Evidence review group report: Botulinum toxin type A for the prophylaxis of headaches in adults with chronic migraine

Produced by Warwick Evidence

Authors:

Dr Pamela Royle, senior research fellow

Dr Ewen Cummins, health economist

Dr Clare Walker, academic clinical fellow

Dr Sam Chong, consultant neurologist

Dr Ngianga-Bakwin Kandala, senior research fellow

Professor Norman Waugh, professor of public health medicine and HTA

Correspondence to;

Professor Norman Waugh, Warwick Evidence (room 102), Warwick Medical School, Gibbett Hill, Coventry CV4 7AL

Email: norman.waugh@warwick.ac.uk Tel: 02476 151585

Date completed: December 2011

This report should be referenced as: Royle P, Cummins E, Walker C, Chong S, Kandala,N, Waugh N. Botulinum toxin type A for the prophylaxis of headaches in adults with chronic migraine: a single technology assessment. Warwick Evidence, 2011.

Sections highlighted in yellow are 'academic in confidence'.
Sections highlighted in aqua are 'commercial in confidence'.

Table of Contents

Lists of Tables and Figures

Tables

Figures

Abbreviations

AMPP	American Migraine Prevalence and Prevention
BDI	Beck Depression Inventory
BSC	Best Supportive Care
BASH	British Association for the Study of Headache
CLAD	Censored Least Absolute Deviations
CDH	Chronic Daily Headache
CM	Chronic Migraine
CTH	Chronic Tension Headache
CTTH	Chronic Tension-Type Headache
EM	Episodic Migraine
EQ-5D	European Quality of Life-5 Dimensions
FSFD	Fixed-Site Fixed-Dose
FTP	Follow-The-Pain
GHQ	General Health Questionnaire
GON	Greater Occipital Nerve
HA	Headache
HDPM	Headache Days Per Month
HIT	Headache Impact Test
HRQL	Health Related Quality of Life
HFEM	High Frequency Episodic Migraine
ICER	Incremental Cost Effectiveness Ratio
IMMPACT	Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
IBMS	International Burden of Migraine Study
ICHD	International Classification of Headache Disorders
IHS	International Headache Society
MO	Medication Overuse
MOH	Medication Overuse Headache
MHRA	Medicines and Healthcare products Regulatory Agency
MIDAS	Migraine Disability Assessment Questionnaire
MSQ	Migraine Specific Quality of Life
MID	Minimally Important Differences
ONB	Occipital Nerve Blocks
OLS	Ordinary Least Squared
PHQ	Patient Health Questionnaire
PREEMPT	Phase III REsearch Evaluating Migraine Prophylaxis Therapy 1
QALY	Quality Assessed Life Year
STAIS	State-Trait Anxiety Inventory State
TTH	Tension Type Headache
TPM	Transition Probability Matrices

Summary

1.1 Scope of the manufacturer submission

The manufacturer gives a brief account of chronic migraine, noting that it may affect as many as 1.6% of the adult population. Chronic migraine is defined as per the NICE scope, except that the scope adds that headaches should have persisted for at least 3 months. As per the NICE scope, the submission focuses on chronic migraine (omitting episodic migraine) and prevention of headaches, not treatment. The manufacturer follows the international consensus of definitions of migraine, wherein people having headaches on 15 or more days a month are classed as having chronic migraine, and those having headaches on 14 or fewer days are classed as having episodic migraine.

1.2 Summary of clinical effectiveness evidence submitted

The Allergan submission relies on data from the PREEMPT trials. PREEMPT is short for "Phase 3 Research Evaluating Migraine Prophylaxis". The two PREEMPT trials were very similar and are pooled in the main analysis.

In the PREEMPT trials, patients were randomised to have injections of Botox or saline placebo, with the same number of injections to the same 31 to 39 sites across 7 muscle groups. The basic regimen was 31 injections given as fixed doses to fixed sites (FSFD), but the physicians treating the patients could add more injections (up to an additional 8) as part of what is called the "follow the pain" (FTP) technique, targeting areas thought to be particularly involved.

The PREEMPT studies had two main phases, firstly the randomised trial over the first 24 weeks, and then an open extension wherein all the patients were given botox injections, using the same FSFD and FTP regimens as in the trial phase.

The main outcome was the number of headache days. At baseline, both groups were having headaches on about 20 days out of 28 days. In the pooled results data at 24 weeks, the Botox group had 8.4 fewer headaches days per 28 days. However, the placebo group also showed a marked improvement, with 6.6 fewer headache days per month (Allergan submission Table 5.26), giving a difference of 1.8 days per month. Most other outcomes favoured Botox, including frequency of headache episodes, frequency of migraine days, cumulative headache hours and frequency of triptan intakes. Most of these differences were clinically quite small but highly statistically significant. One outcome which showed no difference was frequency of acute analgesia use where the difference was not significant ($p = 0.257$).

These data were presented for all the 1384 patients in the PREEMPT trials. The group that resembled the NICE decision problem group, namely those who had previously tried 3 or more prophylactic treatments, was smaller in number (479). They had a smaller reduction from baseline of 7.4 fewer headache days in the Botox group, and 4.7 in the placebo group, so the difference between groups was slightly greater at 2.4 days.

Another way of reporting results is by the proportion of patients having good responses, for example those having a 50% or more decrease in headache days. At week 24, 47% of the Botox group and 35% of the placebo group were 50% responders.

1.3 Summary of ERG critique of clinical evidence submitted

The trials were generally of good quality. In the PREEMPT 1 trial, patients in the Botox group had at baseline a significantly lower frequency of migraine episodes (11.5 vs 12.7, p=0.006) and frequency of headache episodes (12.3 versus 13.4, p=0.023), and significantly more cumulative hours of headache occurring on headache days (295.7 versus 274.9), p =0.022) compared to those in the placebo group. The manufacturer suggests that this was a chance phenomenon.

The ERG's main concern was about whether blinding was maintained. In previous Botox trials, 70% of participants receiving Botox correctly guessed what they had received, because of the changes in muscle tone, such as reduced wrinkling of foreheads. If a significant proportion of the Botox group realised they had had Botox, would that have created a greater placebo effect than was seen in the placebo group? How much of the difference in headache days and other outcomes could have been due to different-sized placebo effects?

As un-blinding is a significant factor in controlled trials of prophylactic treatment of chronic migraines, the International Headache Society (HIS) guidelines recommend that subjects and investigators should be questioned at the end of the trial regarding their opinion as to whether the subject was assigned to active or placebo group during study. The ERG note that this was not done in the PREEMPT trials.

The original primary outcome in PREEMPT 1 was to have been frequency of headache episodes, but this was changed to headache days for reasons that seem reasonable to the ERG. There was no difference in the frequency of headache episodes in PREEMPT 1 at 24 weeks, but there was in headache days (reductions of 7.8 with Botox, and 6.4 with placebo (p = 0.006). The difference was greater in PREEMPT 2, with reductions of 8.7 headache days in the Botox group and 6.3 in the placebo group.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

The manufacturer presents a cost utility Markov model with a 12 week cycle and a two year time horizon. For the base case the manufacturer concentrates upon the patient population with 1 or more prior prophylaxis treatments; the 1-prior group. Those with 3 or more prior prophylaxis treatments, the 3-prior group, are presented as a sensitivity analysis.

The model employs 6 health states defined by ranges of headache days per month (HDPM). The best three health states span the episodic migraine range of 0 to 14 HDPM, while the worst three health states span the chronic migraine range of 15 to 28 HDPM. Movements between these health states and treatment discontinuations are governed by transition probability matrices (TPMs). These TPMs are derived from patient level data counts. The first two cycles use patient level data from the RCT. For the placebo arm, the 2nd cycle TPM is repeatedly applied over the extrapolation period in the placebo arm.

For the Botox arm, a common 3rd, 4th and 5th cycle TPM is calculated based upon the summed 12 weekly movements between health states as drawn from the Botox-Botox open label arm. While unusual, the resulting TPM is repeatedly applied over the extrapolation period in the Botox arm.

For the Botox arm the model applies a negative stopping rule due to lack of efficacy at week 24. If patients have not improved by at least two health states they discontinue and have the placebo TPM applied to them.

For the Botox arm the model also applies a positive stopping rule due to efficacy at week 48. If patients are in any of the best three health states they are assumed to come off Botox treatment and remain stable thereafter.

Utility values are drawn from MSQ values among the 1-prior patient population. These are converted to utilities through a mapping function estimated from MSQ and EQ-5D data within the manufacturer commissioned International Burden of Migraine Study (IBMS). Utilities for the health states are differentiated by treatment arm.

Botox administration is costed at 30 minutes of consultant time. The IBMS also provides the data for the additional resource use associated with the individual health states, though this appears to rely upon a probably small subset of Scottish patients' data.

The base case deterministic estimate for the 1-prior group presented by the manufacturer is an additional cost of £549 and a patient gain of 0.09 QALYs to yield a cost effectiveness estimate of £5,828 per QALY.

Sensitivity analyses supplied by the manufacturer show the cost effectiveness estimate for the 1-prior group to be sensitive to:

- The administration cost assumed for Botox. The reference cost for a neurology outpatient follow up visit worsens the cost effectiveness to £7,972 per QALY.
- The costs associated with the health states of the model. Varying only the hospital admission cost by 20% varies the cost effectiveness from £5,360 per QALY to £6,252 per QALY.
- The time horizon of the model. A one year horizon worsens the cost effectiveness to £14,098 per QALY while a 5 year time horizon improves it to £300 per QALY.
- The stopping rules. Not applying the negative stopping rule worsens the cost effectiveness estimate to £7,946 per QALY. Not applying the positive stopping rule worsens the cost effectiveness estimate to £12,486 per QALY. Not applying either worsens the cost effectiveness estimate to £15,294 per QALY.

For the 3-prior group the deterministic estimates are an additional cost of £543 and a patient gain of 0.09 QALYs, with a cost effectiveness estimate of £6,083 per QALY. The sensitivity analyses are not presented by the manufacturer for the 3-prior group.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

Given the NICE scope, concentrating upon the 1-prior group is surprising. It may have been better to concentrate upon the 3-prior group, and present the 1-prior group as a sensitivity analysis.
The validation data presented by the manufacturer raises some concerns. This appears to suggest that the model may overestimate the average net reduction in mean headache days per month (HDPM) at week 24 by more than 50%. Given that model extends to two years, some further explanation of this is warranted.

The construction of the transition probability matrices (TPMs) for cycles 3, 4 and 5 in the Botox arm of the model is unusual in that the data is not specific to the time frames of these cycles, but is rather pooled between them. The resulting common TPM is also used for extrapolation within the Botox arm. Alternative approaches may worsen the cost effectiveness estimate by between 5% and 10%.

The ERG views the negative stopping rule due to lack of efficacy at 24 weeks as reasonable, though there may be some debate as to whether this would require an improvement of two health states or

only one health state. An improvement of only one health state within two cycles worsens the cost effectiveness estimate for the 3-prior group from £6,083 per QALY to £8,354 per QALY.

A key structural element within the modelling is the likelihood of a positive stopping rule at 48 weeks being possible in practice, and if possible how long these patients will remain stable and not require retreatment. Removing this positive stopping rule roughly doubles the cost effectiveness estimate for the 3-prior group from £6,083 per QALY to £12,542 per QALY.

The model submitted by the manufacturer is strongly non-linear, with the probabilistic estimate of cost effectiveness being more than double that of the deterministic estimate. Some elements of the probabilistic modelling as submitted by the manufacturer seem unwarranted; e.g. treating the identity matrix as having an uninformed prior and modelling it probabilistically. Removing these elements reduces the degree of non-linearity but does not eliminate it. The cost effectiveness estimate for the 3-prior group increases from a deterministic £6,083 per QALY to a probabilistic £14,004 per QALY using the manufacturer model and to a probabilistic £11,447 per QALY if the ERG revisions are applied.

The ERG has some concerns around the estimation of utilities. Botox is anticipated to result in additional quality of life gains over placebo for a given health state. The manufacturer supplies some supporting data for this result. But a particular concern is that the utility values and Botox increments are estimated from the 1-prior patient population, and then applied to the 3-prior patient population. The manufacturer has not demonstrated that similarly large utility increments for a given health state would apply within the 3-prior patient population. These within health state utility increments account for around half of the anticipated patient gain from Botox.

The cost of a Botox administration based upon 30 minutes of consultant time may be too optimistic. It seems more appropriate to apply the consultant led follow-up neurology outpatient cost to this.
The ERG views the resource use associated with the health states of the model as too high. In particular hospital admission rates appear to be too high. The resource use data may have relied upon a very small sample size for chronic migraine based upon the Scottish subset of the UK data within the IBMS.

1.6 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG has revised the base case of the manufacturer to apply:

- The cost of a neurology outpatient visit for Botox administration, and monitoring among the placebo group and those discontinuing treatment

- The resource use of the overall International Burden of Migraine Study (IBMS) rather than the poorly reported Scottish subset
- A minor revision to the average cost per A&E visit

These result in a deterministic cost effectiveness estimate for the 3-prior group of £10,257 per QALY. Removing the positive stopping rule at one year worsens this to £17,517 per QALY. The probabilistic central estimates, based upon 1,000 iterations, are £16,165 per QALY with the positive stopping rule and £26,494 per QALY without the positive stopping rule.

Estimating the cycle 3, 4 and 5 TPMs more in line with the patient data that relates to their time span would probably further worsen these estimates, by a reasonable but not dramatic amount.

The impact of estimating utilities specific to the 3-prior group, using both the Migraine Specific Quality of Life (MSQ) mapping exercise and the joint MSQ plus Headache Impact Test-6 (HIT) mapping exercise, is unknown.

1.7. Other uncertainties

The dose used in the PREEMPT trials was 155 units for the FSFD regimen, with up to another 40 units for the FTP component, if used. This requires a 200 Unit vial of Botox to be used. Were a dose of 150 Units to be used, one 100 Unit vial and one 50 Unit vial could be used. The dose would be 3% less, but the drug cost would be 28% less. The ERG has seen nothing to convince us that a dose of 155 Units is cost-effective compared to 150 Units, or indeed less than 150 Units.

As regards the FTP add-on dosage, the decrease in headache free days was less in the FTP group, raising doubts about the cost-effectiveness of FTP. However it may be that they were more severely affected at baseline.

A striking feature of the PREEMPT, and indeed other headache trials, was the large placebo effect. Most of the effect was placebo effect, and as mentioned above, the ERG wondered if it was possible that the Botox group had a greater placebo effect, due to un-blinding, than the placebo group. It might not be unexpected for a new treatment involving multiple injections and close attention, to have a considerable placebo effect. However, two issues arise. Firstly, could the injection of saline have had a non-placebo effect? Given the transient nature of any effect that saline might have, this seems unlikely. Secondly, if Botox was not approved for use in the NHS, the placebo response would be lost unless an alternative was found.

1.8 Research needs

If Botox were to be approved by NICE, we suggest that further trials be done, including;

- A range of doses, comparing the PREEMPT 155 unit FSFD regimen against doses below 150 units. Would a lower dose be equally effective but at lower cost?
- An RCT of FSFD versus FSFD plus the FTP extra dosage. Is the extra FTP dosage cost-effective?

2 Background

Botulinum toxin is produced by the organism *Clostridium botulinum*, and can cause severe food poisoning. (It was once called "sausage poisoning" – the Latin for sausage is botulus). There are seven serotypes, but types A, B and E have accounted for most human botulism.

One effect is muscle paralysis and in botulism, there can be problems with breathing.

The toxin has been adapted for therapeutic use. It was first approved by the Food and Drug Administration in 1989 for strabismus and blepharospasm. Other approved indications have included cervical dystonia, axillary hyperhidrosis (excessive sweating from the armpits) and cosmetic use, such as removing frown lines from the forehead.

In most of these uses, the aim is to reduce activity in muscles. It has been tried in tension and other headaches, but the focus of this review is on its use in migraine.

We will refer to the therapeutic form in short as Botox, rather than use the full names, botulinum toxin type A (used in the NICE scope) or onabotulinumtoxin A (as used in the Allergan trials).

The proprietary name for the Allergan form is Botox. There are other forms produced by other companies, such as Dysport, produced by Medicis Pharmaceutical Corporation and Xeomin (Merz Aesthetics) that are also approved by the FDA for some purposes. (Note that dilutions and dosages are different so one cannot simply be substituted for the other).

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer's description (page 20 of the submission by Allergan) is brief. Key points made are;

- The prevalence of chronic migraine (CM) is estimated to be 1.6% of the adult population. This is higher than the figure given in the NICE scope – but see below
- It is considered to result in progression from episodic migraine (EM)
- CM is defined as 15 or more headaches days a month, of which 8 are migrainous – as per the NICE scope, except the scope specifies that this should have persisted for at least 3 months.
- CM can sometimes revert to EM.

Defining chronic migraine: ERG commentary

Headaches are defined as chronic when people have them for more than 15 days a month. This seems somewhat arbitrary. It means that someone with migraine who has headaches on 9 to 14 days a month, is defined as having "high frequency episodic migraine" (HFEM). A one-day a month increase would shift her into the chronic migraine category.

Chronic headache can be classed as chronic migraine, or as chronic tension-type headache (CTTH). However as a recent study states, the boundary between CM and CTTH is controversial, and some people can have both.[1] It is important in this appraisal because Botox is licensed only for CM. If there are problems distinguishing CM from CTTH, the number of patients who might be treated, could rise sharply.

The NICE scope uses the International Headache Society definition of headaches on 15 or more days a month for at least 3 months where the attacks meet the criteria for migraine without aura on at least 8 days per month. This creates a problem if on some of the 8 days, the sufferer has auras. The scope goes on to say;

"*Despite these criteria, in clinical practice there is a lack of consensus regarding the definition of chronic migraine*."

The evolution of the disorder of "chronic migraine"

Migraine.

The International Classification of Headache Disorders (ICHD) I (1988) classified migraine without aura as headache attacks lasting from 4-72 hours with at least two of the following characteristics - unilateral location, pulsating quality, moderate or severe intensity and aggravated by physical activity.[2] At least one of nausea and/or vomiting or photo and phonophobia must be present. Other causes must be excluded.

Migraine with aura describes headache episodes with similar characteristics as migraine without aura, with the addition of fully reversible neurological symptoms preceding the development of headache.

Chronic headache.

In the 1990s it was recognised in specialist centres that a proportion of complex patients were suffering with daily or near daily headaches.[3] They often were heavy users of analgesics and other drugs. The term "chronic daily headache" (CDH) had been applied to this condition, however, some neurologists objected to this term believing it to be a general description of a problem, rather than a

specific patho-physiological diagnosis. It was noted that it arose in individuals who had previously suffered from an episodic headache disorder, usually migraine without aura or tension type headache.

The first ICHD classification described only one chronic headache - chronic tension headache (CTH) (previously called CDH) as a headache present for at least 15 days a month for 6 months. The quality of the pain had to fulfil two of the following criteria; must be pressing or tightening, of mild or moderate severity (inhibiting, but not prohibiting activities), with bilateral location and not aggravated by physical activity. In addition there must be no vomiting and only one of nausea, photophobia, or phonophobia could be experienced. Other causes must be excluded by history or examination. Episodic tension type headache (TTH) had a similar quality of headache but lasted from 30 minutes up to 7 days.

The problem arose when an individual fulfilled the frequency and chronicity criteria of CTH, but presented with a pulsating headache, aggravated by physical activity and with nausea, more in keeping with the pain of migraine without aura. There was no category for this type of individual.

It was further recognised that one individual may suffer from more than one type of headache and that a mixture of chronic and episodic headaches may co-exist. Whether TTH and migraine without aura were a continuum as suggested in ICHD-I or separate entities as suggested by others, this mixed clinical picture had been acknowledged in previous classifications which provided for "chronic mixed headaches" or "combination headaches".[3] The ICHD recommended that an individual's headache disorder be coded according to all the types of headache they experienced. Therefore, one person may have had migraine without aura, CTH and if overusing medication, headache induced by chronic substance use. They also recommended that the headache days per month of suffering each type of headache were recorded with the code.

Chronic migraine

Early CM and Transformed migraine (TM)

The current concept of CM suggests that migraine headaches may change over time and emerged from the "transformed migraine" (TM) entity first coined in the early 1980s and revised in the 1990s. TM was a CDH of long duration (at least 15 days/month) found in patients with episodic migraine with a headache duration of at least 4 hours (if untreated).[4] The transformation was a history of increase in headache frequency with a decrease in migrainous features. Patients were distinguished by the presence or absence of medication overuse headache (MOH). This transformation was difficult to show in practice as patients had difficulty remembering the characteristics of prior headaches. Therefore, TM was revised in 1996 so that instead of a clear history of transformation in a migraine patient, a history of migraine, or a history of escalation over 3 months, or a current headache that

except for duration met the criteria for migraine was accepted. This remains the most inclusive of all chronic migraine definitions.

CM as a diagnostic label first appeared in a paper by Manzoni in 1995 as migraine with an unfavourable evolution which fulfilled diagnostic criteria for migraine with headache for at least 6 days a week for one year.[5] The term CM, although with a greatly altered definition, first appeared in ICHD-II in 2004.[6] TM was not included as a category as it was felt that not all patients remembered the evolution of their headaches, and some didn't evolve. [7]

ICHD-II

Because of the chronicity of the disorder, CM was added as a complication of migraine. At first it was defined as headache fulfilling the definition of migraine without aura occurring on 15 or more days per month for more than 3 months and in the absence of medication overuse. Medication overuse headaches (MOH)) were thought to be responsible for the majority of headache symptoms in these chronic cases, therefore CM could only be diagnosed 2 months after medication overuse had ceased.

ICHD-IIR

The changing definitions of TM/CM captured different groups. Field testing of criteria on 638 patients with CDH by Bigal et al in 2006 revealed 87% of the patients met the revised criteria for TM whereas 6% of the group met strict ICHD-II CM criteria rising to 10% when those with MO were included.[8] The main problem was the necessity for 15 days of headache fulfilling the criteria migraine without aura. Percentages increased when the following definitions were considered – migraine or probable migraine on at least 15 days per month (48%), migraine or probable migraine on at least 50% of the 15 or more headache days (88%), migraine on at least 8 days per month or responding to migraine specific treatment (95%), 15 or more headache days per month and 8 or more days of migraine or probable migraine.[8]

In June 2006, both CM and MOH were revised as the definition of CM was too exclusive. The Headache Classification Committee justified its decision to reduce the number of migraine days by arguing that as migraine without aura attacks develop, they often go through a phase where they fulfil the criteria for TTH before the characteristic pulsating migraine headache commences. If a triptan was taken correctly the patient may not progress to the migraine stage. The only way to diagnose CM would therefore be to withhold treatment. Secondly, as triptans have been found to be of little effect in TTH, if a triptan aborted an attack, this suggested this headache would have developed into migraine.[9]

Briefly, the revised CM criteria were a headache (both TTH and/or migraine), occurring on at least 15 days per month for at least 3 months, occurring in a patient who has had at least 5 attacks of migraine without aura, and on 8 or more days per month. These headaches must fulfil criteria for migraine

without aura and symptoms be relieved when treated with triptans or ergot. There should be no overuse of medication.

A note about MOH

Medication overuse in chronic headache is a frequently observed phenomenon. It was included in ICHD-II, but its definition depended upon stopping the overuse and the patient showing improvement. In other words it could only be diagnosed when the patient no longer had it. Oleson et al remarked that not all patients improve after the discontinuation of acute medications. Some do, some become newly responsive to prophylactic medication and some patients show no improvement at all.[9]

MOH was revised in ICHD-IIR appendix and was diagnosed where a headache was present on 15 or more days a month, with regular overuse for 3 months or more of one acute/symptomatic treatment drugs, defined as:

- ergotamine, triptans, opioids or combination analgesics on 10 or more days a month
- simple analgesics or combination of ergotamine, triptans, analgesics, opioids on 15 or more days a month.

In addition the headache has developed or markedly worsened during medication overuse.

So someone who takes 2 paracetamol tables on 15 days a month, would be classed as a medication over-user. N.B. Not all patients with medication overuse will have MOH.

Effects of changing definition: comments and controversies

There is ongoing debate over the current definition. Definitions which have followed the Silberstein model became predominant in the 1990s to mid 2000s. Manzoni et al have suggested further subdivisions of the present definitions.[10] They propose adding frequency into the general definition of migraine without aura thereby making CM a subdivision of all migraine without aura when headaches are experienced 10-20 days a month. Further they suggest a category of TM which would capture those with headaches more than 20 days per month who never have more than 5 consecutive days without a headache. The TM category allows a more mixed pattern of headache, whereas the headache of CM must meet migraine without aura criteria. Manzoni also criticised the MOH definition as too wide.[10] He summarises, "the only sure data we have are those about ergotamine, caffeine and combined medication including barbiturates or codeine or prochlorperazine." There is also the difficult problem that the direction of causality is uncertain with CDH and MOH – do patients overuse their medicines because of the frequency and nature of their pain, or does the overuse of medicines cause the chronicity of their headaches?

One of the original aims had been to provide a classification for chronic headache patients which would allow them to be entered into clinical trials, thereby providing testing of new treatment options for their particular problems. The absence of MO is necessary for the diagnosis of CM. However, MO is so common that the IHS has advised that patients with MO be included in trials to produce evidence for effective treatment.[11] However, they add that patients with MOH should be stratified accordingly.[12]

Box 1 Other definitions

Episodic migraine
Both migraine with aura and migraine without aura are usually episodic conditions. Headaches usually last for between 4 and 72 hours and then the person recovers. It may be days, weeks or months before they suffer another attack.

