086 (0.140) | 0.007 (0.153) |
| 0 out of 100 (0%) | Spain | - | - | -0.029 (0.132) | 0.179 (0.154) |
| Serious Infections | | | | | |
| 0 out of 100 (0%) | France | - | - | 0.343 (0.142)* | 0.480 (0.152)** |
| 0 out of 100 (0%) | Spain | - | - | 0.300 (0.136)* | 0.564 (0.154)*** |
| 3 out of 100 (3%) | France | - | - | 0.227 (0.127) | 0.294 (0.134)* |
| 3 out of 100 (3%) | Spain | - | - | 0.131 (0.121) | 0.238 (0.137) |
| Speed of Onset | | | | | |
| 1 week | France | - | - | -0.064 (0.135) | -0.072 (0.143) |
| 1 week | Spain | - | - | 0.022 (0.129) | 0.036 (0.142) |
| 2 days | France | - | - | -0.043 (0.127) | -0.016 (0.136) |
| 2 days | Spain | - | - | -0.080 (0.121) | -0.035 (0.137) |
| Flare Management | | | | | |
| Yes | France | - | - | -0.085 (0.098) | -0.073 (0.106) |
| Yes | Spain | - | - | -0.130 (0.093) | -0.120 (0.108) |
| Long-term Disease Management | | | | | |
| Should not be used long-term | France | - | - | 0.033 (0.144) | 0.036 (0.149) |
| Should not be used long-term | Spain | - | - | 0.172 (0.136) | 0.224 (0.153) |

| | | | | | |
|---|---|---|---|---|---|
| Yes, with the possibility for pauses | France | - | - | 0.034 (0.123) | 0.087 (0.129) |
| Yes, with the possibility for pauses | Spain | - | - | 0.153 (0.121) | 0.299 (0.135)* |
| Administration | | | | | |
| Oral pill, once or twice daily | France | - | - | -0.042 (0.119) | 0.002 (0.130) |
| Oral pill, once or twice daily | Spain | - | - | -0.152 (0.111) | -0.098 (0.132) |
| Check-ups | | | | | |
| Occasional check-ups required | France | - | - | -0.010 (0.138) | 0.042 (0.148) |
| Occasional check-ups required | Spain | - | - | -0.223 (0.132) | -0.189 (0.153) |
| No check-ups required | France | - | - | -0.043 (0.130) | 0.017 (0.140) |
| No check-ups required | Spain | - | - | -0.249 (0.124)* | -0.195 (0.144) |
| Country of residence | | | | | |
| France | Overall | - | -0.148 (0.084) | - | - |
| Spain | Overall | - | -0.280 (0.084)*** | - | - |
| UK | Overall | - | Reference | - | - |
| 4. Model information | | | | | |
| Parameters | - | 18 | 20 | 54 | 54 |
| LL | - | -4866.9 | -4861.2 | -4833.7 | -4833.7 |
| AIC | - | 9769.8 | 9762.4 | 9775.4 | 9775.4 |
| BIC | - | 9886.6 | 9892.2 | 10125.7 | 10125.7 |
| APR | - | 8.30% | 8.40% | 8.20% | 8.20% |

Abbreviations: AIC, Akaike information criterion; APR, Adjusted McFadden Pseudo $R^2$; BIC, Bayesian information criterion; HMNL, heteroskedastic multinomial logit; IMNL, interacted multinomial logit; LL, log-likelihood; MLE, maximum likelihood estimate; MNL, multinomial logit; RIMNL, re-estimated interacted multinomial logit; SE, standard error

Significance: *** P-value < 0.001, ** P-value < .01, * P-value < .05

**Online Supplemental Table 3. Validity assessments**

| Assessment | Full sample N=404 | France N=114 | Spain N=145 | UK N=145 |
|---|---|---|---|---|
| Choice stability, n (%) | | | | |
| Passed the test | 260 (64) | 71 (62) | 94 (65) | 95 (66) |
| Failed the test | 144 (36) | 43 (38) | 51 (35) | 50 (34) |
| Choice dominance [a], n (%) | | | | |
| Passed the test | 359 (89) | 109 (96) | 130 (90) | 120 (83) |
| Failed the test | 45 (11) | 5 (4) | 15 (10) | 25 (17) |
| Serial non-participation [b], n (%) | | | | |
| Never select the same option | 384 (95) | 108 (95) | 136 (94) | 140 (97) |
| Always select treatment A | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Always select treatment B | 1 (0) | 1 (1) | 0 (0) | 0 (0) |
| Always select old treatment | 19 (5) | 5 (4) | 9 (6) | 5 (3) |
| Dominated decision making [c], n (%) | | | | |
| Itch reduction | 6 (1) | 1 (1) | 2 (1) | 3 (2) |
| Skin appearance | 1 (<1) | 0 (0) | 1 (1) | 0 (0) |
| Eye inflammation | 3 (1) | 1 (1) | 1 (1) | 1 (1) |
| Serious infections | 8 (2) | 3 (3) | 3 (2) | 2 (1) |
| Speed of onset | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Flare management | 1 (<1) | 0 (0) | 0 (0) | 1 (1) |
| Long-term disease management | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Administration | 21 (5) | 8 (7) | 8 (6) | 5 (3) |
| Check-ups | 2 (<1) | 1 (1) | 0 (0) | 1 (1) |
| None | 362 (90) | 100 (88) | 130 (90) | 132 (91) |
| Response time for DCE choice task section only [d], n (%) | | | | |
| Adequate | 391 (97) | 111 (97) | 143 (99) | 137 (95) |
| Inadequate | 13 (3) | 3 (3) | 2 (1) | 8 (5) |
| Time to complete DCE choice task section only, n (%) | | | | |
| <5 min | 236 (58) | 64 (56) | 93 (64) | 79 (54) |
| 5-10 min | 123 (30) | 38 (33) | 38 (26) | 47 (32) |
| 10-15 min | 28 (7) | 7 (6) | 9 (6) | 12 (8) |
| 15-20 min | 4 (1) | 1 (1) | 2 (1) | 1 (1) |
| >20 min | 13 (3) | 4 (4) | 3 (2) | 6 (4) |