Migraine without aura
This is the most common form of migraine. It is a recurrent headache disorder with attacks lasting 4-72 hours if untreated or unsuccessfully treated. The headache is characterised by at least two of the following features – unilateral location, pulsating quality, moderate or severe intensity, and is aggravated by, or causes the individual to avoid physical activity. In addition either nausea and vomiting or photo- and phonophobia will be present. (Adapted from ICHD- II)
Five attacks must be experienced before the diagnosis is made

Migraine with aura
About one third of migraine sufferers report auras. Migraine with aura is a recurrent disorder with attacks of reversible neurological symptoms (the aura) that develop gradually over 5-20 minutes and last for less than an hour. The aura generally precedes and symptoms have usually departed before the headache commences. The headache usually has the characteristics of a typical migraine headache. Some individuals experience aura without developing a headache. Symptoms may affect sight, or speech or other senses. Visual symptoms may include flickering lights, spots, lines or loss of vision. Pins and needles or numbness are examples of sensory symptoms. (From NHS Clinical Knowledge Summaries and ICHD-II)

Common migraine
Common migraine is the old term for migraine without aura

Classic(al) migraine
Classic or classical migraine is the old term for migraine with aura ICHDII

Imploding and exploding headaches
Burstein et al[13] classified headaches as "exploding" wherein people described a feeling of pressure inside their head, and "imploding" where people described the headache as being crushing, from outside to inside, in type. The latter were more likely to respond to Botox.

Prophylactic medication
These are medicines taken regularly (usually daily) for a period of time to try and reduce the

frequency and severity of headaches. They are taken whether or not the patient has a headache at the time.

Acute or rescue medicines
In contrast to prophylactic medicines these are taken when a headache commences. They may include simple pain relieving drugs or medicines such as triptans which are migraine specific medicines taken when a migraine begins to stop that attack becoming worse.

Placebo effect
Even if the therapy is irrelevant to the patient's condition, the patient's attitude to his or her illness, and indeed the illness itself, may be improved by a feeling that something is being done about it. [14]

Photophobia
An abnormal sensitivity to light. In migraine this is usually temporary and sufferers seek a darkened room in which to rest.

Phonophobia
An abnormal sensitivity to sound. Again, a temporary symptom of migraine.

Medication over-use
The definition of medication over-use varies by medication;
Ergotamine, triptans, opioids, combination analgesics - $\geq$ 10 days a month and $\geq$ 2 days a week
Simple analgesics - $\geq$ 15 days a month and 2 or more days per week.

Box 2 IHS Classification ICHD-II. Appendix A1.5.1 Chronic Migraine

Chronic migraine [15]
A. Headache (tension-type and/or migraine) on ≥15 days per month for at least 3 months*
B. Occurring in a patient who has had at least five attacks fulfilling criteria for 1.1 Migraine without aura
C. On ≥ 8 days per month for at least 3 months headache has fulfilled C1 and/or C2 below, that is, has fulfilled criteria for pain and associated symptoms of migraine without aura
1. Has at least two of a-d
(a) unilateral location
(b) pulsating quality
(c) moderate or severe pain intensity
(d) aggravation by or causing avoidance of routine physical activity (e.g. walking or climbing stairs)
and at least one of a or b (a) nausea and/or vomiting (b) photophobia and phonophobia .
2. Treated and relieved by triptan(s) or ergot before the expected development of C1 above
D. No medication overuse† and not attributed to another causative disorder‡
*Characterization of frequently recurring headache generally requires a headache diary to record information on pain and associated symptoms day-by-day for at least 1 month.
†Medication overuse as defined under 8.2 *Medication-overuse headache*.
‡History and physical and neurological examinations do not suggest any of the disorders listed in groups 5-12, or history and/or physical and/or neurological examinations do suggest such a disorder but it is ruled out by appropriate investigations, or such disorder is present but headache does not develop in close temporal relation to the disorder.

Prevalence

Prevalence of chronic migraine in the UK

It is difficult to know the true prevalence of chronic migraine in the UK. The NICE scope states that the prevalence of CM in the UK is not known, but may be around 1 in 1000. This figure seems too low. This is partly because of the evolving nature of the definitions of chronic migraine and because only one prevalence study has been performed in adults. Steiner et al study in 2003 looked at the prevalence of migraine in England using a telephone survey of 4007 individuals aged 16-65.[16] CM was not described separately and the paper does not give a detailed enough breakdown of frequency to estimate the prevalence of CM. However, 7.6% of males and 18.3% of females reported migraine with or without aura within the last year using similar diagnostic criteria to IHS. These are similar to WHO figures. A WHO factsheet quotes European and American studies showing that 6-8% of men and 15-18% of women experience migraine each year.[17]

In 2009 Natoli et al produced a systematic review of the global prevalence of chronic migraine.[18] They found 16 papers representing 12 individual studies; 6 were European, 4 in the Americas and 2 in the Western Pacific. The 12 studies gave 14 prevalence estimates (the French study provided 3 estimates depending upon definition).

Of the European studies using a definition including 15+ migraine days a month, 1 year prevalence was low. A small Danish study had no-one reporting CM and a Norwegian study up to 0.7%. French estimates ranged from 0.86 – 2.1% using increasingly broad definitions similar to CDH with migraine or migrainous disorders (ICHD-I).

A Brazilian study gave a prevalence of 5.1% and a US study of 4.1%. The latter used the broad "transformed migraine" definition. The former was a relatively small study (n=625) and a definition of CDH with migraine or migrainous disorder. However, it remains double the level of the French study (2.1%) which used a similar definition, perhaps reflecting a true regional difference.

A literature review considered 107 worldwide publications and reported that almost 3% of the population had chronic headaches (defined as 15 or more days a month).[19] In a large German population study, 0.28% of individuals met a definition of CM that required 15 days of migraine headaches a month, but included those overusing medication. A 'pure' definition of CM, excluding those with medication overuse, gave a prevalence of 0.09%. So depending on the definition used, the proportion treated could vary by 3-fold.[20]

Extrapolating to the UK and excluding those studies with a significantly more permissive definition than that in the NICE scope, a prevalence in the region of 1% seems reasonable.

The burden of migraine

Little work on populations has been done in the UK but several large scale studies of chronic headache morbidity have been performed in the Europe and US. Quantifying the burden of CM on sufferers depends on the definition used. However defined, studies have shown that CM causes significant suffering and disruption to the life of the patient.

A 2003 Irish outpatient clinic study by Cassidy et al compared chronic and episodic headaches including CM, EM, and chronic and episodic TTH.[21] This population is likely to have more severe and intractable problems than the general population. The focus was on psychiatric co-morbidities. They found higher depression scores in those with more frequent headaches, but no difference in anxiety. The chronic headache group had two-fold higher MIDAS disability scores than those with episodic headaches. The CM group scored the highest on GHQ-28 (anxiety and depression), Beck Depression Inventory (BDI) and State-Trait Anxiety Inventory (STAI-S) showing an increased burden of psychological distress. They also missed double the number of workdays than the chronic tension-type headache (CTTH) group, missing an average of 5 a month. Applying several models they found that frequency of headache predicted MIDAS disability, whereas frequency and neuroticism predicted GHQ and BDI depression.

The validated self-administered questionnaire-based American Migraine Prevalence and Prevention (AMPP) study compared respondents with CM and EM (n=11904). Both groups were more likely to be Caucasian, overweight and had similar education levels. CM sufferers reported lower household income levels, were less likely to be employed full time and nearly twice as likely to be occupationally disabled.[22]

They found 57.4% of patients with CM missed more than 5 days of work or school over a 3 month period with a similar percentage reporting reduced productivity in household chores and a third reporting 5 days of lost family activities. This is similar to the Cassidy study above.

The International Burden of Migraine Study (IBMS) was a cross-sectional web-based survey conducted in 2009 with participants from Australia, Canada, France, Germany, Italy, Spain, UK, Taiwan and the US. Once again CM and EM were compared and they found similar socio-demographic characteristics with greater numbers of Caucasians and Hispanics, more overweight and fewer in full-time employment in the CM group, with no difference in education levels.[23]

The headache duration of the CM group was double that of the EM group lasting 24.1 hours if medication was taken and 65.1 hours without medication. Nearly 90% of the CM group described their headache intensity as severe or very severe (7/10 or greater on a linear scale) and 92% described their headache pain as severe.

To assess HRQoL, questionnaires were employed. The MIDAS headache disability questionnaire showed the CM group had significantly raised scores (67.67 versus 13.57 for EM) and the proportion with scores above 21 indicating grade IV-A severe or IV-B very severe were 78% for CM versus 23%. The Migraine-Specific Quality of Life Questionnaire (MSQ) is divided into 3 domains assessing the limitations in a sufferer's daily social and work-related activities, how migraines prevent such activities, and the emotions associated with migraine. The higher the score, the better the quality of life, and those with CM scored an average of 7.9, 6.0 and 13.5 points lower than the EM group (adjusted model).

Lanteri-Minet's systematic review in 2011 included 16 studies reporting quality of life measures, most commonly the SF-36 but include chronic headaches which do not have migrainous features.[24] Three out of four studies reported lower quality of life scores for those with CM compared to other forms of CDH. Estimates of missed school or work days for CM varied from <1 -5 per month. One study considered CDH with or without MOH and concluded MOH was associated with greater disability and productivity loss.

Co-morbidity

There are significant co-morbidities reported with CM which have an impact on the individual's experience of health.

Psychiatric and pain disorders were reported more often by CM patients in the AMPP. These include anxiety, bipolar disorder, depression (assessed both by PHQ-9 and self-report of a professional's diagnosis OR 2.00), chronic pain (OR 2.49) and arthritis. Respiratory and circulatory diseases were also more common in the CM group – allergies and hay fever, asthma, bronchitis and chronic bronchitis (OR 1.99), circulation problems, COPD, hypertension, high cholesterol, sinusitis and stroke. Odds ratios are all raised in the range 1.23 - 2.49 and are statistically significant.[22]

The IBMS confirmed an increase in psychiatric disorders especially anxiety and depression, pain disorders and circulatory disorders in the CM group.

In summary, CM sufferers experience pain and limitation of their functioning which is greater than patients with other headache disorders. The presence of co-morbidities, especially other pain syndromes make symptoms more difficult to treat. Depression and pain have a synergistic relationship, one affecting the other, and this is recognised in the British Association for the Study of Headace (BASH) guidelines which state "migraine treatments may fail unless these underlying problems are dealt with." [25]

2.2 Critique of manufacturer's overview of current service provision

This is very brief, the main points being;

- Only 17% of CM patients are seen by a specialist (not defined, but presumably mainly neurologists?) (page 21)
- Only 20% of patients with CM have had the diagnosed confirmed by a neurologist (citing Bigal et al[26])
- Patients with CM are initially treated with a range of oral prophylactics as per the BASH guidelines

The current pathway of care is outlined in figure 2.2 on page 23, (as reproduced in Figure 1 below) which shows where Botox would fit into the pathway

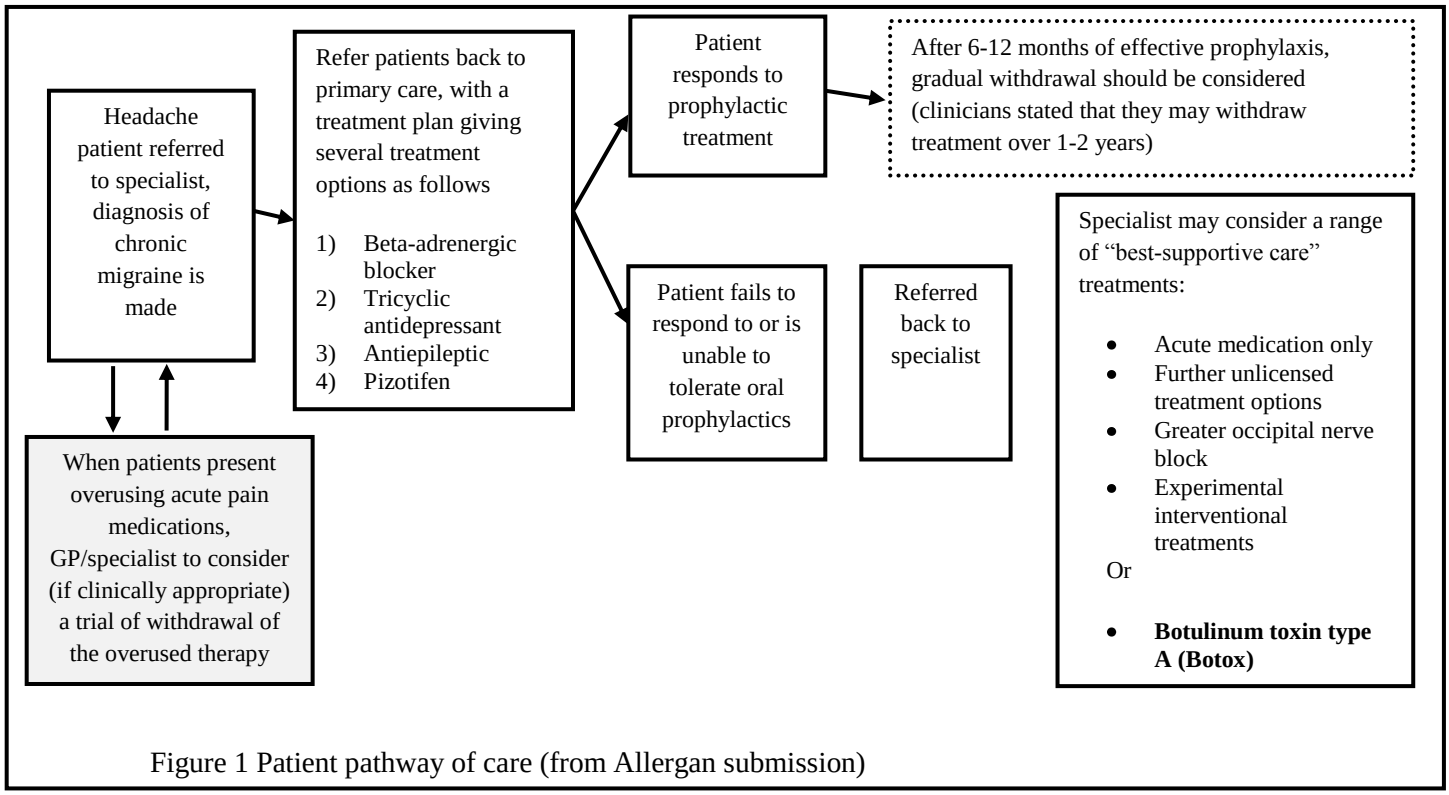


Figure 1 Patient pathway of care (from Allergan submission)

The NICE scope gives a more detailed account of current therapy;

"Preventative (also called prophylactic) treatment of migraine can be an important component of chronic migraine management. The goals of preventive therapy are to reduce the frequency, painfulness, and/or duration of migraines, and to increase the effectiveness of medication that is taken at the earliest signs of a migraine headache (known as abortive therapy). Preventative treatment of migraines may also help to avoid medication overuse headache, otherwise known as rebound

headache, which is linked to overuse of pain medications, and is a common problem among people with migraines.

Preventative interventions can take many forms including nutritional supplements, lifestyle alterations such as increased exercise and avoidance of migraine triggers, and prophylactic migraine medications. Prophylactic chronic migraine medications are generally considered for people who have at least two attacks a month, whose attacks are increasing in frequency, whose attacks cause significant disability despite abortive treatment, or who cannot take abortive treatment for migraine attacks. Prophylactic migraine medications include betablockers (propranolol, atenolol,metoprolol, nadolol and timolol), sodium valproate, topiramate, antidepressants (amitriptyline, nortriptyline, imipramine, desipramine), pizotifen, gabapentin and cyproheptadine. "

However, the Allergan submission notes that they expect the place of Botox to be in those who have failed on oral prophylactics, so they might reasonably argue that some of the NICE text is not relevant.

UK Practice: the British Association for the Study of Headaches guidelines

The British Association for the Study of Headaches (BASH) produced guidance on the management of migraine in 2010.[25] Although chronic migraine is not mentioned as a specific subtype, the most severe end of the stepwise ladder of management described is applicable to migraines which are chronic. The guidance is summarised in some detail below. This enables an assessment of the generalisability of the results from PREEMPT to UK practice.

6.4.13 Limits to acute therapy: frequency of use

"Over-frequent use of drugs for acute intervention may be one criterion for prophylaxis ...On a regular basis

a) use of *triptans on 10 or more days a month* or *analgesics on 15 or more days a month* is inappropriate for migraine and is associated with a clear risk of MOH;

b) use of either on *two or more days every week* calls for close enquiry into how it is used, and review of the diagnosis."

6.5.1 Indications for prophylaxis

"Prophylaxis is used to reduce the number of attacks in circumstances when acute therapy, used appropriately, gives inadequate symptom control.....Over–frequent use of acute therapy is also a

criterion for migraine prophylaxis, but prophylactic drugs are inappropriate and will be ineffective for medication overuse headache. This condition must first be excluded."

Therefore, it would be expected that patients would not be overusing medication without trying prophylaxis. However, 36% of patients in PREEMPT had never received oral prophylaxis. As prophylactics do not help in MOH a period of withdrawal of overused acute medication seems indicated by the advice above, prior to use of prophylaxis.

6.5.2 Dose-titration

"Most prophylactics are used within a dose range, and in general must be up-titrated slowly to an effective dose (or to the maximum dose) in order to avoid side-effects that will precipitate premature discontinuation. This can lead to a delay in efficacy which itself, unfortunately, sometimes triggers discontinuation. Careful explanation is needed."

6.5.3 Duration of use

"...Drugs that are effective should be continued for 4-6 months, then withdrawal considered to establish continued need. Withdrawal is best achieved by tapering the dose over 2-3 weeks.

Prophylactic drugs that are apparently not effective should not be discontinued too soon since efficacy may be slow to develop, particularly when dose-titration is necessary...In practice patients usually decide when they stop medication, so careful explanation is needed lest they be labelled non-responders inappropriately (eventually, perhaps, to all drugs). There is not absolute guide but, in the absence of unacceptable side-effects, 6-8 weeks is a reasonable trial following dose-titration...."

6.5.4 & 5 First and second-line prophylactic drugs

BASH bases its recommendations on a mixture of evidence and expert opinion. It describes a "good" formal evidence base for beta-blockers, topiramate and valpoate, and an adequate one for amitriptyline.

It recommends atenolol or bisoprolol (but further evidence of efficacy needed for the latter), or amitriptyline (or less sedating alternatives, again requiring further evidence).

It recommends topiramate and sodium valproate as second line agents.

The ERG notes that the evidence for comparing topiramate with Botox includes 2 trials. [27,28] Both were described as pilots, and had 60 and 59 patients respectively. They both found topiramate and Botox to be of similar efficacy. Neither study gives details of what prior prophylactics had been tried, but in one study (Cady 2011[27]) it appears that 41% had not been on any prior prophylactic. Hence they are probably not relevant to the NICE decision problem group.

6.5.6 Third line prophylactic drugs

BASH advises consideration of gabapentin. Methysergide is generally considered (on limited formal evidence) to be the most effective prophylactic but is held in reserve because of its association with retroperitoneal fibrosis (although that is said not to occur in courses of less than 6 months duration) and a severe rebound headache on withdrawal for which some experts cover with prednisilone.

BASH also advises a combination of beta-blockers and amitriptyline as a synergistic effect is claimed, without formal evidence.

6.5.7 Other drugs used in prophylaxis but with limited or uncertain efficacy

These include pizotifen and clonidine which are widely used but with little trials evidence of efficacy. Verapamil too has limited trials evidence and SSRIs are of uncertain value,

Lisinopril, montelukast, candesartan, riboflavin and co-enzyme q-10 all show potential benefit from RCTs, but further research is needed before they can be recommended.

Prophylaxis for hormone related migraine.

"An effect on hormones on migraine is common, and greater for migraine without aura. Evidence suggests estrogen withdrawal triggers migraine in some women."

"Menopause itself commonly exacerbates migraine and symptoms can be relieved with optimised HRT".

This suggests that consideration of a woman's hormonal status may be useful and may point to therapeutic options away from usual migraine prophylaxis and acute therapies. As the majority of patients in PREEMPT were female this may be of importance.

<u>Difficulties with prophylaxis</u>

The population identified by the NICE Scope is expected to have failed to respond to at least 3 prior prophylactic therapies. As the above BASH guidelines suggest, the decision to try prophylaxis is based on inadequate symptom control and over frequent use of acute medications. The primary care guidelines website NHS Clinical Knowledge Summaries (CKS) suggests beta blockers or amitriptyline as prophylactic agents suggesting an expectation that at least one of these be initiated in primary care.[29]

There is limited information about the prophylactic agents which PREEMPT participants had previously tried. The MHRA report (p33) does comment that at the time of phase 3 studies there was no consensus about prophylactic medication and "investigators were instructed to record the past use

of any of the medications listed in the guideline provided [which included all of the BASH first, second and third line agents from the 2007 guideline], irrespective of whether the treatment was prescribed as headache prophylaxis or not. Overall 64% of enrolled subjects had a history of prior headache prophylaxis medication use when this broad list of possible medications was used." .[30]

Although nearly two thirds of participants (879/1384) had tried some form of prophylaxis it is uncertain how many had tried a medication with an evidence base; 41.5% of enrolled subjects had a history of prior BASH first-line migraine headache prophylaxis (MHRA report table 10, page 32). However it is uncertain how many of these trials of medication were accompanied by the kind of careful counselling advocated above intended to aid adherence to prophylaxis and help define a true non-response to therapy. It is therefore uncertain how many participants had been supported to continue with a medication trial for at least 6-8 weeks before it was considered a failure.

97.5% of the subgroup of patients with a history of use of 3 or more prophylactic medicines discontinued their use due to lack of efficacy and/or due to side effects, which are the very circumstances BASH caution about.

In summary, a substantial proportion of the PREEMPT population had not followed similar patient journeys to that anticipated by the NICE scope. A third had received no oral prophylaxis and only 41.5% had received a first-line drug as recommended by BASH. It is impossible to know how much explanation, counselling and support during prophylaxis trials the participants had received and how many of them may have continued with and received benefit from these agents with this help.

Current service provision

The ERG noted the comment that only 17% of people with CM are currently seen by specialists. If Botox is approved by NICE for use, presumably by specialists, then the burden on neurology services might increase considerably. We therefore asked Allergan if they had details on the expected time costs for specialist services, and received the following helpful reply;

Manufacturer response to Q A.2:

"Unfortunately the market research source cited did not capture adequate information to address the ERGs questions around the frequency and duration of visits".

> In order to provide further information, we conducted a brief email survey of UK Neurologists treating chronic migraine patients. Six responses were received and these are summarised below in Table 1

This survey demonstrates that assumptions made for the placebo arm of the economic model, in terms of frequency of follow up are potentially conservative. Advisers responding represented a good geographical spread across England and Scotland and both secondary and tertiary settings.

Note that the follow-up consultations would not allow sufficient time for Botox injections.

Table 1 Survey of UK Neurologists treating chronic migraine patients

Question	A1	A2	A3	A4	A5	A6
Thinking about patients presenting with a diagnosis of chronic migraine to your specialist clinics, how long would an initial consultation last?	30 mins	20-30 mins	30 mins	30 mins	30 mins	30 mins
How long might a follow up consultation last?	15 mins	10-15 mins	15 mins	15 mins	15 mins	15 mins
Following an initial assessment and perhaps initiation of oral prophylaxis or advice regarding acute medications, when would a patient be followed up?	3-6 months	2-6 months	3 months	3-6 months	3 months	No routine f/u*
How many times might a patient with chronic migraine be seen in your clinical before being discharged? (average, recognising wide variation)	4	3	4	4	4-10	*Can have multiple attendances

3 Critique of manufacturer's definition of decision problem

The Allergan account of the decision problem (submission, page 27) is similar to that of NICE. In brief;

3.1 Population

The population is as defined by NICE: adults with headaches on at least 15 days per month with migraine type headaches on at least 8 days per month; who have failed to respond to at least three prior preventive drugs, and in whom medication overuse has been appropriately managed.

3.2 Intervention

Botox injections

3.3 Comparators

"Standard management" not specified by NICE or Allergan, but since prophylactic measures have failed, standard management is presumably "rescue medications" such as analgesics during attacks.

3.4 Outcomes

Largely as per NICE

3.5 Other relevant factors

Economic analysis: The submission states "as per NICE methods". However the timescale of the modelling is only two years – this is discussed by the ERG later.

Some issues arise. Firstly, a minor point - the Allergan account of the decision problem does not mention that the definition of CM requires a 3-month duration of headaches of the defined frequency and type. Secondly, as discussed later, the main evidence comes from the PREEMPT trials which recruited a much broader patient group, in which medication over-use had often not been managed. Obviously, recruitment to PREEMPT preceded the NICE statement of the decision problem, and it could be argued that the broader group with its medication over-users aids the generalisability of the trials. We return to these issues in the clinical effectiveness section

4. Clinical effectiveness

4.1Critique of the manufacturer's systematic review

The manufacturer's approach was to update the searches carried out by the Scottish Intercollegiate Guidelines Group (SIGN) for the 2008 guideline (SIGN 107) and to add searches for interventions not covered by SIGN, such as nerve block.[31]

The Allergan submission notes (page 37) the difficulty which would be had if they had tried to limit searches to the decision problem subgroup (DP subgroup) and so no attempt was made to restrict searches to that subgroup. The submission also notes problems with definitions and terms, and so studies of patients with "transformed migraine" and "chronic daily headache" were retrieved and included if relevant.

The review identified over 20 papers from 7 trials. However most trials were not relevant to the decision problem. Four were against oral prophylactic drugs, and hence in a group in whom these had not failed. The three trials of Botox against placebo (in effect of Botox + rescue drugs versus placebo plus rescue drugs) included the two PREEMPT trials, discussed below, and the Freitag 2008 study.[32] The submission excludes the Freitag trial from further consideration, on the grounds that it was a small (60 patients) pilot study with a high drop-out rate (30%), and did not report prior prophylactic use. The ERG concurs, but for completeness, details of this study are included as Appendix 1.

In summary, the Allergan review of the evidence concluded that the only relevant evidence on clinical effectiveness comes from the PREEMPT trials.

4.1.1 Searches

The ERG carried out its own searches for previous reviews, trials, safety data, and studies of the epidemiology and costs of chronic migraine. No relevant trials were found that were not mentioned in the Allergan submission. Some relevant studies that were not used in the industry submission were found and used by the ERG, but these would not have been indexed in databases at the time the searches for the Allergan submission were done.

We noted previous reviews of Botox for episodic migraine, which concluded that it was ineffective.

4.2 Critique of the PREEMPT trials

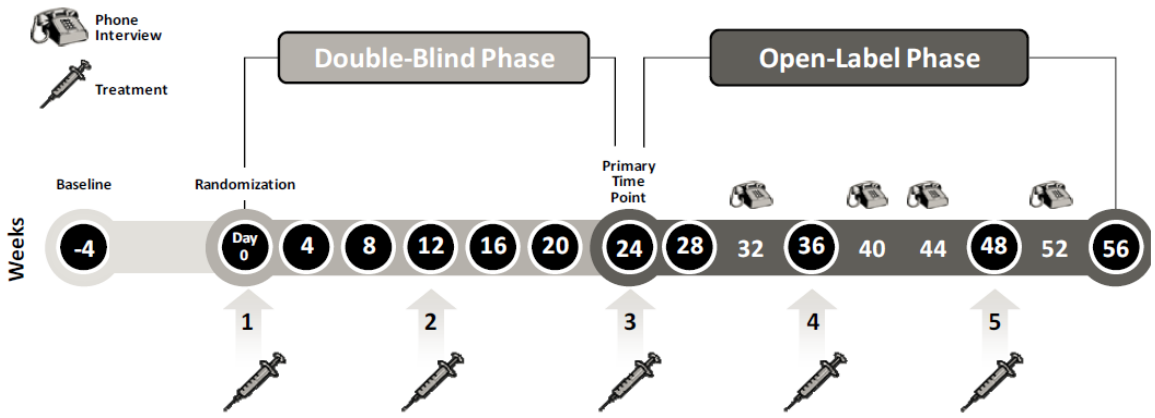
The key evidence to support the efficacy and safety of Botox in the chronic migraine population is taken from a pooled analysis of two pivotal phase 3 double-blind placebo controlled trials, called the PREEMPT (Phase 3 REsearch Evaluating Migraine Prophylaxis Therapy (PREEMPT) studies. [33]

4.2.1 Study design

The two PREEMPT studies were identical in design, apart from the designation of the primary and secondary endpoints.

They included a 4-week screening/baseline phase (week -4 to day 0), followed by a 24-week, double-blind, randomised, placebo-controlled, parallel-group phase, which was then followed by a 32-week open-label extension phase, as shown in the figure below.[33-36] Patients were randomised 1:1 to Botox or placebo. Randomisation was stratified according to acute medication over-use or not during the 28-day baseline. Patients had to provide diary data on ≥20 of 28 days and have had ≥15 headache days, of which each consisted of ≥4 hours of continuous headache and of which ≥50% were migraine or probable migraine days (referred to hereafter as migraine days); they also had to have had ≥4 distinct headache episodes, each lasting ≥4 hours.

Participants were told at entry to the trials that there would be an extension study in which they would all receive Botox.



* In-office visits were conducted every 4 weeks, except when noted by telephone.

Figure 2 Schematic of phase 3 study design - PREEMPT 1 and PREEMPT 2

PREEMPT 1 was conducted from 2006 to 2008 at 56 North American sites. Its primary outcome measure was the mean change from baseline in frequency of headache episodes during the 28 day period ending with Week 24.[34]

The PREEMPT 2 study ran at almost the same time as PREEMPT 1, at 50 North American and 16 European sites.[36] Its primary outcome measure was the change in frequency of headache days. The PREEMPT 1 study protocol and statistical analysis plan were amended, prior to treatment unmasking, to change the primary endpoint from frequency of headache episodes to frequency of headache days.

Why was there two nearly identical studies?

PREEMPT 1 and PREEMPT 2 took place at nearly the same time – the former started 2 weeks later and ended 4 weeks earlier. Market and Solomon asked why were the 2 studies not designed as 1 larger multi-center, multi-country study?[37] Dodick et al responded that the reason is that regulatory authorities have expressed a wish for more than one trial.[38]

It appears that regulatory would ideally like independent confirmation, but they accept non-independent evidence from virtually the same trial sponsored by the manufacturer. The FDA Guidance for Industry Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products[39] states in Section 4 that:

> *"Reliance on a Single, Multicenter Study - Caveats*
>
> *While acknowledging the persuasiveness of a single, internally consistent, strong multicenter study, it must be appreciated that even a strong result can represent an isolated or biased result, especially if that study is the only study suggesting efficacy among similar studies"*

4.2.2 Quality assessment

A quality assessment of both trials is shown in Table 2 and Table 3.

Table 2 Quality assessment of PREEMPT 1

Domain	**ERG Group's assessment**	**Description**
Adequate random sequence generation?	Yes	Randomization sequence generated
Adequate allocation concealment?	Yes	Computer generated

Blinding of participants adequate?	Uncertain	Unblinding may have occurred due to the side-effects of Botox
Incomplete outcome data addressed?	Yes	Reasons for discontinuations due to adverse events, loss to follow-up and other reasons were reported for all patients
Free of selecting reporting?	Yes	All relevant outcomes were reported, as specified in ClinicalTrials.gov, identifier NCT00156910
Groups comparable at baseline?	No*	In the PREEMPT 1 trials patients in the Botox group had significantly lower frequency of migraine episodes (11.5 vs 12.7, p=0.006); frequency of headache episodes (12.3 vs 13.4, p=0.023) and more cumulative hours of headache occurring on headache days (295.7 vs 274.9, p =0.022) compared to those in the placebo group
Intention to treat analysis?	Yes	All efficacy analyses used the intent-to-treat population, which included all randomized patients
Sample size calculation done?	Yes	PREEMPT 1: Planned enrolment for the study was approximately 650 patients. For headache episode frequency, the week 24 minimum retained sample size of n=240 per group, with standard deviation of 5.5, would have >90% power to detect ≥1.75 between-group difference in mean change from baseline, using a two-sided alpha=0.05. Due to the long duration of the study (56 weeks per patient), a larger sample size than what was needed for the week 24 primary efficacy analysis was planned to ensure sufficient sample size at the end of the study for long-term safety evaluations (>150 patients with five active treatment cycles).
Other comments		As unblinding is a significant factor in controlled trials of prophylactic treatment of chronic migraines, the IHS guidelines recommend that subjects and investigators should be questioned at the end of the trial regarding their opinion as to whether the subject was assigned to active or placebo group during study. The ERG note that this was not done in the PREEMPT trials.

The ERG requested that the manufacturer provide details of the observed distributions of baseline variables in the UK sample in the CM subgroups in the PREEMPT 1 and 2 trials. The manufacturer responded that: " *Unfortunately there is insufficient data available to allow this analysis to be conducted. Only PREEMPT 2 included UK sites and the total patient numbers recruited across 3 UK sites are insufficient to permit meaningful analysis*"

Table 3 Quality assessment of PREEMPT 2

Domain	ERG Group's assessment	Description
Adequate random sequence generation?	Yes	Randomization sequence generated
Adequate allocation concealment?	Yes	Computer generated
Blinding of participants adequate?	Uncertain	Unblinding may have occurred due to the side-effects of Botox
Incomplete outcome data addressed?	Yes	Reasons for discontinuations due to adverse events, loss to follow-up and other reasons were reported for all patients
Free of selecting reporting?	Yes	All relevant outcomes were reported, as specified in ClinicalTrials.gov, identifier NCT00168428
Groups comparable at baseline?	yes	Groups were comparable at baseline for all outcomes
Intention to treat analysis?	Yes	All efficacy analyses used the intent-to-treat population, which included all randomized patients
Sample size calculation done?	Yes	In PREEMPT 2, enrolment of 650 patients was planned. For headache day frequency, a week 24 minimum sample size of n=325 per group, with standard deviation of 6.7, would have >90% power to detect ≥1.75 between-group difference in mean change from baseline, using a two-sided a=0.05. Due to the long duration of the study (56 weeks per patient), a larger sample size than what was needed for the week 24 primary efficacy analysis was planned to ensure sufficient sample size at the end of the study for long-term safety evaluations (>150 patients with five active treatment cycles).
Other comments		As unblinding is a significant factor in controlled trials of prophylactic treatment of chronic migraines, the IHS guidelines recommend that subjects and investigators should be questioned at the end of the trial regarding their opinion as to whether subject was assigned to active or placebo group during study. The ERG note that this was not done in the PREEMPT trials.

Patient data from the studies were pooled and the efficacy and safety measures reanalysed, using the same primary endpoint as study PREEMPT 2 (frequency of headache days). The primary endpoint for the pooled analysis was mean change from baseline in frequency of headache days at 24 weeks.

Secondary endpoints were mean change from baseline to week 24 in frequency of migraine/probable migraine days, frequency of moderate/severe headache days, total cumulative hours of headache on headache days, frequency of headache episodes, frequency of migraine/probable migraine episodes, frequency of acute headache pain medication intakes, and the proportion of patients with severe (≥60) Headache Impact Test-6 score at week 24.

All efficacy analyses were also analysed for the medication overuse stratum.

Change in endpoint

The original primary endpoint in PREEMPT 1 was headache episodes, with headache days as the secondary endpoint. However, the International Headache Society issued new clinical trial guidelines for evaluating headache prophylaxis in CM and recommended headache days as the primary endpoint. The FDA also preferred headache days as the endpoint. The PREEMPT investigators therefore changed the primary outcome to headache days. The authors note that this was done without knowledge of the results of PREEMPT 2, before unblinding was done.[38]

The pooled analysis also had headache days as the primary outcome.

The ERG regard the change in primary outcome in PREEMPT 1 as logical and justifiable.

4.2.3 Injection sites

In the PREEMPT studies patients had Botox injections, or placebo (saline) injections using the same injection regimen with clinicians blinded to the treatment allocation for the first 24 weeks. In the 24-week, double-blind phase, patients received a series of 31-39 intramuscular (IM) injections of Botox or placebo (saline) at day 0 and week 12. In the 32-week, open-label phase all patients continuing in the studies received Botox treatment at week 24, week 36 and week 48.

Box 3 shows injection sites. NB the large number of injections. We return to this when considering the placebo effect.

Box 3

DOSAGE AND ADMINISTRATION Chronic Migraine

The recommended dilution is 200 Units/4 mL or 100 Units/2 mL, with a final concentration of 5 Units per 0.1 mL . The recommended dose for treating chronic migraine is 155 Units administered intramuscularly (IM) using a sterile 30-gauge, 0.5 inch needle as 0.1 mL (5 Units) injections per each site. Injections should be divided across 7 specific head/neck muscle areas as specified in the diagrams and Table 2 below. A one inch needle may be needed in the neck region for patients with thick neck muscles. With the exception of the procerus muscle, which should be injected at one site (midline), all muscles should be injected bilaterally with half the number of injection sites administered to the left, and half to the right side of the head and neck. The recommended re-treatment schedule is every 12 weeks.

Diagrams 1-4: Recommended Injection Sites (A thru G) for Chronic Migraine

4.2.4 Patients and flow through study

A total of 1384 adults was randomized to Botox (n=688) or placebo (n=696). Since the two PREEMPT trials were very similar, pooled analysis provided greater power, and could show highly significant differences ($P \leq .004$) favouring Botox over placebo for the change from baseline in frequencies of headache episodes and migraine episodes - a result observed in PREEMPT-2 but not in PREEMPT-1.

Patients

The recruits were men or women aged 18 to 65 years with chronic migraine, defined as ≥15 headache days during the first 28 days of the baseline period, with each day consisting of ≥4 hours of continuous headache, of which at least 50% were migraine or probable migraine days, and to have had ≥4 distinct headache episodes each lasting ≥4 hours.

Exclusions

- Patients were excluded if they used headache prophylactic medications within 4 weeks prior to the start of baseline

- Patients whose headaches were attributed to other disorders such as medication overuse. However, chronic migraine patients with protocol defined excessive use (overuse) of acute medications were included. Many chronic migraine sufferers report inadequate pain relief with acute treatments, resulting in frequent intake in an attempt to relieve their severe and frequent symptoms.

- Patients with fibromyalgia were excluded. It is reported that 11% of people with CM have fibromyalgia so this seemed odd. We sought clarification from Allergan, who replied;

Manufacturer response to Q A.23:

"Fibromyalgia is a complex pain syndrome and common cause of chronic widespread pain. Although pain is the dominant feature, fibromyalgia is also associated with other symptoms such as fatigue, problems sleeping, stiffness, problems with concentration, depression, anxiety, headaches, migraine, and paraesthesia. The PREEMPT studies were designed with specific exclusion criteria for patients with a concurrent diagnosis of fibromyalgia or other co-morbid chronic pain conditions, in order to avoid confounding the efficacy results. Patients with co-morbid chronic pain conditions, other than the disorder of interest, were not permitted to receive treatments for this condition since there is overlap with such potential treatments and migraine prophylaxis (e.g., antidepressants). This could not only confound study results, but enrolling patients into such a long study (56-weeks) during which they could not receive treatment for the chronic pain condition was felt to be unethical and not in the patient's best interest. "

The ERG is uncertain as to the justification for this. People with CM have other co-morbidities. The IBMS reported that 26.5% had chronic pain, of whom under half had fibromyalgia. It is not clear why only those in the chronic pain group with fibromyalgia were excluded. In the IBMS , 34% were reported as having depression. PREEMPT excluded some with depression (Beck score of >28 at day 1 of baseline).

Patients could be discontinued prematurely from the study due to adverse events, lack of efficacy, pregnancy, protocol violation, personal reasons, lost to follow-up, or other reasons.

Flow through study.

Figure 3 (taken from the Allergan submission) showed the patient flow of the patients in the 24 week double blind phase and the 32 week open label phase.

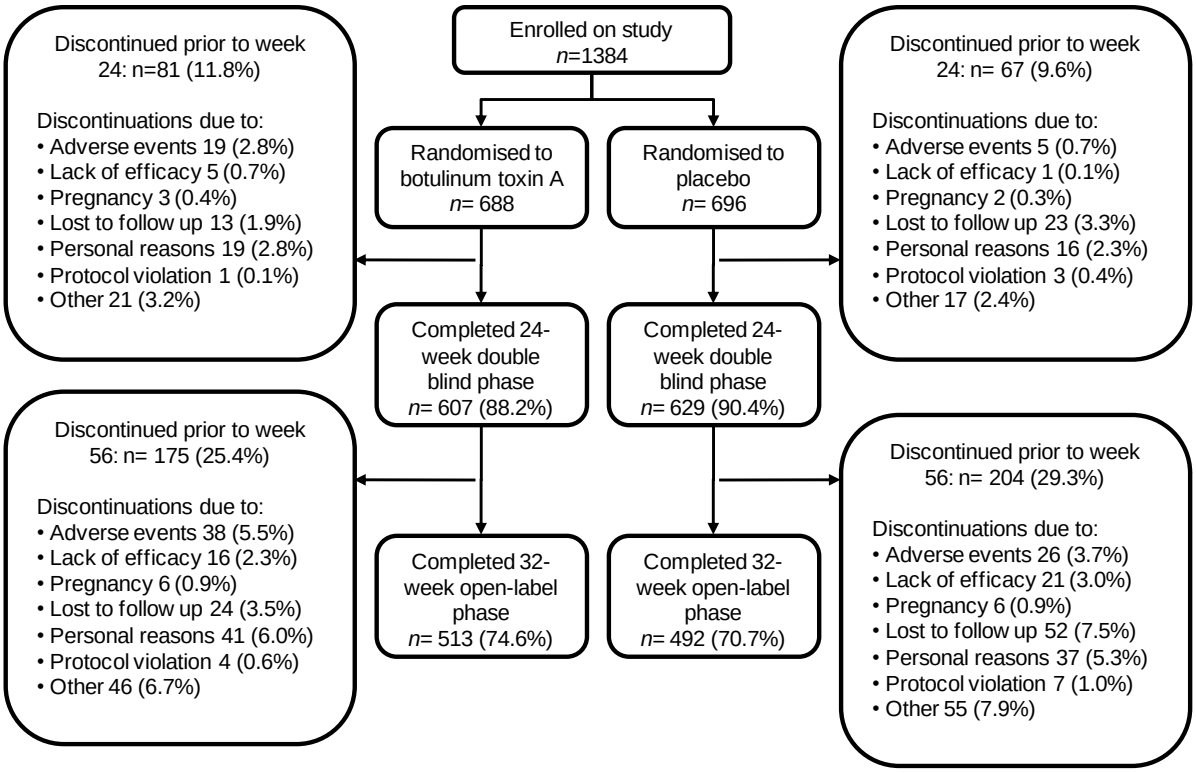


Figure 3 Summary of the patient flow in the pooled analysis

4.2.5 Interventions

Fixed-site fixed-dose injection regimen (FSDF) vs follow-the-pain (FTP) strategy

The standard PREEMPT regimen was 155 U of Botox given as 31 fixed-site fixed-dose (FSFD) injections into 7 head or neck muscles. At the discretion of the physicians, up to 40 additional units could be given at eight sites across 3 head and neck muscles (temporalis, occipitalis and/or trapezius), using the follow-the-pain (FTP) method.[40,41] The maximum dose was 195 U across 39 sites. The decision to give the additional FTP units was based on criteria such as patient-reported location of predominant pain, and severity of the muscle tenderness.

The PREEMPT phase 3 studies were not designed to evaluate the FSFD versus the FTP injection paradigm. Patients were not randomised to FSFD or FSFD + FTP, and so it was not possible to assess the benefits of the additional FTP injections.

All patients treated in the phase 3 studies received the FSFD 155 U injection paradigm. In the double blind phase, 48.6% received FSFD and 51.4% received FTP.

Comparator

Placebo (saline) at day 0 and week 12, with the same number of injection sites.

4.2.6 Baseline matching

A table of baseline characteristics for the 3 or more prior prophylactic treatments population is compared to all PREEMPT patients in the pooled population. See Table 4 below

The *P* values refer to the comparison of the proportions in Botox and placebo groups.

Matching was good in the "decision problem" subgroup, although it is noted that the percentage of patients overusing acute pain medications in the 3 or more prior group is approximately 6% higher than the pooled population of all patients during baseline i.e. (71.6% vs ~65.5% respectively)

However, at baseline the Botox group had significantly fewer headache episodes (12.2 vs 13.0; P = .004) and migraine episodes (11.4 vs 12.2; P = .004), and significantly more total cumulative hours of headache occurring on headache days (295.9 vs 281.2; P = .021). Most patients overused acute pain medications during the 28-day baseline; however, very few (1.7%) had opioid overuse.

These baseline imbalances arose from PREEMPT 1 where there were significant differences in the Botox and placebo groups in some variables at baseline, shown in Table 5 below, from the Allergan submission.

Appendix 2 shows the baseline characteristics of the two trials side by side.

Table 4 Baseline characteristics

	3 or more prior prophylactic treatments population			Pooled PREEMPT phase 3 studies – all patients		
Characteristics	**Pooled sample (n=479)**	**Range**	**P value**	**Botox (n=688)**	**Placebo (n=696)**	**P value**
Mean age (yrs)	[redacted]	[redacted]	[redacted]	41.1	41.5	0.579
Mean time since onset of chronic migraine (yrs)	[redacted]	[redacted]	[redacted]	19.4	19.0	0.488
Women (%)	[redacted]		[redacted]	87.6	85.2	0.185
Caucasian (%)	[redacted]		[redacted]	89.7	90.5	0.602
Mean body mass index (kg/m2)	[redacted]	[redacted]	[redacted]	26.7	27.2	0.080
Mean headache episodes during baseline (SD)	[redacted]		[redacted]	12.2 (5.25)	13.0 (5.5)	0.004*
Mean headache days during baseline (SD)	[redacted]		[redacted]	19.9 (3.68)	19.8 (3.68)	0.498
Mean migraine episodes during baseline (SD)	[redacted]	[redacted]	[redacted]	11.4 (5.02)	12.2 (5.42)	0.004*
Mean migraine days during baseline (SD)	[redacted]	[redacted]	[redacted]	19.1 (3.99)	18.9 (4.05)	0.328
Mean moderate/severe headache days during baseline (SD)	[redacted]	[redacted]	[redacted]	18.1 (4.12)	18.0 (4.25)	0.705
Cumulative headache hours occurring on headache days during baseline (SD)	[redacted]	[redacted]	[redacted]	295.93 (116.88)	281.22 (114.74)	0.021*
Patients who overused acute headache pain medications during baseline (%)	[redacted]	[redacted]	[redacted]	64.8	66.1	0.450
Patients who had previously used ≥1 headache prophylaxis medication (%)	[redacted]	[redacted]	[redacted]	64.8	66.1	0.450
Patients with severe (≥60) HIT-6 score	[redacted]	[redacted]	[redacted]	93.5	92.7	0.565
Mean HIT-6 score	[redacted]	[redacted]	[redacted]	65.5	65.4	0.638

*P=<0.05

Table 5 Baseline disease characteristics for primary and secondary efficacy variables in study PREEMPT 1

Characteristics	BTX (n=341)	PBO (n=338)	p-value
Mean age (yrs)	41.2	42.1	0.317
Mean time since onset of chronic migraine (yrs)	20.3	20.6	0.839
Women (%)	89.1	85.8	0.187
Caucasian (%)	89.4	91.4	0.381
Mean body mass index (kg/m2)	26.7	27.3	0.147
Mean headache episodes during baseline (SD)	12.3 (5.23)	13.4 (5.71)	0.023
Mean headache days during baseline (SD)	20.0 (3.73)	19.8 (3.71)	0.571
Mean migraine days during baseline (SD)	19.1 (4.04)	19.1 (4.05)	0.978
Mean migraine episodes during baseline (SD)	11.5 (5.06)	12.7 (5.72)	0.006
Mean moderate/severe headache days during baseline (SD)	18.1 (4.22)	18.3 (4.23)	0.674
Cumulative headache hours occurring on headache days during baseline (SD)	295.7 (116.81)	274.9 (110.90)	0.022
Patients who overused acute headache pain medications during baseline (%)	66.3	69.8	0.322
Patients who had previously used ≥1 headache prophylaxis medication (%)	59.5	64.2	0.210
Patients with severe (≥60) HIT-6 score	94.4	94.7	0.888
Mean HIT-6 score	65.4	65.8	0.297

We asked Allergan to comment on possible reasons for the underlined differences. In two cases, the Botox group had fewer episodes, and in the third case the placebo group had fewer headache hours. They replied as follows.

Manufacturer response to Q A. 11

"In PREEMPT 1 there were significant differences in the baseline variables for mean headache episodes, mean migraine episodes and headache hours. This imbalance was an outcome of the randomization process. By chance alone, patients in the placebo group had more headache and migraine episodes and fewer cumulative hours, but similar numbers of headache days, migraine days and moderate/severe days than patients randomized to placebo. The study design did not control for baseline headache characteristics".

4.2.7 Results

Effectiveness

We focus here mainly on the pooled results. The results of the two PREEMPT trials are provided in– Appendix 3 (adapted from Table 8 of MHRA report) which shows the mean changes from baseline in primary and secondary efficacy variables at week 24 for the PREEMPT 1 and 2 trials individually. All outcomes, except for the frequency of acute headache pain medication intakes (which showed a non-significant change in both trials), showed a significant change from baseline for PREEMPT 2.

However, there were also non-significant changes in PREEMPT 1 for the following outcomes of: frequency of headache episodes (the primary efficacy variable in PREEMPT 1), frequency of migraine/probable migraine headache episodes, frequency of acute headache pain medication days, proportion of patients with 50% reduction in headache days, proportion of patients with 50% reduction in headache episodes. All of these outcomes, except for the latter one, showed a significant change in the pooled analysis.

For the outcome of change in frequency of headache days, PREEMPT 1 showed a 39% decrease (7.8/20 days) in the Botox group and 32.3% (6.4/19.8) in the placebo group, giving a difference between the groups of 6.7%. For PREEMPT 2 the comparative decreases were 45.2% (9.0/19.9) for Botox and 34% (6.7/19.7) for placebo, giving a difference of 11.2%. Although the differences were both statistically significant, it appeared that Botox was substantially more effective for this outcome than placebo in the PREEMPT 2 trial than in the PREEMPT 1 trial.

For the outcome of frequency of headache episodes, the PREEMPT 1 study showed a 42.3% (5.2/12.3 days) decrease in the Botox group and a 39.6% (5.3/13.4) decrease in the placebo group, giving a non-significant difference between the arms of only 2.7%. PREEMPT 2 showed a 44.2 % (5.3/12.0) for the Botox group and a 36.2% (4.6/12.7) for the placebo group, giving a statistically significant difference between the groups of 8%. Therefore, Botox was also more effective for this outcome in the PREEMPT 2 than the PREEMPT 1 trial.

The main results of the pooled PREEMPT studies (

Table 6) show that there are statistically significant differences for all outcomes, except frequency of acute pain medication intakes.