Abbreviations: DCE, discrete choice experiment

[a] A respondent was considered to have failed the test if they chose the inferior (dominated) option as their preferred treatment.

[b] A respondent was classified as a serial non-participant if they choose the same option for all 12 experimental choice tasks.

[c] Decision making was considered dominated when the respondent choses the best option on one attribute in all 12 experimental tasks.

[d] Response times in the lower 10% of the distribution were classed as too fast, and those in the upper 10% of the distribution as too slow. A participant was considered to have had an adequate response time if <80% of choice tasks were answered too fast or too slow.

**Online Supplemental Table 4. Multinomial logit results: maximum likelihood estimates**

| Attribute | Level | MLE (SE) | 95% CI |
|---|---|---|---|
| Alternative specific constant | Old treatment | 1.46 (0.12)*** | [1.23; 1.69] |
| | Option A | -0.04 (0.04) | [-0.11; 0.03] |
| Itch reduction | 2 out of 10 (20%) | Reference | - |
| | 4 out of 10 (40%) | 0.59 (0.06)*** | [0.47; 0.71] |
| | 5 out of 10 (50%) | 0.76 (0.06)*** | [ 0.65; 0.87] |
| Skin appearance | 1 out of 10 (10%) | Reference | - |
| | 2 out of 10 (20%) | 0.21 (0.06)*** | [ 0.10; 0.33] |
| | 4 out of 10 (40%) | 0.48 (0.06)*** | [ 0.36; 0.60] |
| Eye inflammation | 20 out of 100 (20%) | Reference | - |
| | 10 out of 100 (10%) | 0.27 (0.05)*** | [ 0.18; 0.37] |
| | 0 out of 100 (0%) | 0.64 (0.06)*** | [ 0.53; 0.75] |
| Serious infections | 6 out of 100 (6%) | Reference | - |
| | 3 out of 100 (3%) | 0.31 (0.05)*** | [ 0.21; 0.40] |
| | 0 out of 100 (0%) | 0.72 (0.06)*** | [ 0.61; 0.83] |
| Speed of onset | 2 weeks | Reference | - |
| | 1 week | 0.01 (0.05) | [-0.09; 0.11] |
| | 2 days | 0.18 (0.05)*** | [ 0.08; 0.27] |
| Flare management | No | Reference | - |
| | Yes | 0.09 (0.04)* | [ 0.01; 0.17] |
| Long-term disease management | Yes, without the possibility for pauses | Reference | - |
| | Should not be used long-term | 0.06 (0.05) | [-0.05; 0.16] |
| | Yes, with the possibility for pauses | 0.36 (0.05)*** | [ 0.27; 0.45] |
| Administration | Injection under the skin, every 2 weeks | Reference | - |
| | Oral pill, once or twice daily | 0.25 (0.05)*** | [ 0.16; 0.35] |
| Check-ups | Frequent check-ups required | Reference | - |
| | Occasional check-ups required | 0.24 (0.05)*** | [ 0.14; 0.35] |
| | No check-ups required | 0.31 (0.05)*** | [ 0.21; 0.41] |
| Number of observations | | 4848 | |

| | |
|---|---|
| Model log-likelihood at convergence | -4867 |
| Adjusted pseudo $R^2$ | 0.08 |
| Bayesian information criterion | 9887 |

Abbreviations: CI, confidence interval; MLE, maximum likelihood estimate; SE, standard error

## ONLINE SUPPLEMENTAL FIGURE LEGENDS

**Online Supplemental Figure 1. Comparisons of estimates between MXL and MNL models**

Abbreviations: MNL, multinomial logit; MXL, mixed logit

**Online Supplemental Figure 2. Comparison of estimates between NL and MNL models**

Abbreviations: MNL, multinomial logit; NL, nested logit

**Online Supplemental Figure 3. MNL results by country**

Abbreviation: CI, confidence interval

**Online Supplemental Figure 4. MNL results by age**

Abbreviation: CI, confidence interval

**Online Supplemental Figure 5. MNL results by gender**

Abbreviation: CI, confidence interval

**Online Supplemental Figure 6. MNL results by Patient Oriented Eczema Measure (POEM) overall score. Clear/Mild: 0–7; Moderate: 8–16; Severe: 17–28**

Abbreviation: CI, confidence interval

**Online Supplemental Figure 7. MNL results by self-reported eczema severity. Mild: very mild/mild; Not mild: moderate/severe/very severe.**

Abbreviation: CI, confidence interval

**Online Supplemental Figure 8. MNL results by experience self-injecting**

Abbreviation: CI, confidence interval