Table 6 Results of the pooled PREEMPT studies

	PREEMPT Pooled data			
Outcome	Botox(n=688)	Placebo (n=696)	Botox-placebo†	*P* value†
Frequency of headache days	-8.4	-6.6	-1.8 (-2.52, -1.13)	<0.001
Frequency of migraine days	-8.2	-6.2	-2.0 (-2.67, -1.27)	<0.001
Frequency of moderate/severe headache days	-7.7	-5.8	-1.9 (-2.62, -1.26)	<0.001
Cumulative total headache hours on headache days	-119.7	-80.5	-39.2 (-48.40, -21.04)	<0.001
Frequency of headache episodes	-5.2	-4.9	-0.3 (-1.17, -0.17)	0.009
Frequency of migraine episodes	-4.9	-4.5	-0.4 (-1.20, -0.23)	0.004
Frequency of acute headache medication days	-6.1	-5.3	-0.8 (-1.53 to -0.15)	0.016
Frequency of triptan intakes	-3.2	-2.1	-1.1 (-1.74, -0.61)	<0.001
Frequency of acute headache pain medication intakes (all categories)	-10.1	-9.4	-0.7 (-2.68 to 0.69)	0.257

†The 95% confidence intervals and P values are adjusted for baseline and for medication overuse stratification, except for HIT-6 and MSQ scores.

One striking feature is the size of the improvement on placebo. As discussed later, this is not unusual in headache trials.

Figure 4 shows the changes in headache days for the Botox and placebo groups for both the RCT phase (weeks 0 to 24) and the open label phase (weeks 24 to 56) when all the participants received Botox. It shows a number of things;

- The placebo effect persists for at least 24 weeks
- The Botox effect is greatest with the first injection, which raises the question as to when responders should be identified – should the negative stopping rule be applied after the first injection? That would make Botox more cost-effective.

- The second injection has much less effect – indeed no more than the second placebo injection. Does that imply that the effect of Botox injections last longer than 12 weeks, and that the interval could be extended? This could also improve cost-effectiveness.
- The Botox group from the RCT have further improvement in the open label phase, when they are told they are on Botox (as discussed later, many would have guessed already from the side-effects). The placebo group improve on Botox but never catch up, although by 56 weeks the difference is small. This suggests some on-going benefit from the RCT stage, which again might suggest that the effect lasts longer than 12 weeks and that the interval could be extended. However the drop after the 24 week injection is greater than after the 12 week one, suggesting a placebo element due to being told they are getting Botox.

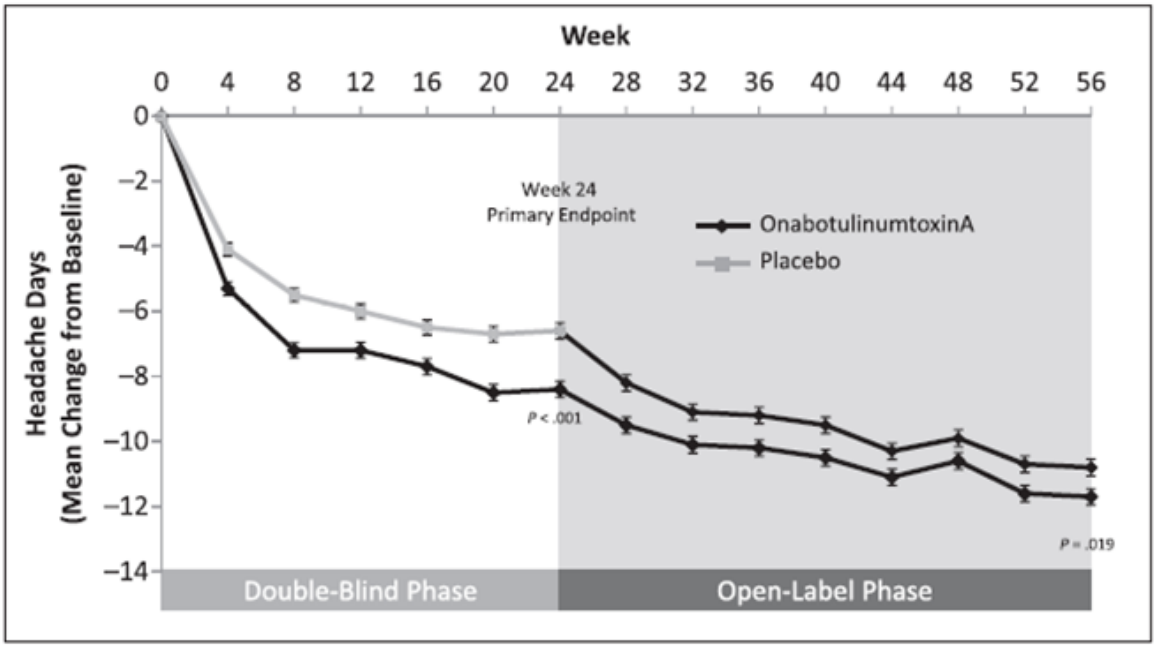


Figure 4 56 week data of the PREEMPT pooled analysis: mean change from baseline in frequency of headache days

The difference between groups is still statistically significant at 56 weeks.

Figure 5 shows a graph of the frequency of acute medication intakes over double blind and open label phase of the study. Curiously, the use of acute rescue medications was not different between the groups at any time, which seems at odds with improvements in other outcomes.

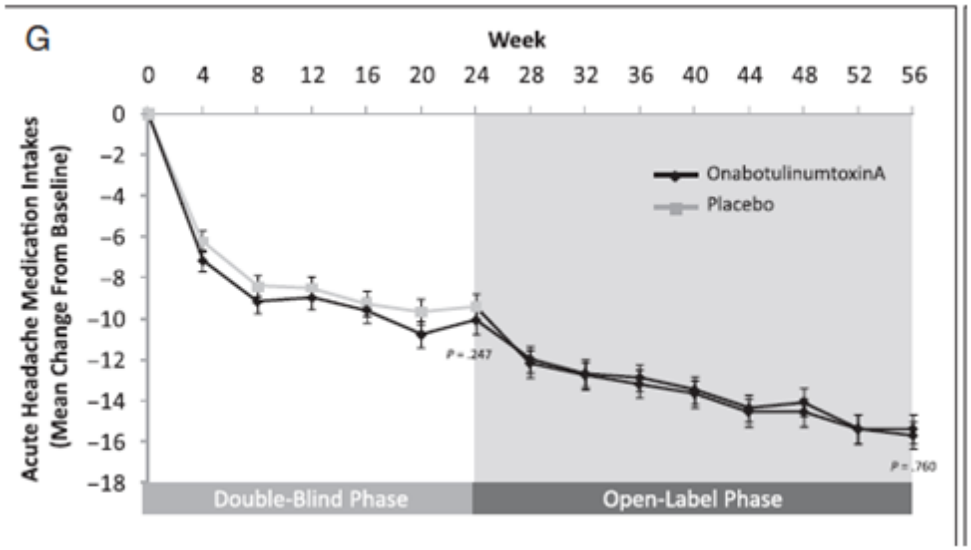
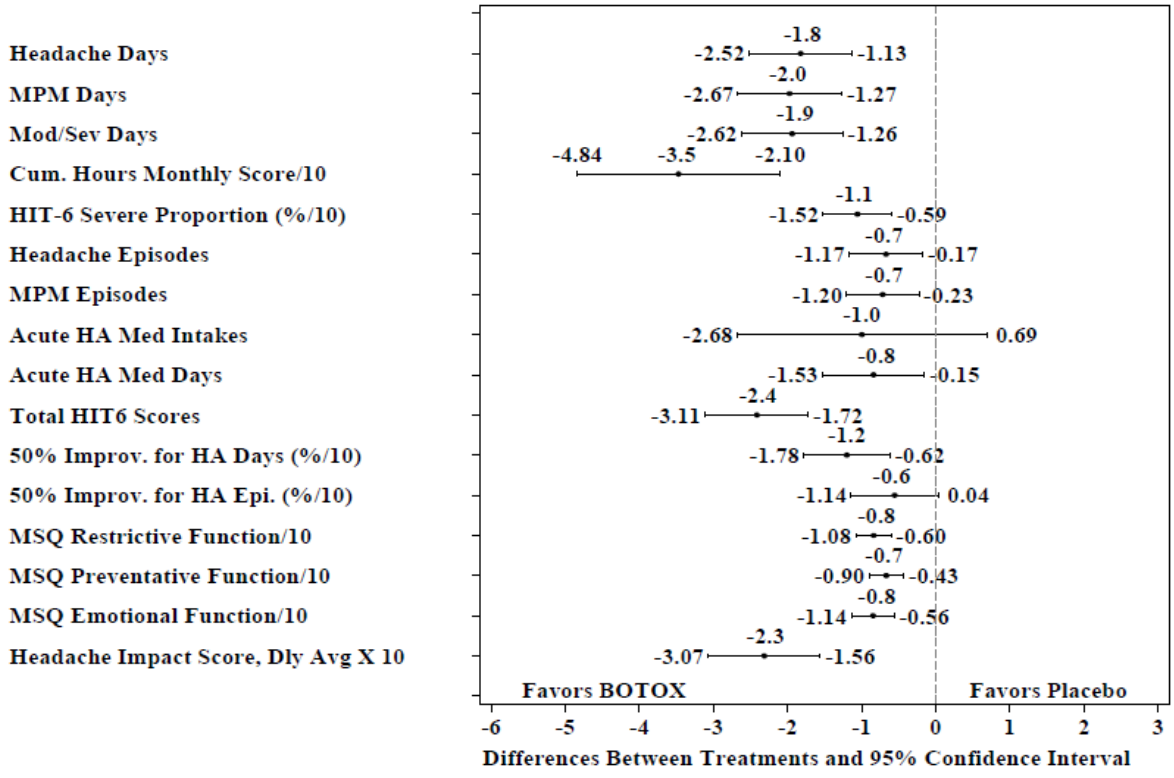


Figure 5 Frequency of acute headache medication intakes (all categories).

Figure 6 below (taken from the MHRA report[30]) shows all outcomes from the pooled PREEMPT studies, with means and 95% CIs



Differences between treatments are displayed as Botox minus placebo, except that 50% improvement is displayed as placebo minus Botox. Thus, negative scores favour Botox. All variables are summarised using mLOCF, except for acute headache med days, MSQ scores and 50% improvement, which are summarised using observed data.

Figure 6 Treatment differences at week 24 for key efficacy variables (ITT population)

4.2.8 Subgroup analyses

Subgroup: BASH guidelines

As detailed in the Background section, the British Association for the Study of Headache (BASH) has produced guidelines for diagnosis and management of migraine, including for prophylaxis.[25]

The MHRA report (page 33) [30] includes a new sub-group analysis in which subjects were re-categorised according to whether they had had first-line prior migraine headache prophylaxis before Botox therapy; 41.5% of subjects had a history of prior BASH first-line migraine headache prophylaxis use

Table 7 Subgroup analysis: headache outcomes according to prior first line BASH medication use

	Had first line BASH prophylaxis before entry (n=575)				Did not have first-line BASH prophylaxis before entry (n=809)			
	Botox (n=277)	Pbo (n=298)	Botox-pbo	P value	Botox (n=411)	Pbo (n=398)	Botox -pbo	P value
Headache episodes	-5.6	-4.6	-1	0.009	-5.3	-4.9	-0.4	0.234
Headache days	-8	-5.6	-2.4	<0.001	-8.7	-7.3	-1.4	0.004
Migraine/probable migraine (MPM) days	-7.8	-5.2	-2.6	<0.001	-8.5	-6.9	-1.6	<0.001
Moderate/severe headache days	-7.4	-4.8	-2.6	<0.001	-8	-6.6	-1.4	0.002
Acute HA Pain medication days	-5.5	-4.7	-0.8	0.093	-6.2	-5.7	-0.5	0.168
≥50 decrease from baseline: HA days	44.6%	31.6%	13%	0.003	49.0%	38.0%	11%	0.006

The difference between Botox and placebo was greater for all outcomes in the prior BASH medication use group. However the decrease in the placebo groups was greater in those with no prior BASH medication use.

Table 8 BASH sub-group – quality of life outcomes and prior first line BASH medication use at week 24 [from MHRA Table 12]

	First line pre-study BASH Med-Use - Yes (n=575				First line pre-study BASH Med Use - No (n=809)			
	Botox n=277	Pbo n=298	Botox-pbo	P value	Botox n=411	Pbo n=398	Botox-pbo	P value
Incidence of total HIT-6 Severe Impact Category Score	71.0%	80.9%	-9.8%	0.006	65.2%	76.1%	-10.9%	<0.001
Mean total HIT-6 score	-4.5	-1.8	2.7	<0.001	-5.0	-2.8	2.2	<0.001
Incidence of ≥5 point improvement on total HIT-6 score	39%	22.1%	16.9%	<0.001	42.1%	27.6%	14.5%	<0.001
Mean migraine specific QoL								
Role Restrictive (RR)	-16.6	-6.3	-10.3	<0.001	-17.2	-10.3	-6.9	<0.001
Role Preventative (RP)	-13.8	-4.3	-9.5	<0.001	-12.6	-8	-4.6	<0.001
Emotional Function (EF)	-18.6	-8.7	-9.9	<0.001	-17.4	-10.1	-7.3	<0.001

Improvements from baseline favouring Botox were statistically significant for all quality of life outcomes for HIT-6 and MSQv2.1 for both sub-groups, as shown in Table 8. These included the proportion of subjects with severe HIT-6 category scores, all 3 domains of the MSQ, total HIT-6 scores and the incidence of subjects with > 5 point improvement from baseline on total HIT-6 score.

The mean change in total HIT-6 score exceeded the minimally important differences (MID) between groups for the BASH prior first-line prophylaxis group, but not for those who had not used BASH first-line prophylaxis.

The MID between groups were reached in the RR and RP domains of MSQ in both sub-groups, except for the EF domain where the MID of 7.5 between groups was not reached for group that had not used BASH first line prophylaxis.

The implication is that prophylaxis according to BASH guidelines should have been tried before Botox is considered, which provides further justification for the NICE position that Botox should be used only after three other prophylactics had been tried.

Subgroup: medication over-users

Table 34 from the AusPAR report[42] presents an analysis of the pooled results for the two PREEMPT studies for subjects with and without prior medication over-use. (The table is reproduced below in Table 9)

The total numbers with medication overuse were 904 (65%) (n=445 in Botox group and 459 in placebo group) and for no such over-use, 480 (35%) (n=243 Botox and n=237 placebo)

Table 9 Sub-group analysis for patients with medication overuse

	Baseline		**Week 24**			
	Botox	Placebo	Botox	Placebo	Botox - pbo	**P-value**
HA days						
Medication overuse	20.1	19.8	-8.2	-6.2	-2	<0.001
No medication overuse	19.6	19.7	-8.8	-7.3	-1.5	0.013
Overall	19.9	19.8	-8.4	-6.6	-1.8	<0.001
HA episodes						
Medication overuse	12.8	13.8	-5.4	-4.9	-0.5	0.028
No medication overuse	10.9	11.4	-5	-4.6	-0.4	0.146
Overall	12.2	13	-5.2	-4.9	-0.3	0.009
Migraine/probable migraine days						
Medication overuse	19.3	19.1	-8.1	-6	-2.1	<0.001
No medication overuse	18.8	18.5	-8.4	-6.6	-1.8	0.004
Overall	19.1	18.9	-8.2	-6.2	-2	<0.001
Moderate/severe HA days						
Medication overuse	18.5	18.4	-7.7	-5.7	-2	<0.001
No medication overuse	17.4	17.3	-7.8	-6	-1.8	0.005
Overall	18.1	18	-7.7	-5.8	-1.9	<0.001

The medication overuse group shows a slightly greater difference between Botox and placebo compared to the non-overuser for all outcomes i.e. headache (HA) days, HA episodes, migraine/probable migraine (MPM) days and cumulative hours of HA on HA days.

Allergan provided some additional data in response to an ERG question, showing that medication over-use fell in both arms of the PREEMPT trial. The pooled data showed falls in the proportions

over-using acute medications, from 65% at baseline for Botox to 27% at week 24, and from 66% to 33% for the placebo group. The confidence intervals for the week 24 figures overlap.

Implication: The NICE scope suggests that Botox should be used only in patients whose medication over-use has been appropriately addressed, presumably because headaches may be due to medication over-use. However if medication over-use, which is very strictly defined (two paracetamol tablets taken on 16 days a month would qualify), is simply due to frequency of headaches, then the results above suggest that Botox might be seen as an appropriate part of management.

Subgroups: age, gender, ethnicity, prior prophylaxis

Table 10 (adapted from Table 10 of the MHRA report) shows that for the main outcome of headache days, there was no significant difference between sub-groups based on age < 40 years versus ≥ 40 years. However there were significant differences based on gender, race and history of medication prophylaxis use. Males, non-Caucasians and those without a history of headache medication prophylaxis did not show a significant difference in headache days at 24 weeks.

Table 10 Sub-analysis: baseline and mean change at week 24 in headache days

	Pooled PREEMPT Studies						
Subgroup Categories	Baseline			Week 24			P value
	Botox	Placebo	Botox-placebo	Botox	Placebo	Botox-placebo	
Male Botox n= 85 Placebo n= 103	20.5	20.4	0.1	-7.9	-7.2	-0.7	0.479
Female Botox n=603 Placebo n = 593	19.8	19.7	0.1	-8.5	-6.5	-2	<0.001
Race							
Caucasian Botox n= 617 Placebo n =630	19.9	19.8	0.1	-8.4	-6.5	-1.9	<0.001

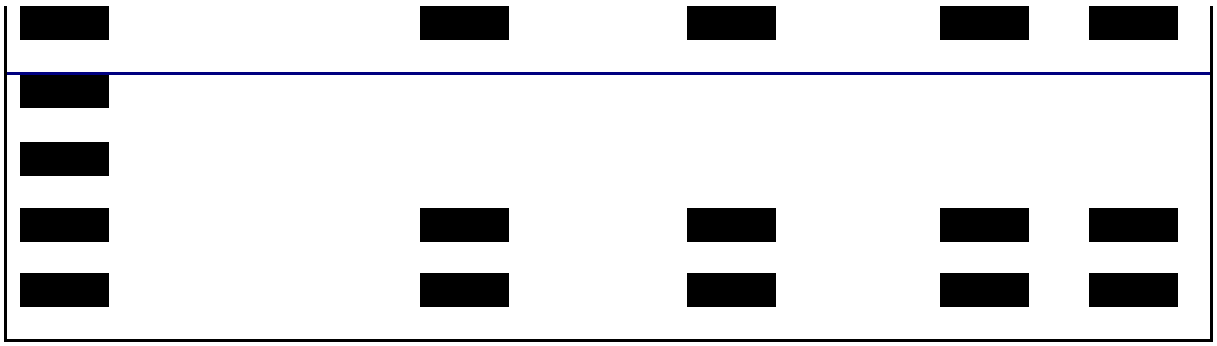
Non-Caucasian Botox n = 71 Placebo n= 66	20.2	19.7	0.5	-8.6	-7.1	-1.5	0.258
History of Headache Medication Prophylaxis							
Yes Botox n=425 Placebo n = 454	20.1	20.1	0	-7.9	-5.6	-2.3	<0.001
No Botox n=263 Placebo n =242	19.7	19.3	0.4	-9.2	-8.3	-0.9	0.197

Subgroups by number of previous oral prophylactics used

Table 5.31 from the Allergan submission (adapted below as Table 11) shows the mean change from baseline for key efficacy outcomes in the pooled analysis of the PREEMPT studies for patient subgroups with increasing levels of previous oral prophylactic medications.

Table 11 Results in pooled analysis for patient sub-groups based on previous oral prophylactic medications

Frequency of headache days (per 28 days)	**BTX**	**PBO**	**Botox-placebo**	**p-value**
[redacted]				
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]



In all the subgroups there is a small difference of [redacted] to [redacted] days favouring Botox, reaching statistical significance for all groups. Botox is effective in all sub-groups based on the number of previous oral prophylactic medication and in those who have previously received topimarate. It can be seen that as the number of previous oral prophylactic medications rises, so does the relative effectiveness of Botox versus placebo, with all the differences between Botox and placebo in the groups with prior prophylactic use being greater than the 1.8 days for all patients in the pooled PREEMPT trials (i.e 8.4 minus 6.6). It also seems that the placebo effect decreases as the number of previous medications increases.

Note to Committee.
On page 92, in the paragraph headed "History of oral prophylactic use", it is stated that subgroup analyses were conducted on patients who had failed to respond to 3 or more oral prophylactics. The text states that Botox was more effective than placebo in "this sub-group (Table 5.30)".
However the figures in Table 5.30 are too large to be just for the 3 or more orals group. The figures given are 420 and 437 = 857, whereas table 5.31 states that the numbers who failed on 3 or more oral drugs were 231 and 248 = 479. The figure of 857 is closer to the figure of 879 in table 5.31 for all patients who had previously received one or more prophylactic drugs. However it is not clear where the figures in table 5.30 come from. The headings of BTX/BTX and PBO/BTX also seem wrong.

Problem: how do we define failure on three or more oral prophylactics? How good is adherence?

Rahimtoola et al [43] evaluated migraine prophylactic medication usage patterns in The Netherlands. Usage patterns in ergotamine and triptan patients commencing treatment for the first time between the years 1992 to 1998 were determined by accessing data from a large prescription database. Patients were characterized as continued, switch or stop use during the patient observation period.

A total of 729 first time users of ergotamine or a triptan had commenced migraine prophylactic treatment following the use of these drugs during the study period. The median duration of migraine prophylaxis was 2.8 months (range 1 day to 6 years), and beta-blockers were by far the migraine prophylactic drugs of first choice for both general practitioners and neurologists.

They found that more than half of the study population had discontinued migraine prophylaxis within 3 months and after one year 74% of the study population had discontinued therapy (stop or switch). A minority, 15%, had demonstrated prolonged adherence of over 2 years.

The type of usage pattern displayed was dependent on the type of migraine prophylactic drug used. Overall, patients undergoing treatment with flunarizine or methysergide were inclined to have discontinued therapy (either stop or switch) more rapidly compared with beta-blockers (RR 1.51; 95% CI 1.13–2.0 and RR 2.02; 95% CI 1.11–3.71, respectively. Age < 40 years (RR 1.9; 95% CI 1.2-3.2) and the concomitant use of non-steroidal anti-inflammatory drugs (RR 3.2; 95% CI 1.2-5.5) or specific abortive migraine drugs resulted in a faster onset of treatment modification (switch).

A review of headache treatment compliance by Rains et al[44] in 2006 concluded that a quarter to a half of patients are noncompliant with prophylactic headache medications and at least 40% are non-adherent with appointment-keeping. Subjective reports of adherence were likely not only to overestimate but also to be discordant with more objective measurements.

Risk factors for noncompliance included:

- Dosing regimens. More frequent and complex dosing regimens decreases compliance.
- Demographic characteristics. The strongest overall predictive demographic variable is probably age, with compliance generally improving with age into the senior years.
- Side effects and costs. Higher medications side effects and costs decrease compliance. Patients are believed not only to reduce their medication intake to diminish the risk of side effects, but also to discover the lowest effective dosage.
- Psychosocial influences. Depressed patients were found 3-fold more likely to be noncompliant than non-depressed patients. The greater the number of psychiatric diagnoses, the lower the adherence.
- Cognitive variables. Cognitive constructs of "self-efficacy" and "locus of control" have been associated with adherence. Self-efficacy, or lower belief that one's actions can impact outcome, correlates with poorer compliance

A retrospective cohort study by Yaldo et al[45] described persistence with migraine prophylactic treatment in patients in one of the largest managed care organizations in the United States. They analysed an integrated research database which provided pharmacy/medical claims data from 5 commercial health insurance plans on adult patients with migraine.

The prophylactic medication identified at index date was used for categorizing patients into 1 of 4 cohorts: amitriptyline, propranolol/timolol, divalproex sodium, or topiramate (reference).
A total of 12,783 patients met the inclusion criteria and were included in the analysis (amitriptyline, 3749; propranolol/timolol, 2718; divalproex sodium, 1644; and topiramate, 4672).

They found that topiramate prescription was associated with greater persistence than prescription of other migraine prophylactic treatments considered in this study ($P < 0.05$). The risks for discontinuing prophylactic treatment were 23%, 6%, and 11% higher with amitriptyline, propranolol/timolol, and divalproex sodium, respectively, compared with topiramate ($P<0.001$, $P = 0.024$, and $P <0.001$). Greater reductions in acute treatment utilization, particularly triptans, were observed among patients prescribed topiramate compared with the other prophylactic cohorts.

The influence of disease features on adherence to prophylactic migraine medication was examined in a cross-sectional study conducted at an outpatient setting at a headache specialist clinic in Sweden by Linde et al[46] A sample of 174 adult migraineurs with a current prescription of pharmacological prophylaxis responded to a questionnaire on adherence to medication. Adherence was self-reported with the Medication Adherence Report Scale (MARS), which consists of five statements about the use of medicines.

One-third (35%) were non- adherent. In a logistic regression analysis, there was no significant association regarding adherence and sex, age, or educational level. Neither demographic characteristics nor any of the disease specific variables were significantly associated with adherence. The mean duration of disease was 26.5 years for adherent migraineurs and 27.3 years for the non-adherent ($p=0.70$), so non-adherence was also a problem among patients with severe migraine.

Evans and Linde[47] recently explored whether a selection of clinical features, apparent to the physician, influenced adherence in a typical migraine population. They found that neither attack frequency, duration of attacks, degree of recovery between attacks nor cardinal symptoms during attacks were significantly associated with adherence. Thus, non-adherence is also a problem among patients with severe migraine.

They reported that a study based on patients in Gothenburg showed tha beliefs about medicines in general did not contribute significantly to adherence in a typical migraine population receiving standard care. Some specific treatment predictors, such as the cost and a satisfactory efficacy/tolerance ratio are essential determinants, which may in part explain why their users of beta-blockers were significantly more adherent than users of amitriptyline. Also, continuous reinforcement from the prescribing physician seems important to maintain adherence

They feel that greater insight into individual needs and fulfilment of the patient's expectations is a prerequisite for successful outcome of prophylaxis. For those patients who do not wish to commence or continue a course of treatment, despite all efforts by the physician, they would consider non-pharmacological treatment, because it may be equally viable to drugs in prophylaxis for migraine and many patients appreciate and prefer "low-tech-high-touch" medicine. Non-pharmacological treatment is considered to activate endogenous, positive mechanisms, and to offer "empowerment"

The BASH guidelines emphasise the need for an adequate period of trial on prophylactic drugs.[25] Problems may occur in practice due to the fact that headaches are self-reported.

4.2.8 Responder analysis

The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) group has emphasised that the interpretation of the meaningfulness of results of randomized clinical trials of treatments for chronic pain involves two distinct processes - interpreting the clinical importance of: 1) individual patient improvements, and 2) group differences.[48]

Group differences are often expressed as means, but individual improvements are often expressed as the proportion of patients achieving a good response, with decreases in pain of 50% or more being considered substantial and such patients classed as responders. The FDA (as quoted in Dworkin et al[48] - the cited report has now been superceded) note that,

"*When defining meaningful change on an individual patient basis (i.e responders), that definition is generally larger than the minimum important difference for application to group mean differences".*

The FDA also says;

"It may be more informative to examine the distribution of responses between treatment groups to more fully characterise the treatment effect."

Dworkin et al argue that such analyses may make it possible to determine whether a sub-group of patients may experience meaningful benefits, even when the overall mean difference is small.

The guidelines of the International Headache Society comment that responder rates may be included as secondary endpoints.[12] These can be defined as either ≥30% or ≥50% reduction in i) headache days with moderate or severe intensity , ii) migraine days, or iii) migraine episodes compared with baseline

period. They say that responder rates have been traditionally defined in migraine as ≥50% reduction, but in a CM population a ≥30 % responder rate can be clinically meaningful.

In the pooled PREEMPT trials, the proportions that achieved a 50% or more reduction in secondary outcome at 24 weeks are show in Table 12. The data is taken from Aurora and colleagues.[35]

Table 12 Percent of patients with a 50% decrease in the outcome from baseline at week 24

Outcome	**Botox (%)**	**Placebo (%)**	**Botox-placebo (%)**	**p-value**
Headache days	47.1	35.1	12.0	<0.001
Migraine days	48.2	36.4	11.8	<0.001
Moderate/severe headache days	49.4	37.5	11.9	<0.001
Total cumulative hours of headache on headache days	50.3	38.9	11.4	<0.001
Headache episodes	48.6	43.1	5.5	0.065
Migraine episodes	48.1	43.4	4.7	0.119

The proportion of patients who demonstrated a 50% decrease from baseline was significant for headache days, migraine days, moderate /severe days and total cumulate hours of headache on headache days. The 50% responder analysis was not significant for headache episodes and migraine episodes.

Dodick et al in a recent meeting abstract reported a 75% responder analysis in the pooled PREEMPT population. At week 24, 22.8% in the Botox group versus 15.5% in placebo group (p=0.002) had a >75% reduction in headache days from baseline.[49] Also, they reported a significantly greater proportion showing a 75% reduction in headache episodes, migraine days, migraine episodes, moderate/severe headache days, and total cumulative hours of headache on headache days (data not given).

There were some differences in responder rates by baseline headache states. Taking response as an improvement by 2 headache states;

- 15-19 days – 52% were responders
- 20-23 days – 44.5%
- 24-28 days – 46.4%

Overall, in the Botox group, 49% were responders

4.2.9 Quality of life measures

Headache Impact Test (HIT-6)

The Headache Impact Test (HIT-6) is a validated questionnaire for discriminating headache impact across episodic and chronic migraine.[50] It comprises six items, each scoring between 6 (never) and 13 (always). The range is 42, so the lowest score is 36 (no impact), and the worst is 78 (most severe). A total score of ≤49 indicates little or no impact, 50-55 indicates some impact, 56-59 indicates substantial impact, and ≥ 60 reflects severe impact

Minimally important differences (MID) in reduction in total HIT-6 scores have been previously estimated by Coeytaux et al.[51] They applied four methods of estimating the MID to data from 71 patients who participated in a clinical trial. Their data suggested that a difference in HIT change scores of 2.3 units (95% CI -4.3 to -0.3) over time between two groups of patients with chronic daily headache reflects improvements in patients' headache condition that may be considered clinically significant.

They also suggested that the change in HIT-6 scores associated with clinical change sufficiently to prompt a patient to report feeling "somewhat better" (i.e. within-person change) is ~3.7 units

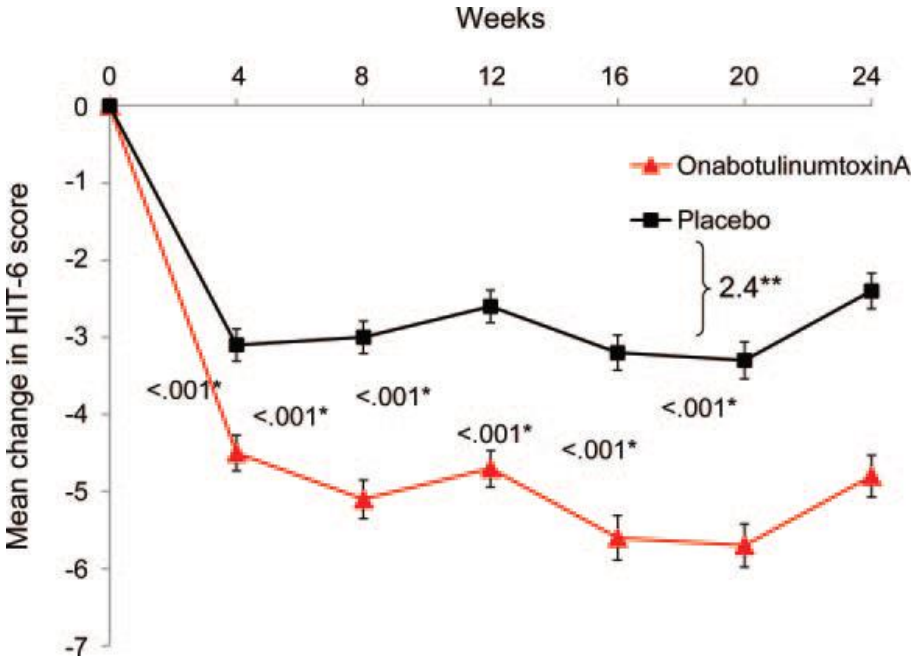
Lipton et al report that a clinically meaningful change for an individual patient has been defined by the user guide for the HIT-6 scale as an increase of 5 or more units.[52]

The PREEMPT data for change in HIT-6 scares are shown in Table 13

Table 13 Baseline scores and change from baseline at week 24 in HIT-6 scores

	Botox (n=688)		Placebo (n=696)		Botox-placebo at week 24	*P* value
	Baseline	Change from baseline at week 24	Baseline	Change from baseline at week 24		
Mean HIT-6 score	65.5	-4.8	65.4	-2.4	-2.4	<0.001
Proportion of patients with a severe (≥60) HIT-6 score	93.5%	67.6%	92.7%	78.2%	10.6%	p<0.001
Proportion of patients with a ≥5 point decrease in total HIT-6 score at 24 weeks		40.8%		25.3%	15.5%	p<0.001

Neither the Botox or placebo group reached the MID of 5 for changes within groups for the mean change in HIT-6 score at week 24. However Figure 7 below (taken from Lipton 2011 [52] shows that there was a statistically significant difference between groups at weeks 4, 8, 12, 16, 20 and 24, with a borderline clinically meaningful (according to Coeytaux et al[51]) between group difference of -2.4 favouring Botox at 24 weeks.



**Significant between-group difference favoring Botox A. **Clinically meaningful between-group difference. Data are mean ± SE.*

Figure 7 Mean change from baseline in total Headache Impact Test (HIT)–6 score

The baseline mean total HIT-6 score was 65 and approximately 93% of patients were severely affected. Most of the change occurred in the first 4 weeks, with very little gain after that relative to week 24. This might be further support for a negative stopping rule after the first set of injections.

If we accept the recommendation by Lipton et al[52], that a meaningful change is 5 or more, then the mean change in HIT-6 score within groups (-4.8 for Botox and -2.4 for placebo) did not reach the MID of ≥5 points i.e. a change that is clinically meaningful for an individual patient.

However, as shown in Figure 8 (taken from[52]) the percentage of patients with a clinically meaningful change in HIT-6 was significantly higher for Botox than placebo at all of weeks 4, 8, 12, 16, 20 and 24. At 24 weeks the proportions were 40.8% vs 25.3% respectively ($p<0.001$). Therefore 59.2% of the Botox group did not show a clinically meaningful improvement in HIT-6 score.

Significantly fewer patients in the Botox group had a severe (≥60) HIT-6 score at all weeks during the double blind phase, and at week 24 the difference of 10.6% was significant (p<0.001).[52]

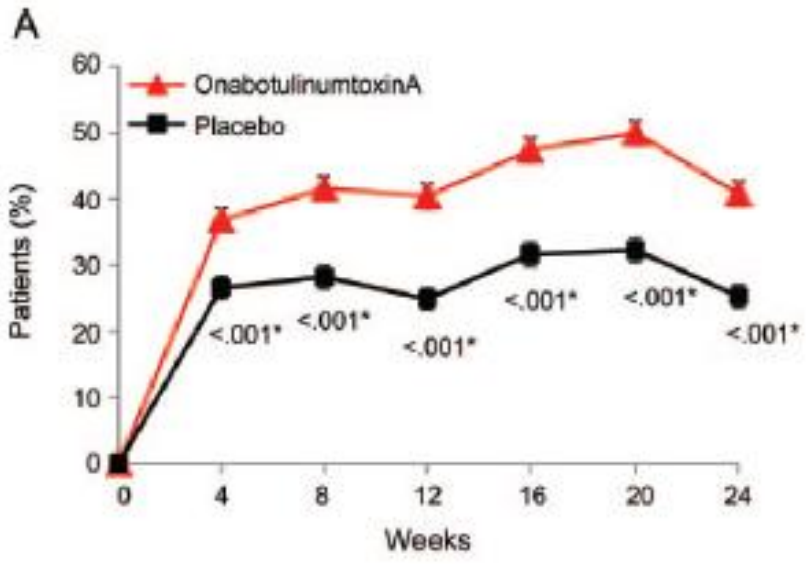


Figure 8 Percentage of patients with a clinically meaningful change in Headache Impact Test (HIT)–6 (>5 points)

Figure 9 below (taken from Lipton et all [52] shows that the percent of patients who were responders (≥50% decrease in number of headache days from baseline) and had a clinically meaningful change in HIT-6 score (≥5 points) at 24 weeks was ~26% in Botox group and ~15% in placebo group (data read from graph) . This difference was significantly higher at all time points measured from week 4 onwards. Therefore, about 75% were not 50% responders with a clinically meaningful improvement.

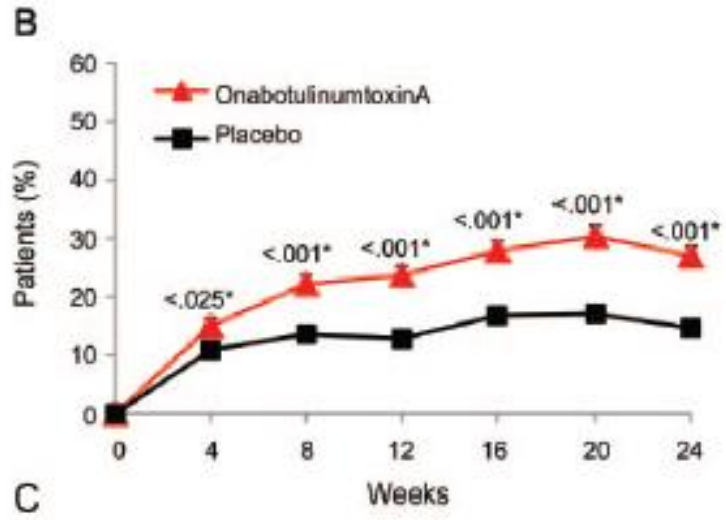


Figure 9 Percent of patients who were responders (≥50 decrease in number of headache days from baseline and a clinically meaningful change in HIT-6 score (≥5 points)

Figure 10 (taken from the Allergan submission) shows the mean change in HIT-6 scores in the double-blind and open-label phase

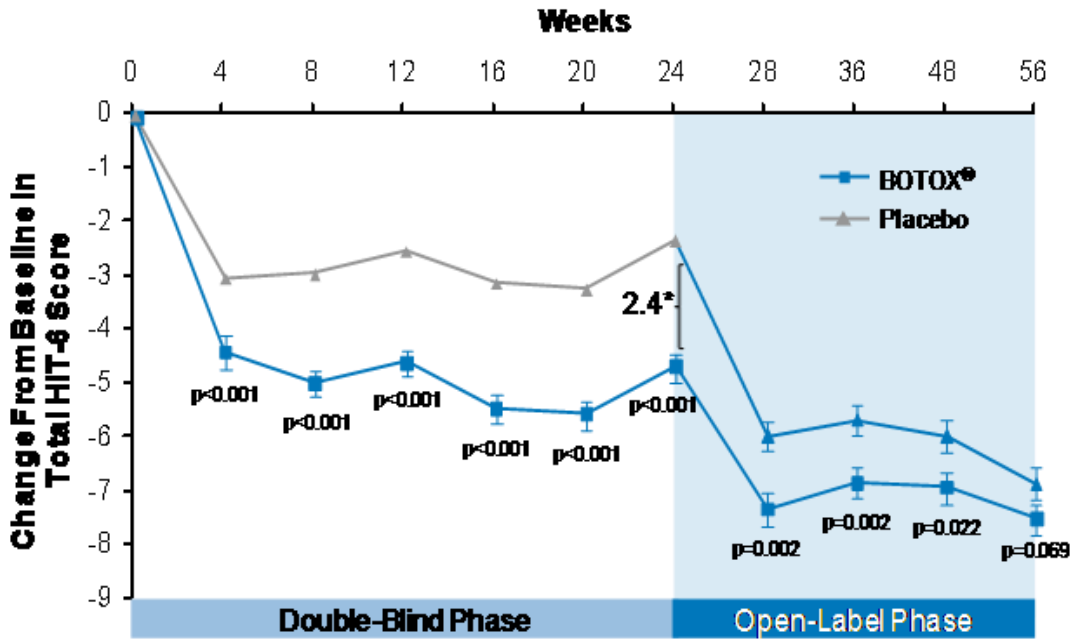


Figure 10 Mean change from baseline at week 24 for HIT-6 results in pooled phase 3 studies (ITT population) (from Allergan submission)

Note the large improvement in HIT-6 score in both groups after first 4 weeks. Note also the reduction between 24 and 28 weeks in the open label phase, which is much larger than the decrease after the 2nd injection at 12 weeks (and after injections at 36 and 48 weeks).

One implication of the lack of change after the 12 week injection is that the effect of Botox lasts longer than 12 weeks, and that it could be given less frequently. A counter-argument to that might be that the effect of the 12-week injection is to maintain the benefits from the baseline injection, but the drop after the 24-week injection might negate that argument in that it shows further reduction not just maintenance.

The difference in effect between the 12 week and 24 week Botox injections needs explained. One explanation might be placebo effect?

MSQ v2.1 (Migraine Specific Quality of Life Questionnaire)

The Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ v2.1) is designed to measure how migraines affect and/or limit daily function across 3 domains; role restrictive (RR), role preventive (RP) and emotional functioning (EF). The questionnaire comprises 14 items across the 3 domains, (7 in RR, 4 in RP, 3 in EF). Raw scores are rescaled to a 0 to 100 scale. The scores range from poorest 0 (poorest) to 100 (best), so an increase in score reflects improvement in HRQoL. [53]

A reliability analysis of the MSQ scales amongst migraine sufferers reported them to be moderately to highly correlated with HIT-6 (*r*= -0.60 to -0.71).[53]

The minimally important differences for changes from baseline between groups for MSQ v 2.1 vary amongst domains, and have been established as RR= 3.2, RP=4.6, EF=7.5.[54] Changes within-groups (individual level) have been established as: RR= +10.9, RP = + 8.3, EF = +12.2.[55]

Table 14 below (taken from data taken in Lipton et al 2011[52]) shows that the baseline MSQ scores were similar between the groups for all domains. The between-group differences in MSQ at week 24 were RR 8.4, RP 6.7 and EF 8.4. Therefore, they all exceeded the MIDs

Table 14 Baseline scores and change from baseline at week 24 in MSQ v2.1 scores

MSQ domains	Botox (n=688)		Placebo (n=696)		Botox-placebo	*P value*
	Baseline	Change from baseline at week 24	Baseline	Change from baseline at week 24		
Role restrictive	38.5	17	38.7	8.6	8.4	<0.001
Role preventive	56.0	13.1	56.1	6.4	6.7	<0.001
Emotional functioning	42.1	17.9	42.4	9.5	8.4	<0.001

The MID for the changes within groups were reached for all MSQ domains for the Botox group, but in none of the placebo groups. However what matters is whether the final differences between the Botox and PBO scores exceeded the clinically meaningful differences.

Table 15 (taken from the response to question A7 in the manufacturer response to clarification questions from the ERG) below shows the percentage of patients having clinically meaningful changes for the MSQ domains

Table 15 Percent of patients having a clinically meaningful change for MSQ domains

[redacted]	[redacted]						
[redacted]	[redacted]			[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]						
[redacted]	[redacted]			[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]						
[redacted]	[redacted]			[redacted]			
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The data shows that the percentage of patients having a clinically meaningful change from baseline within groups was [redacted] for Botox than placebo for each domain of MSQ at week 12 and week 24. Numbers at 24 weeks for Botox vs placebo were: [redacted].

The data for clinically meaningful changes in MSQ domains are also shown in graphical form below in Figure 11, Figure 12, and Figure 13. (These were also taken from taken from the manufacturer response to clarification questions from the ERG.)

Figure 11 [redacted]

Figure 12 [redacted]

Figure 13 [redacted]

4.2.10 Safety

In the PREEMPT trials, adverse events were common, but most were not treatment related. Treatment related adverse events (i.e. events that the investigator thought might have related to study medication) were reported in 29% of those on Botox and 13% of those on placebo. However only one serious treatment-related event was report (in the Botox group), and the numbers of people discontinuing treatment were quite small – 3.8% (26) in the Botox group and 1.2% (8) in the placebo group.

Table 9 of the FDA prescribing information provides a useful summary of adverse events[56]. This is reproduced in

Table 16 below.

Table 16 Adverse Reactions Reported by >2% of Botox treated patients and more frequent than in PBO-treated Patients in the PREEMPT Trials

Adverse Reactions by Body Systems	**BOTOX 155 Units-195 Units N=687)**	**Placebo (N=692)**
Nervous system disorders		
Headache	32 (4.7%)	22 (3.2%)
Migraine	26 (4%)	18 (3%)
Facial paresis	15 (2%)	0 (0%)
Eye Disorders		
Eyelid ptosis	25 (4%)	2 (<1%)
Infections and Infestations		
Bronchitis	17 (3%)	11 (2%)
Muscoloskeletal and connective tissue disorders		
Neck pain	60 (9%)	19 (3%)
Muscoloskeletal stiffness	25 (4%)	6 (1%)
Muscular weakness	24 (4%)	2 (<1%)
Myalgia	21 (3%)	6 (1%)
Muscoloskeletal pain	18 (3%)	10 (1%)
Muscle spasms	13 (2%)	6(1%)
General disorders and administration site conditions		
Injection site pain	23 (3%)	14 (2%)
Vascular disorders		
Hypertension	11 (2%)	7 (1%)

However, outwith the trials, concern has been raised about dysphagia with Botox treatment (though not for migraine as yet), and the FDA has issued the following statement (box). [56]

WARNING: DISTANT SPREAD OF TOXIN EFFECT

Postmarketing reports indicate that the effects of BOTOX and all botulinum toxin products may spread from the area of injection to produce symptoms consistent with botulinum toxin effects. These may include asthenia, generalized muscle weakness, diplopia, ptosis, dysphagia, dysphonia, dysarthria, urinary incontinence and breathing difficulties. These symptoms have been reported hours to weeks after injection. Swallowing and breathing difficulties can be life threatening and there have been reports of death. The risk of symptoms is probably greatest in children treated for spasticity but symptoms can also occur in adults treated for spasticity and other conditions, particularly in those patients who have an underlying condition that would predispose them to these symptoms. In unapproved uses, including spasticity in children, and in approved indications, cases of spread of effect have been reported at doses comparable to those used to treat cervical dystonia and at lower doses. [See Warnings and Precautions (5.2)]

This follows a few accounts of problems such as dysphagia. Rossi et al[57],Thobois et al [58] and Tuite and Laang[59] all reported single cases of dysphagia after the use of Botox in cervical dystonia. Kwek et al [60] reported a case of dysphagia after Botox injected into the orbicularis oculi muscle. Comella et al [61] reported dysphagia after Botox treatment in a series of 18 patients with spasmodic torticollis, a third of whom developed dysphagia after Botox treatment.

Allergan are sponsoring a 15 month observational study in the UK, to monitor the use of Botox, and to collect further safety data, in particular on the incidence of dysphagia. Patients will have Botox injections for migraine, using the 155-195 PREEMPT regimen, at 12-weekly intervals, for one year. Brief details are on the clinical trials website.[62]

Chapman et al carried out a review of the literature on dysphagia in people with dystonia treated with Botox, and reported dysphagia rate round 10% with Botox.[63] They also noted that rates were higher with the Dysport form of Botox, where higher numbers of units were used. There are several forms of botulinum neurotoxins on the market, and they are not interchangeable in terms of units.

4.3 Placebo effects

One of the striking features in the PREEMPT trials was the size of the placebo effect – a reduction in headache days which was not far short of the reduction amongst those on Botox. The mean change from baseline in the primary outcome, frequency of headache days for the 28 day period ending with week 24, was 8.4 days for Botox and 6.6 days for placebo. Therefore, the reduction in frequency of migraines in Botox group is 44.2% (8.4/19.9) compared to placebo 33.3% (6.6/19.8), giving a 10.9% difference between groups in favour of Botox.

This strong placebo effect has been noted in previous trials of Botox for various forms of headache.

Mathew et al carried out an RCT of Botox (105 to 260 units) compared to placebo injections in chronic daily headache (about 61% migraine).[64] The Botox group had a mean reduction of 6.7 headache days a month, but this was not statistically significant because the placebo group improved by 5.2 days. The placebo response was sustained for at least 9 months.

Silberstein et al compared three doses of Botox (75, 150 and 225 units) with placebo.[65] They reported that the placebo response was higher than expected – there were mean improvements in headache free days of 6.0, 7.9 and 7.9 for Botox 225 U, 150 U and 75 U respectively, but an improvement of 8.0 days for the placebo group, and these results were in groups that did not respond to placebo in a run in phase.

There have been several useful reviews or meta-analyses. De Craen et al carried out a meta-analysis of 22 trials to determine the comparative placebo effect of subcutaneous vs. oral administration in the treatment of migraine.[66] The headache relief rates were combined from the placebo arms of randomised clinical trials assessing the value of sumatriptan in acute treatment of migraine.

The main outcome measure was the proportion of patients reclassified from severe or moderate headache to no or mild headache severity two hours after the beginning of treatment. Adequate pain relief was achieved by 32% (279/862) of the group receiving subcutaneous placebo and by 26% (222/865) of those receiving oral placebo (6.7% difference; 95% CI 2.4-11.0%). This difference remained after adjusting for a number of possible confounding variables.

Hence, at least in acute migraine, subcutaneous administration gives a greater placebo effect than oral placebo.

De Craen et al hypothesised that patients who had subcutaneous placebo injection may have felt they were receiving more powerful treatment, which could have induced a release of endogenous opiates

Van der Kuy et al noted that the placebo response in migraine ranged from 6% to 44% depending on patient characteristics and trial design, and was difficult to quantify in open-label trials.[67] They therefore carried out a meta-analysis of double-blind, placebo controlled migraine trials (going back to 1978) to quantify the placebo response of prophylactic therapy in migraine studies. They included 22 studies, in which 19 different products were compared with placebo. Their analysis included 2013 patients, of which 828 (30%) were treated with placebo. All the drugs and placebos were oral.

A reduction in migraine attacks of 50% or more (i.e. 50% responders) was seen in 23.5% (95% C.I. 18.3-28.8%) of the patients in the placebo groups and 45.5% (95% C.I. 37.4-53.6%) in the active groups. A reduction in migraine attacks of 16.8% (95% C.I. 10.9-22.6%) was observed in the placebo groups and 41.8% (95% C.I. 36.9-46.6%) in the active groups.

Macedo et al did a meta-analysis to evaluate the placebo response rate in migraine prophylaxis in all published clinical trials since 1988 and to estimate the influence of study design in response variability.[68] Clinical trials had to meet the following criteria to be included in the present analysis: comparison of an oral, active drug with placebo for migraine prophylaxis; randomised and double-blind clinical trials; description of results for both groups (active drug and placebo); be published in English; publication date after 1988 and migraine diagnosis according to IHS criteria. The outcomes studied were proportions of patients who improved (reduction in migraine attacks of 50% or more); attacks per month, and patients with adverse events.

32 papers were considered. The pooled estimate of the placebo response (percentage of patients under placebo who had a greater than 50% reduction in migraine attacks) was 21% (95% CI, 13 to 28%).

The placebo response rates were significantly higher in studies with a parallel design (22%) than those in cross-over studies (10%) ($p<0.01$). The placebo response was also higher in European studies (25.4%) than in those performed in North America (16.8%) ($p<0.0001$).

Adverse events occurred in 30% of the patients who took a placebo, and the percentage of patients with adverse events was significantly higher in the North American studies (63%) than in those conducted in Europe (22%) ($p<0.002$).

Macedo et al also carried out a meta-analysis of the placebo response in acute migraine, and how it may be influenced by some of the characteristics of clinical trials.[69] Clinical trials had to meet the following criteria to be included: (1) comparison of an active drug with placebo for the treatment of acute migraine; (2) randomized and double-blind design; (3) accurate description of the results for both groups [active drug(s) and placebo]; (4) published in English after 1988; (5) diagnosis of migraine in agreement with The International Headache Society criteria. They included 98 papers published between 1990 and 2004, with a total of 35,481 patients, 11,793 in the placebo group.

After 2 hours, 28.6% of the patients in the placebo group improved and 8.8% were pain-free. The percentage of pain-free patients was the highest in the placebo and active drug groups in which the placebo or drug had been administered subcutaneously, in parallel design studies (vs. cross-over trials) and in studies performed in Europe (vs. North America). Adverse events in the placebo group were significantly higher in studies performed in North America. 29% of those in the placebo groups reported improvement. When active treatment was administered, 58% of patients improved and 29% were pain-free.

Several other studies have assessed the effect of the method of administration on the size of the placebo effect.

Kaptcheck et al compared sham acupuncture injections with an oral placebo in a randomised controlled trial of two placebo treatments, in patients with persistent arm pain due to repetitive strain injury.[70] There was significantly greater improvement in the arm pain scale in the sham acupuncture group than in the placebo pill group (-0.33 v -0.15, $P= 0.0001$) and on the symptom severity scale (-0.07 v -0.05, $P = 0.02$).

Diener et al noted that there was greater variability in studies of prophylaxis of migraine that in studies of treatment of acute attacks.[71] They suggested that this is probably due in part to the different primary end-points used in studies of migraine prophylaxis, and to the inherent variability in response measured over a period of months compared with one measured over a period of hours. They noted that response rates to sham acupuncture were higher than for sham oral drug treatment.

Speciali et al argue that there are three reasons for improvement in migraine trials;[72]

- spontaneous improvement including regression to the mean of people recruited while in severe headache states
- the placebo effect
- the therapeutic effect of the active intervention

They also suggested that the site of injection might affect the placebo response, with injections around the site of the pain in head and neck having a greater placebo response than e.g. systemic injections. Schwedt et al also reported that injections in neck and shoulders increased placebo response.[73]

Solomon reviewed the placebo effects as seen in the Botox for chronic migraine trials.[74] He hypothesised that the placebo effect might be responsible for all the effects seen. He argues that several factors that increase the placebo effect applied in the Botox trials;

- the increased attention in the trial
- regression to the mean or spontaneous improvement
- the injected placebo - invasive placebos have more effect
- the number of injections – minimum of 31
- unblinding because of the rapid relaxation of forehead muscles, resulting in reduced wrinkling

Solomon notes that the PREEMPT authors argue that "*The presence of a placebo effect suggests that the blind was maintained"* but points out that unlike in some other trials, the patients were not asked at the end what they thought they had received. In the trial by Mathew et al comparing Botox and placebo, patients were asked what they thought they had received, and after the first injection; 70% guessed correctly.[64] So the majority of patients were unblinded. The guidelines for trials of prophylaxis in chronic migraine recommend that subjects and investigators be asked at the end of trials, what treatment they thought they had received or administered.[12] This was not done in the PREEMPT trials.

This raises the possibility that the small difference seen in the PREEMPT trials arose because of a differential placebo effect – i.e. both groups had a placebo effect, but because some or most of the Botox group realised they had had Botox, they had a greater placebo effect. As noted at the start of this section, the difference in the reduction in migraines was only 11%. It might not take many Botox patients to be unblinded, to explain that difference. Conversely, as Solomon argues, some placebo patients might have realised that they had been given placebo and had a "nocebo" effect, being disappointed and therefore reducing their placebo effect.

Olesen has also argued that the facial muscle weakness resulting from Botox could lead to unblinding, and could explain the 10% difference between Botox and placebo.[75]

Interestingly, there was no difference between groups in the use of acute rescue medications.

The PREEMPT investigators acknowledged that the marked placebo response in the trials led to problems of interpretation, but that that response implied that the placebo group remained unaware of what treatment they had had, and so at least they were unblinded.[33] Aurora argued that the injections sites and doses for Botox in PREEMPT differed from those used for cosmetic purposes.[35]

The ERG wonders if it would take only a small percentage in the Botox group to be unblinded and report a better response, and a small percentage in the placebo group to be unblinded, disappointed and report more headaches, to cause the small difference in self-reported headache days.

5 Cost-effectiveness

5.1 The costs of migraine

The International Burden of Migraine Study

The industry submission uses data from this study to assess the costs of migraine, and hence the reduction in costs that could be achieved if the frequency of migraine was reduced. This section gives details of the IBMS study and examines some of the assumptions used in the Allergan submission.

Details of the study design are given by Payne and colleagues.[76] The IBMS recruited patients with migraine to a web-based survey. Participants had to have an email address. Recruits had to meet the ICDG02 criteria for migraine. 11897 did so, and 82% of these completed the survey. Of those who participated, 5.7% had chronic migraine (CM: headaches on 15 or more days per month with at least 8 headache days with features typical of migraine).

The 9715 recruits came from ten countries: Australia, Brazil, Canada, France, Germany, Italy, Spain, Taiwan, the UK and the USA. Their care and therefore resource use would be expected to vary amongst countries. The proportion classified as having CM varied, from 1.2% in Taiwan to 10.7% in Australia. The figure for the UK was 5.3%.

The study was funded by Allergan, and the company was involved in all aspects of the study, including analysis and writing up.

The study provides very useful data on the demographics and clinical characteristics of people with migraine. Those with CM had a higher level of co-morbidities that those with episodic migraine (EM), including;

- Depression 34%
- Anxiety 34%
- Chronic pain 27%
- Sinusitis or sinus infection 25%
- Hypertension 20%
- High cholesterol 18%
- Arthritis 17%
- PMS/menstrual problems 15%
- Palpitations or arrhythmias 12%
- Fibromyalgia 12%

(All figures rounded to whole numbers). Only 13% reported having no co-morbidities. 92% described their headaches as severe. Overall, 42% of those with CM had cardiovascular risk factors and 46% had psychiatric disorders.

A second paper from the IBMS provided more details on co-morbidities and disability, but also provides data on resource use.[23] These data came from 9 countries, Brazil not being included. The highest numbers in the CM sample came from the USA (21%), with most other countries providing about 11%. Taiwan provided only 1%.

The paper by Blumenfeld et al provides considerable data on resource use. For example, in the three months prior to the survey, 48% had visited a primary care provider (though this could be just to collect a repeat prescription), 24% had seen a neurologist or headache specialist, 8% had visited an emergency department, and 3% had had a hospital "visit" (which appears to mean admission overnight). However the results are not split by country, and health services vary considerably amongst countries

Resource use was higher in CM than EM, which may not be surprising in view of their higher levels of co-morbidities. The paper does not say that subjects recorded only hospital visits/admissions for headache-related reasons.

A third paper from IBMS reported the costs of care, but this paper makes clear that patients were asked to record health care contacts, including admissions and investigations, for headache-related reasons.[77]

The costs were for subjects in the USA and Canada only. They show marked difference between the two countries, for example in diagnostic testing. Costs in the USA included MRI, CT, EEG and ECG, that are unlikely to be required for the care of long-standing CM. In the USA, 6% of the costs were for diagnostic testing, in Canada 12%. It should be noted that the BASH guidelines for the UK state that;[25]

"*Investigations, including neuroimaging, do not contribute to the diagnosis of migraine"*

The largest cost elements were medications, which made up 72% of CM costs in the USA and 41% in Canada. There were no hospital admissions. There was a curious difference in the reported balance of costs between preventive and acute medications, with a ratio of 3.7 in the USA (i.e. almost 4 times as much spent on acute medications as preventive ones) where the ratio in Canada was much higher, around 10.

The costs also included "other specialist and health professionals visits costs" but no breakdown is given of specialties. Costs were higher in those with co-morbidities particularly for those with vascular disease risk factors and psychiatric co-morbidities.

These papers provide a valuable resource but we have concerns about the applicability of some costs, such as CT and MRI, to routine care in the UK. It is understandable that some patients presenting with severe headaches may cause problems of differential diagnosis, but the patients as described in the NICE scope have had CM for some time (years in the case of patients in the PREEMPT trials), and one would not expect such costs to be necessary.

We also note that in these countries, there were no headache-related hospital admissions. The industry submission (Table 6.18, reproduced below) assumes that there are "hospitalisation" costs, varying amongst bands. The ERG was struck by the large jump in costs between health states 10-14 and 15-19. These costs are presumably averages across all patients. So in Table 6.18, the £3.20 per GP visit in the 0-3 migraines group is total costs of GP visits divided by total number of patients in that health state.

The jump in costs is due mainly to ER visits and hospital admission. The ERG is sceptical that ER visits would be required in the UK, where there is out of hours GP provision.

Table 17 Costs associated with each health state in the primary economic model

Cost per Health State	GP Visits	ER visits	Hospitalisation	Triptan Usage	Cost per 12 week cycle
0 - 3 migraines per 28 days	£3.20	£10.91	£17.51	£6.30	£37.92
4 - 9 migraines per 28 days	£9.60	£25.46	£46.69	£16.98	£98.74
10 - 14 migraines per 28 days	£9.60	£25.46	£46.69	£16.98	£98.74
15 - 19 migraines per 28 days	£18.56	£57.29	£186.78	£24.42	£287.05
20 - 23 migraines per 28 days	£18.56	£57.29	£186.78	£24.42	£287.05
24+ migraines per 28 days	£18.56	£57.29	£186.78	£24.42	£287.05

The costs in the above table are based on the resources use data in table 6.17 of the Allergan submission, as shown in Table 18 below.

Table 18 Resource use associated with each health state in the primary economic model

Health State	GP Visits	ER visits	Hospitalisation	Triptan Usage
0 - 3 migraines per 28 days	0.10	0.12	0.03	1.88
4 - 9 migraines per 28 days	0.30	0.28	0.08	5.07
10 - 14 migraines per 28 days	0.30	0.28	0.08	5.07
15 - 19 migraines per 28 days	0.58	0.63	0.32	7.29
20 - 23 migraines per 28 days	0.58	0.63	0.32	7.29
24+ migraines per 28 days	0.58	0.63	0.32	7.29

The Allergan submission states that;

"In the primary economic model, quantities of resources are taken from the International Burden of Migraine Study (Section 6.4.6), as this data was collected in the study with results presented in Table 6.17. Resource costs are taken from NHS reference costs (Section 6.5.2,) with combined results for these values, with the cost per health state, shown in Table 6.18."

However none of the published papers from the IBMS, mentioned above, give evidence to support these figures. The ERG requested clarification and received the following statement from Allergan;

> *"The Blumenfeld et al. study was referenced to provide a description of the study methods and patient population surveyed in the International Burden of Migraine Study. The resource use by health states has been presented in poster form as the Health Resource Use and Costs for Migraineurs in Scotland and thus is not considered AIC. The full reference for this poster presentation is shown below.*
>
> *Bloudek LM, Hansen RN, Liu L, Batty AJ, Varon SF, Lipton RB, Sullivan SD. Health Resource Utilization and Costs for Migraineurs in Scotland. International Society for Pharmacoeconomics and Outcomes Research 16th Annual Conference, May 21-25, 2011, Baltimore, MD [PND32] "*

This abstract was not referenced in the Allergan submission, and was not included in the pack of references attached to the submission. The details in the abstract are extremely scanty, with no details of number of patients, numbers of admissions, reasons for admissions, or number of "ER" visits. None of the authors are from Scotland – they come mainly from Allergan with one author from Bresmed Health Solutions in Sheffield (who may have produced, or helped with, the Allergan submission to NICE). Furthermore, they contradict the published data from the IMBMS study with hospital admissions being rare in Canada and the USA.

They are also not in keeping with the findings of a very large American study, the American Migraine Prevalence and Prevention (AMPP) Study, in which 359 patients (4.6%) had transformed (i.e. chronic) migraine. The mean number of hospital days a year was 0.261 and the mean number of ER visits was 0.06 per patient per annum (NB not applicable to the UK). Note however that these are averages across all patients with CM, and the 15 and above days per month groups might require more health care.[78]

However in the absence of data to the contrary, and taking into account the published IBMS data from the USA and Canada, the ERG is unconvinced that the level of resource use and costs given in the tables above, are justified.

If we accepted the costs in table 6.17, there would be a large drop in costs of care if Botox worked. The alleged savings would help to make Botox cost-effective.

5.2 .Summary and critique of manufacturer's submitted economic evaluation by the ERG

The manufacturer set out to do a review of the previous cost-effectiveness evidence. The search strategy and proposed inclusions were appropriate. However no published studies were found.

5.2.1 NICE reference case checklist

Attribute	**Reference case and TA Methods guidance**	**Does the *de novo* economic evaluation match the reference case**
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Botox is seen as adjunctive to standard therapy within the model.
Patient group	As per NICE scope: "*Adults with headaches on at least 15 days per month of which at least 8 days are associated with chronic migraine and whose condition has failed to respond to at least three prior pharmacological prophylaxis therapies and medication overuse has been appropriately managed."*	No. Apparently in order to preserve reasonable numbers of patients for the analysis the base case is presented for the subgroup of PREEMPT patients with ≥ 1 prior oral prophylactics. It is also not restricted to the patient group in whom medication overuse has been appropriately managed. This is explored in a sensitivity analysis which analyses the ≥ 3

		prior oral prophylactics subgroup. Analyses are also presented for those not overusing medication as defined within the pivotal trials, though these patient subgroups are quite small.
Perspective costs	NHS & Personal Social Services	Yes.
Perspective benefits	All health effects on individuals	Yes.
Form of economic evaluation	Cost-effectiveness analysis	Yes. Cost utility analysis.
Time horizon	Sufficient to capture differences in costs and outcomes	2 years. RCT data is available to week 24 with additional open label trial data to week 56, with extrapolation thereafter. Within the 2 year time horizon the model is sensitive to assumptions around stopping rules and continuation of benefits. The sensitivity of the model to these assumptions increases as the time horizon and period of extrapolation is extended.
Synthesis of evidence on outcomes	Systematic review	No. Due to the comparator being placebo as within the pivotal trials this was not required.
Outcome measure	Quality adjusted life years	Yes.
Health states for QALY	Described using a standardised and validated instrument	A validated instrument was used to assess health states, the migraine specific questionnaire (MSQ) for the base case but this did not provide a direct utility measure. A mapping exercise was carried out using data from the manufacturer commissioned International Burden of Migraine

		Study (IBMS) which collected MSQ, HIT-6 and EQ5D HRQoL data from respondents.
Benefit valuation	Time-trade off or standard gamble	Time trade off for the IBMS EQ-5D data.
Source of preference data for valuation of changes in HRQL	Representative sample of the public	Yes. Standard UK tariff.
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes.
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes.
Probabilistic modelling	Probabilistic modelling	Yes. The results of the manufacturer probabilistic modelling are very different from those of the manufacturer deterministic modelling. The central estimate of cost effectiveness from the probabilistic modelling is 111% more than the deterministic estimate; i.e. more than double.
Sensitivity analysis		A good range of univariate sensitivity analyses and scenario analyses are presented.

5.2.2 Model structure

The manufacturer base case presents a cost utility markov model with 12 week cycles and a time horizon of 108 weeks, or 9 cycles. The health states of the model are specified in terms of the number of headache days per month (HDPM) with the following six health states which span the following ranges:

HS1:0-3 HDPM: range 4 HDPM

HS2:4-9 HDPM: range 6 HDPM

HS3:10-14 HDPM: range 5 HDPM

HS4:15-19 HDPM: range 5 HDPM

HS5:20-23 HDPM: range 4 HDPM

HS6:24-28 HDPM: range 5 HDPM

In what follows these health states will be referred to individually, and also as episodic migraine for HS1 to HS3 and as chronic migraine for HS4 to HS6. Botox is administered 12 weekly over 30 minutes by a consultant neurologist.

To estimate health-state utility values from disease-specific quality-of-life scores in individuals with migraine the evidence from the manufacturer is based on a recent study by

[redacted]

For the base case, each of these health states is associated with a treatment specific HRQoL estimated from an MSQ to EQ-5D utility mapping exercise estimated using data from the International Burden of Migraine Study (IBMS). Note that the manufacturer has also produced another mapping function that incorporates both the HIT-6 and the MSQ.

The health states are also associated with resource use in terms of GP visits, A&E visits and hospital admissions, reported to be drawn from Scottish data within the IBMS study. This data is grouped into just three health states, with HS1 of infrequent migraine being the only health state with its own

specific set of resource use estimates. Episodic migraine excluding HS1; i.e. HS2 and HS3, has a common set of resource use estimates, as does chronic migraine; i.e. HS4, HS5 and HS6.

The model derives 12 weekly 7 by 7 transition probability matrices (TPMs) from individual patient counts, which encompass transitions between the six health states plus an additional transition for discontinuations from therapy.

At the end of the end of cycle 2 there is the option of applying a Botox treatment negative stopping rule due to lack of efficacy, which the base case takes to be a failure to improve by at least two health states from baseline. Those that discontinue have the placebo cycle 2 TPM applied.

At the end of cycle 4 there is the option of applying a Botox treatment positive stopping rule due to efficacy, in that those achieving episodic migraine are assumed to cease treatment and remain stable thereafter. Those remaining with chronic migraine are assumed to remain on Botox treatment for the remainder of the model.

Whether the positive stopping rule due to efficacy at one year for those with episodic migraine is reasonable is an open question, as is whether they would remain in a stable health state thereafter. But the two year time horizon may in effect suggest that the cycle of treatment will recur at the start of the second year, if patients relapse back into chronic migraine. This could give rise to the same ICER occurring sequentially over a longer time horizon. A longer time horizon could in effect be seen as encompassing a number of two year models.

A recent small study by Rothrock and colleagues (available only as a meeting abstract) followed 100 patients, all of whom had responded well (50% or more reduction in headache days) to Botox, for 2 years. 8 relapsed back to chronic migraine, 24 remained largely free of headache for more than 6 months, and 68 remained responders but needed regular (roughly 3- monthly) Botox injections.[79]

But the degree to which subsequent two year cycles will mirror the first two year cycle will depend upon at least two factors:

- Whether the negative stopping rule due to lack of efficacy in the first two cycles is applied, with these patients remaining off Botox treatment into the longer term.
- The rate at which those patients who have the positive stopping rule due to efficacy at one year applied subsequently relapse and require retreatment with Botox.

The first point could suggest that after an initial two year cycle of treatment those remaining and being retreated with Botox may tend to be the subset with a better response to Botox which could improve the cost effectiveness estimate. But it may also argue for a longer time horizon within the modelling and a more explicit consideration of therapy subsequent to Botox failure. The second point

is currently a matter of conjecture, with the manufacturer base case approximating an average time to retreatment of one year.

Appendix 17 of the submission includes a treatment pathway model which estimates the costs of Botox treatment followed by greater occipital nerve (GON) block for the 30% treatment failure rate. The comparator is assumed to be GON block, the failure rate of which is 60% as drawn from expert opinion. Subsequent failures within both arms move onto an experimental basket of treatments. Over a 1 year time horizon this resulted in estimates of a total cost on the Botox arm of £2,557 compared to a total cost in the GON block arm of £4,416. This model is not presented within the submission, and in the light of the NICE scope this model is not considered further. There is the suggestion by the manufacturer of possible savings from avoiding subsequent 2nd line and 3rd line therapies. But as discussed later, the ERG regards GON as currently an unproven treatment, and as such it should not be funded by the NHS and should not be included as a cost that would be incurred if Botox were not available.

5.2.3 Population

The manufacturer presents results from the pooled PREEMPT trials. The base case presented by the manufacturer restricts this to the 1-prior patient population, with an additional analysis being presented for the 3-prior patient population. Note that within this the 1-prior patient population relates to those with at least 1-prior treatment as summarised in figure 6.1 of the Allergan submission.

For the 1-prior subgroup the patient population is 425 for Botox and 454 for placebo, while for the 3-prior subgroup these are reduced to 231 for Botox and 248 for placebo. These are the patient numbers pooled across patients who were and patients who were not overusing acute headache pain medication at baseline. Further restricting the patient population to those who were not overusing acute headache pain medication at baseline reduces patient numbers to 125 for Botox and 131 for placebo in the 1-prior subgroup, and 67 for Botox and 69 for placebo in the 3-prior subgroup.

Given the model structure and the requirement to populate 7 by 7 TPMs, as summarised below, it seems reasonable to seek to preserve the patient numbers available for analysis. In the light of the NICE scope it also seems reasonable to have focused upon the number of prior prophylaxis treatments, and to only present this further restricted to the patient population who were not overusing acute headache pain medication at baseline as a sensitivity analysis.

5.2.4 Interventions and comparators

Botox is seen as being largely adjunctive to standard care, though its effectiveness in terms of improving the patient distribution between health states is modelled as reducing the requirement for triptan medication.

5.2.5 Perspective, time horizon and discounting

he base case model perspective is to direct patient benefits using QALYs, and NHS and PSS costs as per the NICE reference case.

The model also features the option of including additional societal costs in terms of lost productivity and wages. This is not presented for the base case, but it is a useful feature given the effect of chronic migraine on time off work.

A two year time horizon is adopted, within which costs and benefits are discounted at 3.5% in line with the NICE reference case.

5.2.6 Treatment effectiveness and extrapolation

The model structure is based upon a 12 weekly model cycle that applies 7 by 7 TPMs derived from individual patient count data. These encompass movements amongst the six health states plus a seventh transition to treatment discontinuation.

A common probability of death is applied during each twelve week cycle, but given the baseline age of 42 years and the time horizon of 2 years this has virtually no impact and can be disregarded. A recent meta-analysis by Schurks and colleagues concluded that migraine is not associated with any increase in mortality.[80]

- Baseline to week 24: cycles 1 and 2

A particular feature of the model is that the TPMs encompass a treatment stopping rule for lack of efficacy during the second cycle of the model, with this discontinuation being modelled as occurring at week 24. In other words for each arm, there are four separate TPMs for weeks 12 to 24:

- a TPM with no stopping rule due to lack of efficacy,
- a TPM in which those not achieving an improvement of at least two health states between baseline and week 24 discontinue treatment
- a TPM in which those not achieving an improvement of at least one health state between baseline and week 24 discontinue treatment

- a TPM in which those not achieving an improvement of at least one health state between baseline and week 24 discontinue treatment

The summary of the last bullet point is the interpretation of the ERG. It is not entirely clear from the submission. Within the electronic copy of the model is referred to as "*Patients discontinue if they did not improve at least 1 health state within 1 cycle*", which could be interpreted as not improving by one health state between baseline and week 12 weeks or between week 12 and week 24.

If a stopping rule is applied for placebo, this affects the week 12 to week 24 TPM. But the TPM applied subsequent to this for extrapolation over week 24 to week 36, and beyond, is the no stopping rule placebo TPM from week 12 to week 24. This is reasonable but it would be simpler to disregard discontinuations within the placebo arm altogether. As outlined later including discontinuations affects the net administration costs to the benefit of Botox.

- Week 24 to week 60: cycles 3, 4 and 5

For the three 12 week model cycles between week 24 and week 60; i.e. cycles 3, 4 and 5, the manufacturer applies the same TPM for the Botox arm. This is calculated by summing the individual patient movements for each of the three periods week 24 to week 36, week 36 to week 48 and week 48 to week 56 into a single pooled matrix of patient movements, and deriving a common TPM for cycles 3, 4 and 5 for the Botox arm based upon these summed patient numbers.

For example, were there to be 12 patients in HS1 at week 24, 14 patients in HS1 at week 36 and 16 patients in HS1 at week 56 these would be summed to give a denominator of 42 patients. If 2 of the 12 patients had moved into HS2 between week 24 and week 36, 3 between week 36 and week 48 and 4 between week 48 and week 56 these would be summed to give a numerator of 9 patients. The 12 weekly probability of moving between HS1 and HS2 in the Botox arm for each of cycles 3, 4 and 5 would then be a common 9/42=21%.

The reason for deriving this common TPM for cycles 3, 4 and 5 within the Botox arm may be due to the patient level data available for cycle 5 being of only 8 weeks duration. But it may also be an attempt to increase the patient counts for each transition, which in turn reduces the uncertainty associated with the TPMs. This may affect results. As summarised below, the manufacturer model results in a deterministic estimate of cost effectiveness that is less than half the probabilistic estimate.

- Week 60 to week 108: cycles 6, 7, 8 and 9

Within the Botox arm the base case applies a treatment stopping rule to those patients who have entered episodic migraine: HS1, HS2 or HS3[1] within the Botox arm. This assumes that these patients are no longer treated with Botox and remain in steady state; i.e. have the identity matrix applied to them. Those remaining with chronic migraine: HS4, HS5 or HS6, remain on treatment and continue to cycle through the model health states as before. Note that despite the negative stopping rule due to lack of efficacy emptying the model cells in the Botox arm for HS5 and HS6 at week 24, there is some subsequent repopulation of these cells given the TPMs being applied.

Patients in the placebo arm continue to cycle through the model health states as before, having the placebo week 12 to week 24 TPM applied to them.

- Treatment discontinuations

Treatment discontinuations as drawn from the trial data, and as modelled by assumption due to the treatment stopping rule due to lack of efficacy at week 12, have the week 12 to week 24 placebo TPM applied to them. The patients incur no direct drug or administration costs, but do incur the health state specific GP visit costs, A&E costs and hospital costs.

5.2.7 Health related quality of life

Utility values were drawn from a mapping exercise using EQ-5D and MSQ data from the IBMS study. As summarised in Blumenfeld and colleagues, this was a manufacturer commissioned web based survey administered between February and April 2009 in nine countries.[23] Potential participants were identified from panels maintained by Synovate Healthcare, Chicago, or its partner companies, and incentivised through redeemable points and entry into monthly sweepstakes.

The 63,001 invited to take part in the survey had previously reported headaches or migraine, except in Taiwan where participants were invited from the pool of general patients available to the partner company. Invitees opted in through a web link and were screened through an apparently validated questionnaire based upon ICHD-II criteria. 10,650 were found to be eligible, and categorized as having either episodic migraine or chronic migraine. Of these, 8,726 completed the survey of whom 499 (6%) had chronic migraine and 8,227 (94%) had episodic migraine. The breakdown of respondent by country and migraine type is as below in

[1] Note that the model permits only a percentage of patients with episodic migraine to have this stopping rule applied. It is also possible within the model to apply a stopping rule to those remaining with chronic migraine at this point.

Table 19.

Table 19 IBMS patient distribution

	Chronic	Episodic
Australia	55 (11%)	461 (6%)
Canada	55 (11%)	626 (6%)
France	57 (11%)	1,404 (6%)
Germany	52 (10%)	1,397 (6%)
Italy	55 (11%)	921 (6%)
Spain	56 (11%)	645 (6%)
Taiwan	8 (2%)	659 (6%)
UK	57 (11%)	1,013 (6%)
USA	104 (21%)	1,101 (6%)
Total	499 (100%)	8,227 (100%)

[redacted]

[redacted]

[redacted]:

- [redacted]
- [redacted]
- [redacted]

- [redacted]

[redacted]

In response to ERG clarification question A.6 the manufacturer outlines that an additional analysis was performed using both the

[redacted]

Table 20 IBMS utility mapping functions

	[redacted]			[redacted]		
	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	* [redacted]					

It is not immediately obvious why separate models were estimated for episodic migraine and chronic migraine. A pooled analysis with dummy variables could have explored the requirement for this, with possible benefits of an increased sample size for estimation of some of the coefficients for chronic migraine, the MSQ model for this having only one statistically significant MSQ coefficient. As the IBMS is the source of the utility mapping data there may be no particular reason to anticipate that it

will capture any difference in adverse events between the Botox arm and the placebo arm, though these may be picked up by the elements of the MSQ.

The ERG reviewed the statistical approach submitted in the main report and notes that in general, the statistical methodologies proposed are suitable to these types of data. However, the ERG identified a number of concerns regarding the robustness of the statistical method used by the manufacturer to map health-state utility values from disease-specific quality-of-life scores in individuals with migraine.

The ERG notes that in chronic migraine, the preferred HIT-6 and MSQ algorithms explained [redacted] and [redacted] of the variance in the training samples which indicates *fair* agreement in terms of the R^2. In the clarification process, the ERG argued that to improve the R^2 to moderate agreement, it was possible to map the EQ-5D utility values using both the HIT and MSQ in the pool data instead of using only the MSQ in the mapping exercise. Following the clarification process, the manufacturer argued that the use of MSQ only was preferred.

The manufacturer indicates that the correlations between the EQ-5D and both the HIT-6 and MSQ dimension scores were stronger in individuals with chronic migraine than those with episodic migraine (Table 4 of Allergan submission). However, it is not clear from the modelling strategy by the manufacturer how this strong correlation was taken into account. The use of OLS certainly could not solve this issue. The ERG did raises some concerns about the use of ordinary least squared (OLS) regression models and suggested the use of an additive model (for instance censored least absolute deviations (CLAD) or Tobit model) which imposes no restrictions on the relationship between dimensions. Following the clarification question, the manufacturer argued that Tobit or CLAD have on the whole in the mapping literature not given distinct advantages over OLS.[81] However the Brazier reference[81] cited by Allergan in their response, actually reports that in the study in which OLS, CLAD and Tobit were used, CLAD performed best.

However, the ERG is still of the view that given the limitation of the OLS in dealing with correlated data, a sensitivity analysis comparing OLS, Tobit or CLAD could have been undertaken. Moreover, EQ-5D utility scores are known to exhibit a ceiling effect, where a large proportion of subjects rate themselves in full health with a utility score of 1, and hence the data can be interpreted as being bounded or censored at 1. The ERG believes that ignoring the bounded nature of the EQ-5D resulted in biased and inconsistent estimates. Following the clarification questions, the manufacturer did agree to the ERG suggestion but still preferred the OLS which the ERG does not believe it is robust enough to take into account the upper censoring limit of 1 of the EQ-5D in the mapping process.

The ERG also notes some uncertainty about the predicted EQ-5D scores in the pooled data for the population in the decision problem. Although, the manufacturer wrote that no statistically significant

difference was observed between mean observed and mean estimated EQ-5D, the manufacturer was not able to provide a graph of the observed and predicted EQ-5D scores for this population owing to the fact that the International Burden of Migraine Study (IBMS) dataset does not allow for the population described in the decision problem to be defined.

Therefore, it is likely that the estimates predicted from the validated sample might have been overestimated by ignoring the above issues.

The utility values applied within the model from the pooled PREEMPT trial data using the IBMS MSQ mapping function are specific to the 1-prior patient population and are summarised below in Table 21.

Table 21 Utilities derived for 1-prior subgroup from IBMS MSQ algorithm

Health state	Botox	Placebo	net
HS1: 0 - 3	0.746	0.724	0.022
HS2: 4 – 9	0.719	0.658	0.061
HS3: 10 – 14	0.652	0.620	0.032
HS4: 15 – 19	0.602	0.568	0.034
HS5: 20 – 24	0.515	0.558	-0.043
HS6: 24 +	0.601	0.479	0.122

There is an apparent inconsistency in the above results, in that in the Botox arm those in HS5 have a lower utility estimate than those in HS6. There are reasonable patient numbers in these health states, 43 and 41, and this may highlight some uncertainty between the MSQ mapping to utilities and correspondence with the model structure as defined in HDPM.

It is also noteworthy that within all but one headache states, those on Botox are given a higher utility than those on placebo. This implies that there are differences in other aspects such as number of headaches in a headache day, duration of headaches, or the proportion of headaches that were migraines.

There is a lack of detail about the proportion of the 1-prior patient population this utility data is based upon. There is also a lack of detail about the time point(s) the utility data is drawn from. It would be a concern if only a proportion of ITT patients' MSQ data was used for the utility estimation. It might

also be a concern if this proportion changed over time between measurement points, in tandem with the changes in the balances between health states. It would be a particular concern if the proportions reporting between the arms evolved differently. In short, the submission appears to provide no definite detail about whose utilities are being estimated, or when. Tables of the number of patients in each health state and the proportion having their MSQ analysed to produce the utility estimates for each relevant time point would have provided reassurance around this.

Applying the ERG preferred mapping function which encompasses both the MSQ and the HIT -6 results in the following utility estimates (see Table 22).

Table 22 Utilities derived for 1-prior subgroup from IBMS MSQ+HIT-6 algorithm

Health state	Botox	Placebo	net
HS1: 0 - 3	[redacted]	[redacted]	[redacted]
HS2: 4 – 9	[redacted]	[redacted]	[redacted]
HS3: 10 – 14	[redacted]	[redacted]	[redacted]
HS4: 15 – 19	[redacted]	[redacted]	[redacted]
HS5: 20 – 24	[redacted]	[redacted]	[redacted]
HS6: 24 +	[redacted]	[redacted]	[redacted]

Applying these utility estimates has minimal impact upon the cost effectiveness estimate. However, it should be noted that these are specific to the 1-prior group. As with all the utility estimates, it would have been reassuring if the manufacturer had also estimated utility values specific to the 3-prior group.

The manufacturer dismisses the utility estimates provided within Brown et al on the grounds that the paper considers a range of migraine states that are not restricted to chronic migraine patients.[82] Brown and colleagues report an average monthly migraine frequency of 6 as compared with 11 to 12 in PREEMPT. The benefit of the Brown and colleagues study is that SF-36 data was collected and mapped to utilities using the Brazier algorithm. This suggested that reductions of ≥75% in migraine frequency, 50% to 75% and <50% were associated with HRQoL changes of 0.124, 0.104 and 0.014 respectively.

5.2.8 Resources and costs

- Direct drug, administration and monitoring costs

Botox arm on therapy

For those on therapy each 12 weekly Botox administration is assumed to require one 200 Unit vial at a cost of £276.40, plus 30 minutes of consultant time. This neurologist consultant time has been valued at £73.00 based upon the 2010 PSSRU Unit Costs of Health and Social Care amount for a medical consultant of £146 per patient related hour. This does not allow for the cost of other staff. In practice, the neurology reference cost is likely to be more relevant, which is considered by the manufacturer as a sensitivity analysis.

As regards time cost for Botox administration, the ERG notes the submission to NICE from Dr Andrew Dowson, a very experienced user, in which he states that 15 minutes would be sufficient time.

Note that the electronic copy of the model has additional placeholders for 10ml saline, a 22 gauge, 1.5 inch needle for reconstitution and a 30 gauge, 0.5 inch needle for treatment. These are set to zero. Including consumables would have little impact upon overall costs or results.

Placebo arm "on therapy"
No direct drug costs are applied within the placebo arm. Based upon the response to ERG clarification question A.2 it is assumed that those on placebo will receive routine monitoring and therapy optimisation at one 15 minute outpatient consultant neurologist appointment every 12 weeks. This is costed at £36.50.

Manufacturer expert opinion suggests an average of 4 to 6 consultations for these patients prior to them being discharged back to GP care. A 3 monthly GP visit would be of similar cost to the £36.50 assumed for the base case neurology outpatient visit.

Botox arm off therapy due to lack of efficacy at week 24 stopping rule
These patients are assumed to incur no routine administration or monitoring costs.

Botox arm off therapy due to efficacy at week 48 stopping rule
These patients incur the same routine monitoring and therapy optimisation as the placebo "on therapy" group, at £36.50 every 12 weeks.

Off therapy due to trial based discontinuations
These patients are assumed to incur no routine administration or monitoring costs.

- Other costs

Adverse event costs

There is no explicit inclusion of costs related to adverse events.

Other costs

The 12 weekly rate of GP visits, A&E visits, hospital admissions and triptan use is drawn from the IBMS, and is apparently based upon the Scottish subset of the UK data as summarised in Bloudek et al [83] and as clarified within the manufacturer response to ERG clarification question. A.1. Within the IBMS participants were asked about "*the frequency of health care professional visits, use of emergency department or urgent care clinic, overnight hospital stay, diagnostic tests, and any headache specific treatments used*". On the basis of this the manufacturer applies the following rates and resultant costs per cycle (Table 23)

Table 23 Other resource use rate and average cost per 12 weeks

12 weekly rates and costs	GP visits		IP admissions		A&E visits		Triptan use	
	Rate	Cost	Rate	Cost	Rate	Cost	Rate	Cost
HS1: 0-3 HDPM	0.10	£3.20	0.03	£17.51	0.12	£10.91	1.88	£6.30
HS2 & HS3: 4-14 HDPM	0.30	£9.60	0.08	£46.69	0.28	£25.46	5.07	£16.98
HS4, HS5 & HS6: 15+ HDPM	0.58	£18.56	0.32	£186.78	0.63	£57.29	7.29	£24.42

Note that that UK subset of patients had a reasonable number of patients in the episodic migraine range of 0-14 HDPM, 1,013 patients. The UK subset of patients in the chronic migraine range of 15+ HDPM was only 55. Further restricting this sample to only Scottish patients, as appears to be the case from the Bloudek and colleagues abstract may result in very small patient numbers and quite unreliable estimates.

The unit costs underlying these estimates are:

- £32 for a GP visit, drawn from the 2010 PSSRU Unit Costs of Health and Social Care.
- £584 for an inpatient admission based upon the non-elective inpatient short stay and long stay reference cost averaged across the HRGs listing headache: AA31Z, PA04A and PA04B. The average length of stay across the 78,230 admissions was 1.5 days. Note that the average day case cost across the same HRGs is £508 the inclusion of which would marginally lower the average unit cost to £582.
- £91 for an A&E attendance, based upon the reference cost average across the various A&E types for the HRG VB09Z: Category 1 investigation with category 1-2 treatment.
- £3.35 per triptan tablet based upon what appears to be market research, within which sumatriptan 50mg, maxalt melt oral lyophilisate 10mg, sumatriptan 100mg, naramig 2.5mg and zomig 2.5mg account for 75% of prescriptions. Non-proprietary sumatriptan is by far the

cheapest of these drugs, being an order of magnitude less at £0.27 per 50mg tablet and £0.39 per 100mg tablet. If sumatriptan was used for all patients, the cost of triptans would be reduced, savings would therefore be reduced, and the ICER would rise.

5.2.9 Cost effectiveness results

The base case cost effectiveness estimates presented by the manufacturer are summarised below (Table 24 and Table 25)

Table 24 Manufacturer base case results and ICER 1-prior: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,839	1.22			
Botox	£2,388	1.31	£549	0.09	£5,828

Table 25 Manufacturer base case results and ICER 3-prior: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,895	1.20			
Botox	£2,438	1.29	£543	0.09	£6,083

The submission does not report the central estimates of the probabilistic modelling. The model copy supplied to the ERG suggests a central estimate from the probabilistic modelling of £12,275 per QALY. But this is based upon only 1,000 iterations and it appears that the probabilistic model has not reliably converged around a central estimate with this number of iterations. For this reason the CEAC presented as figure 6.9 of the submission and reproduced below (Figure 14) should be interpreted with some caution, as should the likelihood of Botox being cost effective at the various thresholds.

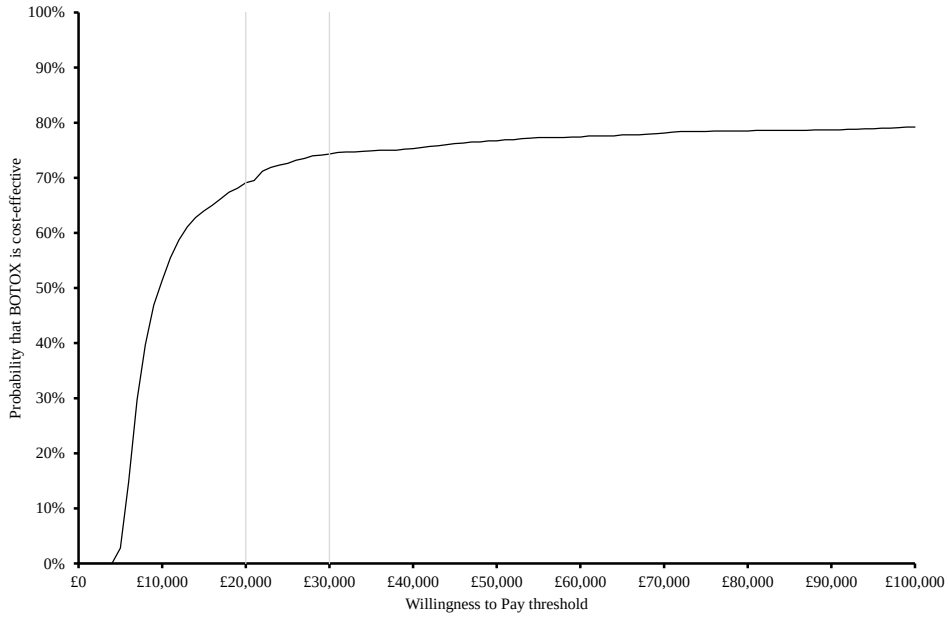


Figure 14 Manufacturer submitted CEAC 1,000 iterations: 1-prior

Table 26 Manufacturer estimates of likelihood of cost effectiveness 1,000 iterations: 1-prior

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	3%
£10,000	51%
£20,000	69%
£30,000	74%
£40,000	75%
£100,000	79%

Re-running the manufacturer submitted model with the patient population set to the 3-prior subgroup and 1,000 iterations results in a central estimate of £13,828 per QALY (Table 26).

5.2.10 Sensitivity analyses

The manufacturer presents a range of univariate and structural sensitivity analyses around: administration costs, patient populations, stopping rules, the time horizon, the calculation of utilities and the adoption of a societal perspective for costs.

The sensitivity analysis concerning the calculation of utilities applies a common set of utilities to each arm as drawn directly from IBMS survey data. The values of these do not appear to have been presented within the submission. The adoption of a societal perspective on costs applies the following estimates for lost hours of work and productivity per 12 weeks, valued at an average of £14.60 per hour. (Table 27) It appears likely that these values were also drawn from the IBMS study.

Table 27 12 weekly lost hours and productivity

Health state	Hours	Value
HS1: 0 - 3	6.31	£92.13
HS2: 4 – 9	16.78	£244.99
HS3: 10 – 14	16.78	£244.99
HS4: 15 – 19	41.35	£603.71
HS5: 20 – 24	41.35	£603.71
HS6: 24 +	41.35	£603.71

Table 28 Sensitivity analyses: patient populations

	Net Costs	Net QALYs	ICER
All patients	£609	0.09	£6,814
1-prior excluding medication over-users	£533	0.09	£5,971
Prior topiramate users	£628	0.08	£8,301

Table 29 Sensitivity analyses: 1-prior patient population

	Net Costs	Net QALYs	ICER
Base case	£549	0.09	£5,828
Administration costs			
Botox administration 15 minutes	£438	0.09	£4,654
Botox administration 50% consultant 50% nurse	£469	0.09	£4,984
Botox administration neurology OP reference costs £139	£751	0.09	£7,972
Botox monitoring additional 10 minutes consultant mid cycle	£623	0.09	£6,611
Placebo monitoring cost halved	£647	0.09	£6,870
Stopping rules			
No negative stopping rule due to lack of efficacy at week 24	£1,008	0.13	£7,946
No positive stopping rule due to lack of efficacy at week 48	£1,144	0.09	£12,486
No stopping rules	£1,883	0.12	£15,294

Minimum 1 health state improvement by week 24	£705	0.12	£6,109
50% of both episodic and chronic patients stop Botox at week 48	£836	0.09	£9,080
Those stopping due to efficacy get placebo TPM from week 60	£621	0.07	£9,503
Common utility values from the IBMS	£549	0.08	£7,025
Societal cost perspective	£380	0.09	£4,033
Hospital admission cost £474	£589	0.09	£6,252
Hospital admission cost £703	£474	0.09	£5,360

In response to the ERG clarification question A.14, the manufacturer provides estimates which apply considerably longer time horizons than the two years of the base case. These are provided for three scenarios:

- Scenario 1: as per the base case assumptions.
- Scenario 2: not applying the positive stopping rule due to efficacy at year 1.
- Scenario 3: applying the positive stopping rule due to efficacy at year 1 but applying the placebo TPM subsequently for this group.

These result in the following cost per QALY deterministic estimates for the 1-prior patient population

Table 30 Time horizon, stopping rule due to efficacy at year 1 and cost effectiveness: 1-prior

Time horizon	Base case	Scenario 2	Scenario 3
24 weeks	£27,162	..	..
1 year	£14,098	..	..
2 years	**£5,828**	£12,486	£9,503
5 years	£300	£9,467	£6,519
10 years	Dominant	£8,662	£5,740
15 years	Dominant	£8,508	£5,562
20 years	Dominant	£8,470	£5,519

These still apply the stopping rule at 24 weeks due to lack of efficacy. Unfortunately, the manufacturer has not explored this additional dimension in its presentation of sensitivity analyses which extend the time horizon. It seems likely that the stopping rule at 24 weeks due to lack of efficacy is one of the main sources of the improved cost effectiveness as the time horizon is lengthened within the analyses which do not apply the positive stopping rule due to efficacy at 1 year. As noted previously, most of the benefit occurs after the first injection, so a stopping rule before the 12 –week injection could be applied.

5.2.11 Model validation and face validity check

The manufacturer presents some validation data comparing the output of the model with the trial results at week 24 for the 1-prior subgroup. This is reproduced as presented within the submission for completeness.

Table 31 Manufacturer presented model validation data

	Botox		Placebo	
Table 6.22	Trial	Model	Trial	Model
Baseline mean HDPM	19.9	20.0	19.8	20.1
Mean change	-8.4	-9.5	-6.7	-6.9
Net mean change	-2.0	-2.6		
	Botox		Placebo	
Table 6.23	Trial	Model	Trial	Model
Baseline N HS1	0.00	0.00	0.00	0.52
Baseline N HS2	0.00	0.00	0.00	0.00
Baseline N HS3	0.00	0.48	1.00	0.52
Baseline N HS4	215.00	216.13	232.00	230.87
Baseline N HS5	127.00	123.78	129.00	132.22
Baseline N HS6	83.00	84.61	92.00	90.39
Week 24 N HS1	55.00	55.00	33.00	33.03
Week 24 N HS2	107.00	106.98	92.00	91.81
Week 24 N HS3	32.00	31.94	91.00	91.04
Week 24 N HS4	14.00	13.96	78.00	78.15
Week 24 N HS5	0.00	0.00	49.00	49.07
Week 24 N HS6	0.00	0.00	69.00	68.69
Week 24 N discontinued	217.00	216.89	42.00	41.97

It is unclear why the headache days at baseline within table 6.22 differ between the two arms in the model output, though this may be due to the model implementation for validation applying the trial based arm specific baseline patient distributions. Slightly bizarrely, the model apparently “estimates”

a small proportion being within HS1 at baseline in the placebo arm though this may be a typo. The model used for the cost effectiveness estimates applies a common average baseline distribution.

The ERG has attempted to replicate the model validation of table 6.22 of the submission using the manufacturer model, but has not been able to. It is not clear from the submission how the values have been calculated, but they appear to show good correspondence between the model and the trial data at week 24 in terms of the distribution between health states.

The correspondence in terms of the net impact upon the mean HDPM is rather poorer, with the model apparently overestimating the effect of Botox at week 24 by 30% on the basis of a net mean change of -2.0 HDPM within the trial compared to -2.6 HDPM within the model. But the -2.0 HDPM appears to have been wrongly calculated with table 6.22 implying a lower value of -8.4 HDPM minus -6.7 HDPM which is only 1.7 HDPM. Using this figure suggests that by week 24 the model has overestimated the net impact by 53%.

This appears to be a quite serious difference, but as far as the ERG is aware the estimate of the mean HDPM at week 24 has no direct impact upon the cost effectiveness estimate of the model. It is just the distribution between health states which determines costs and QALYs within the model. That said, the discrepancy warrants further explanation.

Table 6.23 reports the Botox arm as having no patients in either of HS5 or HS6, with 51% having discontinued, so it appears to include the negative stopping rule due to lack of efficacy at week 24. In itself this should not affect the HDPM and the distribution between the health states at week 24. While it would only be possible for the Botox arm, the manufacturer could also have presented validation data for week 48 and for week 60 for the model not applying any treatment stopping rule due to lack of efficacy at week 24. There are some concerns around the handling of the TPMs for cycles 3, 4 and 5 of the model within the Botox arm.

5.3 ERG cross checks and critique

Base case results

Re-running the manufacturer model, the base case deterministic results cross check with those of the manufacturer for both the 1-prior subgroup and the 3-prior subgroup.

Table 32 Manufacturer base case results and ICER 1-prior: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,839	1.22			
Botox	£2,388	1.31	£549	0.09	£5,828

Table 33 Manufacturer base case results and ICER 3-prior: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,895	1.20			
Botox	£2,438	1.29	£543	0.09	£6,083

The probabilistic results vary from those submitted by the manufacturer for the 1-prior subgroup. This may be due to the limited number of iterations possible within the submitted model. Revising the manufacturer model to permit 20,000 iterations and running over this number of iterations produces the following probabilistic results for the 1-prior subgroup.

Table 34 ERG re-run probabilistic 20,000 iterations results: 1-Prior

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,591	1.29			
Botox	£2,342	1.35	£751	0.06	£12,888

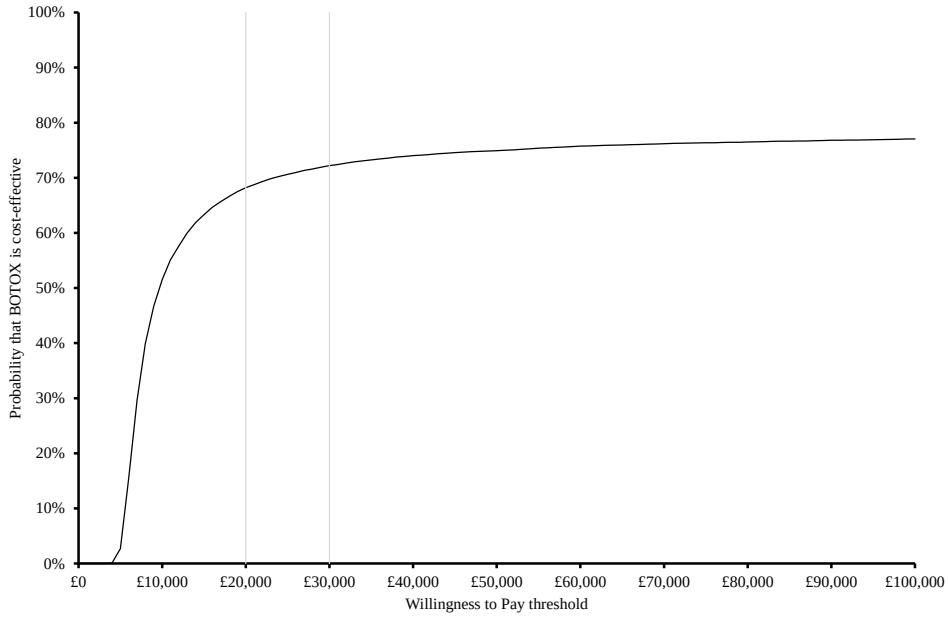


Figure 15 ERG re-run probabilistic 20,000 iterations CEACs:1-*Prior*

Table 35 ERG re-run probabilistic 20,000 iterations cost effectiveness likelihood: 1-Prior

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	3%
£10,000	51%
£20,000	68%
£30,000	72%
£40,000	74%
£100,000	77%

Similarly revising the manufacturer model for the 3-prior subgroup with 20,000 iterations results in the following.

Table 36 ERG re-run probabilistic 20,000 iterations results: 3-Prior

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,689	1.27			
Botox	£2,438	1.33	£749	0.05	£14,004

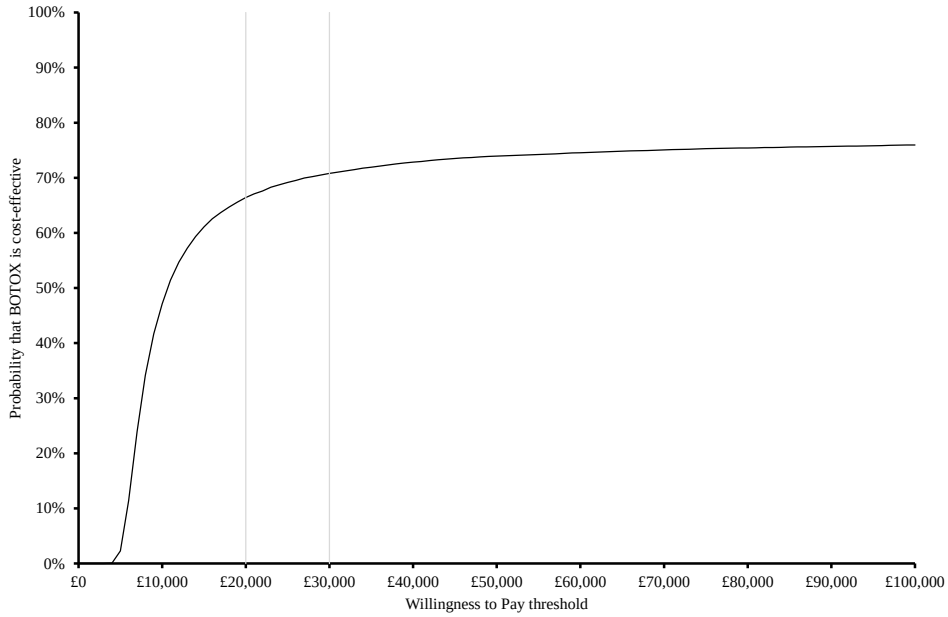


Figure 16 ERG re-run probabilistic 20,000 iterations CEACs: 3-Prior

Table 37 ERG re-run probabilistic 20,000 iterations cost effectiveness likelihood: 3-Prior

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	2%
£10,000	47%
£20,000	66%
£30,000	71%
£40,000	73%
£100,000	76%

Data Inputs

Correspondence between written submission and sources cited

The ERG has not been able to verify the additional resource use estimated by the manufacturer from the IBMS study. The 12 week rates applied within the submission are presented below (Table 38), alongside the proportion of respondents reporting this resource use and the average rate of this resource use as summarised in Blumenfeld et al.[23] Note that Blumenfeld et al estimates are only split by those with episodic migraine of 0-14 HDPM, and those with chronic migraine of 15+ HDPM.

Table 38 Cross check of 12 weekly rates of other resource use

12 weekly resource use rates	GP visits		IP admissions		A&E visits		Triptan use	
	Sub.	Paper	Sub.	Paper	Sub.	Paper	Sub.	Paper
HS1: 0-3 HDPM	[redacted]	26.4%	[redacted]	1.8%	[redacted]	4.9%	[redacted]	.
HS2 & HS3: 4-14 HDPM	[redacted]	0.69	[redacted]	0.03	[redacted]	0.10	[redacted]	
HS4, HS5 & HS6: 15+ HDPM	[redacted]	48.0% 2.07	[redacted]	2.8% 0.09	[redacted]	8.3% 0.41	[redacted]	..

While the text of Blumenfeld and colleagues is not entirely unambiguous, the ERG interpretation is that for those with chronic migraine 48% reported visiting a GP within the last 3 months. Across all those with chronic migraine the average number of visits was 2.07; i.e. among those visiting a GP the average number of visits was 2.07 / 0.48 = 4.31 visits: around once every three weeks, which seems high.

There appear to be wide disparities between the rates in table 6.17 of the submission and the rates provided in table 6 of Blumenfeld et al.[23] Blumenfeld and colleagues report that *"hospitalisations were quite rare… but slightly more common for CM than EM (2.8% vs. 1.8%, p=.086)".* The ERG interpretation is that within the IBMS the 3 monthly rate of hospitalisations among those with episodic migraine was 0.03 while among those with chronic migraine it was 0.09.

In response to ERG clarification question A1 the manufacturer states that the resource use by health states has been presented in poster form at the annual ISPOR 2011Conference: *Health Resource Use and Costs for Migraineurs in Scotland*, Bloudek and colleagues.[83] The abstract of this summarises the research as splitting adult Scottish migraineurs into six health states which appear likely to correspond to those of the manufacturer model. Reference costs and costs from the PSSRU Unit Costs of Health and Social Care were then applied. No details are provided within the abstract, but the resulting 3 monthly cost estimates ranged from £35 for the best health state, to £226 for the worst with these latter costs apparently being largely driven by hospital admission costs. These are less than the £37 for HS1 and the £287 for health states HS3, HS4 and HS5 of the submission.

The disparity between the results of Blumenfeld and colleagues and Bloudek and colleagues appears to be quite large. A possible explanation may be a very limited number of patients with chronic migraine within the Scottish subset of the UK IBMS patient data. A small absolute number of these Scottish patients reporting hospitalisations could result in a high rate. Within the manufacturer submission the standard error for the 0.32 estimate is reported as 0.14.

No data, such as the number of patients or centres, were provided within the Bloudek abstract or the industry submission or in response to the ERG clarification question A1. This is despite Bloudek and the majority of authors being manufacturer employees and the IBMS being a manufacturer commissioned survey. Due to this lack of data, the ERG is of the opinion that the published IBMS resource use estimates from Blumenfeld and colleagues are preferable to those of Bloudek and colleagues.

The ERG has not been able to replicate the average cost per A&E visit. The ERG attempt for HRG VB09Z: Category 1 investigation with category 1-2 treatment results in an average of £77.33 rather than the £90.94 of the manufacturer calculation (Table 39).

Table 39 VB09Z: Category 1 investigation with category 1-2 treatment average cost

A&E VB09Z	N	Unit cost
A&E Leading to Admitted	695,258	£96.67
A&E Not Leading to Admitted	2,563,838	£90.41
A&E Minor Injury Service Leading to Admitted	174,854	£45.34
A&E Minor Injury Service Not Leading to Admitted	586,121	£48.81
A&E Walk In Centres Leading to Admitted	111,049	£43.67
A&E Walk In Centres Not Leading to Admitted	408,751	£35.02
Non 24 hr A&E Leading to Admitted	90,312	£42.39
Non 24 hr A&E Not Leading to Admitted	99,605	£72.54
	Mean	£77.33

Note also that it may be unlikely for all migraine patients attending A&E to have VB09Z applied, as this includes investigations. In North America neuroimaging is used much more frequently than in the UK, where the BASH guidelines state that such investigations are unnecessary for diagnosing migraine. The less intensive VB11Z: No investigation with no significant treatment has an average unit cost of £60.31.

Correspondence between written submission and electronic model

There is a minor discrepancy in the HRQoL value for HS2 within the Botox arm which in the table 6.11 of the main submission is given as 0.719, while in the table of appendix 13 and the electronic model 0.710.

Construction of the electronic model

The manufacturer model is constructed in quite a convoluted manner and is not readily amenable to cross checking by tracing individual cells' antecedents. The formulae are lengthy, typically unique to each cell and difficult to trace through the model. The cohort flow within the Botox arm is also broken up into six different worksheets, with a break point occurring at the fifth cycle. This again complicates cross checking and the ability of the ERG to be able to confidently state that the model is error free.

For this reason the ERG has attempted to rebuild the deterministic manufacturer model as a cross check relying upon: matrix multiplication, keeping the cohort flow on a single worksheet and providing an overall cohort flow cross check sum. The results of this rebuild of the manufacturer model show good correspondence with the aggregate QALYs and direct drug and administration costs. The results for the other net cost offsets are not quite as closely in line with those of the manufacturer model, differing by around 10-15%.

Table 40 ERG deterministic model rebuild cross check

	QALYs	Drug	Admin	GP	Hosp	A&E	Triptan	Total
Placebo	1.19	£0	£265	£120	£1,022	£356	£174	£1,937
Botox	1.28	£831	£295	£104	£834	£306	£156	£2,526
Net	0.093	£831	£30	-£16	-£189	-£50	-£18	£589
ICER	£6,341							

Note that the above is a rebuild employing the manufacturer base case assumptions and parameter values. It does not imply that the ERG agrees with these assumptions and parameter values. The reasons for the discrepancies between the ERG cross check model rebuild and the manufacturer submitted model and not clear: it may be a difference of interpretation; or, an error within the manufacturer model; or, an error within the ERG model rebuild cross check. In the light of this and the general spirit of openness and transparency, the ERG model rebuild cross check was sent to NICE for forwarding to the manufacturer for checking on 2nd December.

But the ERG views the cost offset resource use parameters as being too high, in particular hospital admission rates. Revising these to more reasonable amounts will reduce the discrepancies between the manufacturer model and the ERG cross check model rebuild.

Model structure and implementation

- Number of modelled health states

The health state bands chosen for the modelling are not justified within the submission. The range of HDPM varies between them:

- 0-3 HDPM: range 4
- 4-9 HDPM: range 6
- 10-14 HDPM: range 5
- 15-19 HDPM: range 5
- 20-23 HDPM: range 4
- 24-28 HDPM: range 5

The reasons for the unequal ranges are not provided. The manufacturer presents a model within which the patient group and associated TPMs can be varied by: the number of prior prophylactic treatments; whether to include or exclude medication over-users; whether to include a stopping rule due to lack of efficacy; and, whether to include a stopping rule due to efficacy.

The same cannot be said for the application of the utilities data. As reviewed below, only one set of utilities are applied within the modelling. These are drawn from the 1-prior patient population, and differentiated by arm. The ranges of health states could be chosen to be aligned with a particular set of utility values. But there would be a degree of swings and roundabouts within this, and the greater concern is that one set of common utility values are applied to the different patient populations being modelled.

The manufacturer does supply a sensitivity analysis which applied a condensed model with only 3 health states: 0-3 HDPM for infrequent migraine, 4-14 HDPM for episodic migraine and 15+ HDPM for chronic migraine. But this condensed model appears to be the same as the full model, with only the utility values for the different health states being changed. The net costs do not change between the two arms.

The additional resource use data applied by the manufacturer is also much more naturally applied to a model which genuinely does model only three health states: 0-3 HDPM for infrequent migraine, 4-14 HDPM for episodic migraine and 15+ HDPM for chronic migraine. A model with only three health states would, taking into account the need to model discontinuations, only require 4 by 4 TPMs to be estimated. Only 16 cells would need to be populated for the estimation of the TPM in contrast to the 49 cells of the TPMs of the submitted model: a reduction of 67%. It is likely that this would significantly reduce the uncertainty within the modelling. It might also reduce the disparity between the deterministic estimates and the probabilistic estimates, though it is not inconceivable that it might exacerbate any ceiling and floor effects.

- Botox TPM calculation for cycles 3, 4 and 5 and extrapolation

The manufacturer model does not readily lend itself to exploring alternative assumptions around the source of the TPMs data for different cycles. There are some concerns that the treatment of the TPMs for cycles 3, 4 and 5 within the Botox arm may not be reasonable.

An exploration of this within the ERG cross check model rebuild can be implemented if the Botox arm TPM implied by the patient movements between week 24 and 36 is applied for cycle 3, and likewise the between week 36 and 48 is applied for cycle 4. This leaves the TPM for cycle 5 and extrapolation which can be as per the manufacturer, based upon the two patient count matrices between week 36 and week 48 and between week 48 and week 56, based upon that between week 48 and week 56, or even just based upon the last full 12 week cycle between week 36 and week 48.

These changes have surprisingly little impact upon results. For the 3-prior subgroup the cost effectiveness estimate changes relatively little from £6,319 per QALY to a maximum of around for £6,700 per QALY for a variety of combinations. Note that the £6,319 per QALY differs from the manufacturer model estimate of £6,083, but the proportionate change would probably be roughly similar.

- Half cycle correction

Half cycle correction does not appear to have been applied within the model. The twelve week cycle is relatively long compared to the base case time horizon of two years, and as a consequence half cycle correction would be anticipated to affect both the Botox and placebo outcomes, and the net amounts between the arms.

But this may have slightly biased the analysis against Botox. Not accounting for the higher proportions moving into the better health states under Botox until the end of the period may slightly underestimate the net benefits of Botox over placebo, and given the assumed cost structure the net cost savings from these sources.

- Stopping rule due to lack of efficacy at week 24:

ERG expert opinion views a stopping rule after two sets of injections due to lack of efficacy as reasonable. How to define a substantial reduction is more problematic. ERG expert opinion suggests that the number of headache days would be required to be reduced by around one third before proceeding with another set of injections. On this basis the manufacturer assumed stopping rule due to

lack of efficacy is reasonable, though improving by only one health state at 24 weeks might also be reasonable.

Model implementation

Within the manufacturer model applying a stopping rule due to lack of efficacy within the placebo arm affects results within the placebo arm. Costs vary in part due to the model not applying any ongoing administration costs to those discontinuing. But the cost offsets also change as do the aggregate QALYs within the placebo arm. This appears to be due to the subset of patients discontinuing due to lack of efficacy being modelled as transferring to the same health state in the subsequent cycle which will not be formally correct: some will discontinue and remain in the same health state, but some will discontinue and improve while some will discontinue and worsen.

The impact upon model outputs for the scenario of applying a stopping rule in the placebo arm is limited, but this consideration will also apply to those modelled as discontinuing within the Botox arm of the model. However, as this discontinuation rule is hypothetical rather than trial based there is no data to support this. It may have been better to assume that those discontinuing due to lack of efficacy in the Botox arm followed the placebo transition probabilities, and not to model any discontinuations within the placebo arm.

Ongoing care costs

Those discontinuing therapy due to lack of efficacy are assumed to incur no ongoing routine care costs. This benefits the Botox arm given the base case assumptions of a stopping rule being applied for Botox but not for placebo. By the 10th cycle around 57% of patients in the Botox arm have discontinued treatment compared to only 35% in the placebo arm. This may seem initially counterintuitive, but stems from the adoption of a negative stopping rule in the Botox arm but not in the placebo arm. Applying a routine care cost per cycle equal to that assumed for placebo would increase costs in both arms, by perhaps around £60-70 in the placebo arm and £160-170 in the Botox arm within the 1-prior population. It seems unreasonable for the application of the stopping rule to have this impact.

- Stopping rule due to efficacy at week 48

ERG expert opinion suggests that this stopping rule is likely to be more controversial. Patients may report a large reduction in headache days only to then quickly develop more frequent attacks again, and would then have to be re-injected. The ERG expert currently only has a small number of patients who are on their 4th set of injections; i.e. at the point at which the positive stopping rule due to efficacy would apply. These patients will definitely not want to stop their treatment, and it is

apparently difficult to argue against the fear of relapse. Another issue is that this rule would be entirely dependent upon patients' reported symptoms. Those not wishing to stop will just report whatever is required in order to remain on treatment.

- Probabilistic modelling

The model can be revised to: remove the priors from the calculation of the probabilistic TPMs; apply a common placebo group week 12 to week 24 probabilistic TPM within an iteration wherever this is used for extrapolation; treat the identity matrix as deterministic; and, retain the inverse gamma whenever the cell count is less than 340[2]. This does reduce the gap between the deterministic results and the probabilistic results but it does not eliminate it. The simulations outlined below (Table 41) apply these changes, and are also run for 20,000 iterations.

Table 41 ERG revision probabilistic 20,000 iterations results: 1-Prior

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,547	1.15			
Botox	£2,280	1.22	£732	0.08	£9,317

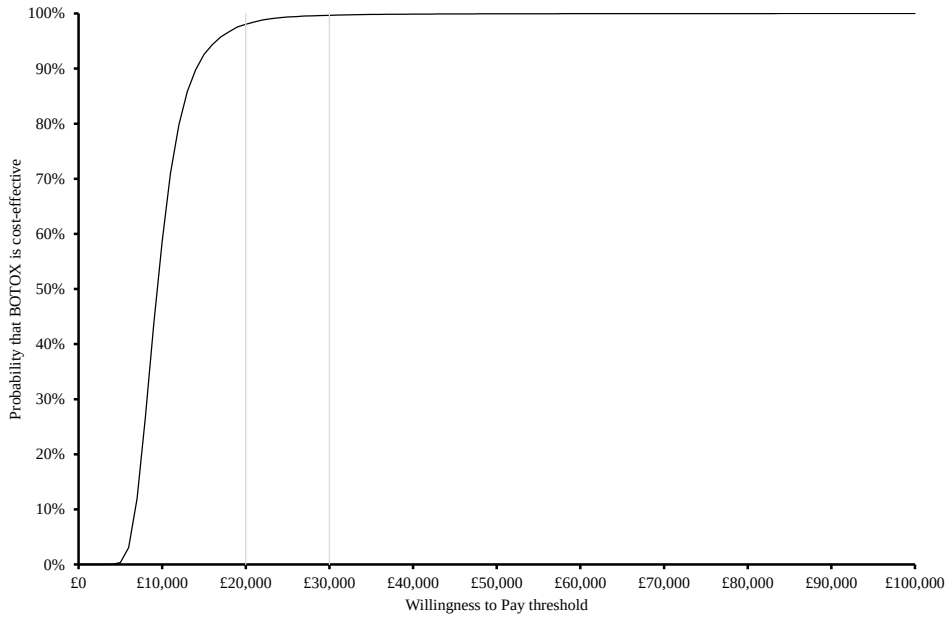


Figure 17 ERG revision probabilistic 20,000 iterations CEACs: 1-Prior

[2] Implemented by removing the references to the prior distributions within matrices D57:J63, D148:J154, D239:J245 and P57:V63 of the *Transitions_Botox* worksheet and within matrices D57:J63, D148:J154 and D239:J245 of the *Transitions_Placebo* worksheet; setting elements on the principle diagonal of these matrices equal to 1 if the observed row count is zero; set the elements of matrix O269:U275 and D269:J275 equal to D178:J184 within the *Transitions_Placebo* worksheet; and calculating the cumulative gamma/normal functions along the following lines: IF(D148>0,IF(D148<340,GAMMAINV(RAND(),D148,1),NORMINV(RAND(),D148,SQRT(D148))),0)

Table 42 ERG revision probabilistic 20,000 iterations cost effectiveness likelihood: 1-Prior

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	0%
£10,000	59%
£20,000	98%
£30,000	100%
£40,000	100%
£100,000	100%

Similarly, revising the manufacturer, for the 3-prior subgroup 20,000 iterations results in the following (Table 43).

Table 43 ERG revision probabilistic 20,000 iterations results: 3-Prior

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,607	1.16			
Botox	£2,343	1.22	£736	0.06	£11,447

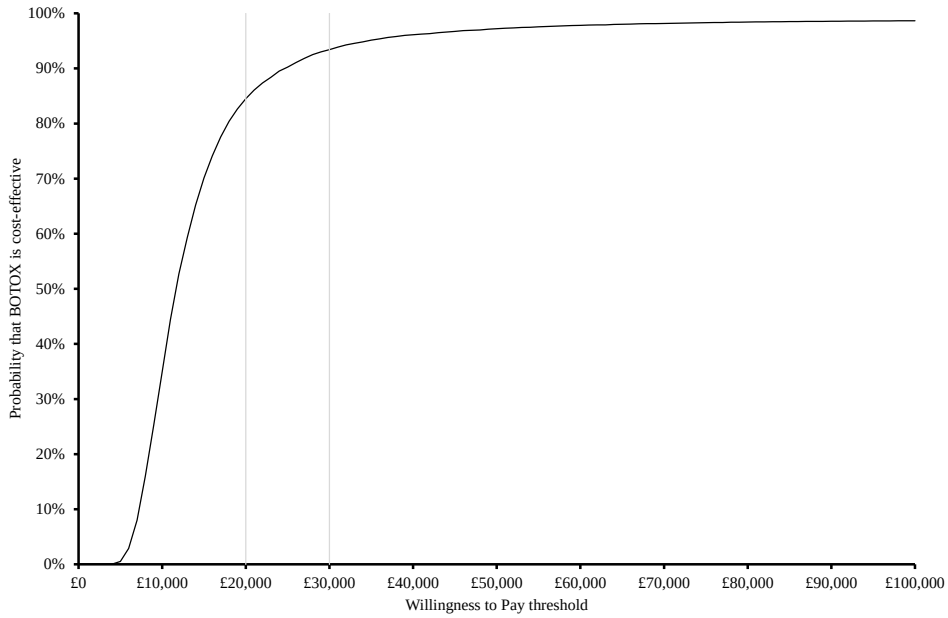


Figure 18 ERG revision probabilistic 20,000 iterations CEACs: 3-Prior

Table 44 ERG revision probabilistic 20,000 iterations cost effectiveness likelihood: 3-Prior

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	1%
£10,000	35%
£20,000	85%
£30,000	93%
£40,000	96%
£100,000	99%

The ERG revisions to the probabilistic modelling result in central estimates for the cost effectiveness of Botox of £9,317 per QALY for the 1-prior group and £11,447 per QALY for the 3-prior group, as compared to deterministic estimates of £5,828 per QALY for the 1-prior group and £6,083 per QALY for the 3-prior group. These probabilistic estimates are less divergent from the deterministic estimates than the probabilistic estimates from the manufacturer submitted model of £12,888 per QALY for the 1-prior and £14,004 per QALY for the 3-prior, but the results are still very strongly non-linear.

This non-linearity appears to be largely due to the probabilistic modelling of the TPMs. A possible explanation may arise in terms of how probabilistic TPMs model patients as moving away from the central estimates of the deterministic modelling. Within the Botox arm there may be a greater ceiling effect upon how much better probabilistic cohort flows can be than the deterministic cohort flow, in terms of patients not being able to improve further than HS1; i.e. HS1 is closer to the average patient experience in the Botox arm than in the placebo arm. Similarly, within the Botox arm there may be less of a floor effect upon how much worse probabilistic cohort flows can be than the deterministic cohort flow; i.e. HS6 is further from the average patient experience in the Botox arm than in the placebo arm.

The extent to which ceiling and floor effects may affect the probabilistic modelling can in part be gauged by the cohort flows for the 1-prior group presented for the deterministic modelling in figure 6.3 and figure 6.4 of the submission. In the second year of the model the balance between the health states for both Botox and placebo is reasonably constant. With regards any ceiling effect, the proportion of patients in HS1 is around 30% in the Botox arm compared to around 10% in the placebo arm. With regards any floor effect, the proportion of patients in HS6 is around 10% in the Botox arm compared to around 16% in the placebo arm.

Given this, other things being equal it may be that the greater is the uncertainty in the TPMs the greater will be the divergence between the deterministic estimate of cost effectiveness and the central probabilistic estimate of cost effectiveness. Some support for this is provided by the ERG revisions to the manufacturer probabilistic modelling resulting in probabilistic estimates which are closer to those of the deterministic estimates.

But this also argues for all results including the sensitivity analyses being presented for both the deterministic modelling and the probabilistic modelling. Due to time constraints the ERG has not been able to undertake further runs of the probabilistic model that apply 20,000 iterations. Any further ERG probabilistic analyses are presented based upon only 1,000 iterations, and the resulting cost effectiveness estimates may not have reliably converged.

- Quality of life values

As previously outlined, the utilities applied to a given health state differ depending upon whether the patient is in the Botox arm or the placebo arm. The utilities are derived from the MSQ data of the 1-prior subgroup, though the time point for this data does not appear to be specified (Table 45).

Table 45 Utilities derived from 1-prior subgroup

Health state	Botox	Placebo	net
HS1: 0 - 3	[redacted]	[redacted]	[redacted]
HS2: 4 – 9	[redacted]	[redacted]	[redacted]
HS3: 10 – 14	[redacted]	[redacted]	[redacted]
HS4: 15 – 19	[redacted]	[redacted]	[redacted]
HS5: 20 – 24	[redacted]	[redacted]	[redacted]
HS6: 24 +	[redacted]	[redacted]	[redacted]

The above utilities appear likely to be specific to the health states experienced by the 1-prior patient. But these utilities and the corresponding gain from Botox over placebo are applied to all the subgroups being modelled, including the 3-prior subgroup as specified in the NICE scope. Having specified subgroup specific TPMs, there is no obvious justification for not having specified subgroup specific utilities. The utility increments for Botox over placebo for a given health state for the 3-prior subgroup have not been reported. It may be an additional reason for having adopted the 1-prior subgroup for the base case rather than the 3-prior subgroup.

Setting the placebo health state utilities equal to those for Botox roughly doubles the ICER. In other words, the estimated utility increments for Botox over placebo for a given health state as summarised in Table 47 above are the source of roughly half the total net patient gain from Botox over placebo.

Section 6.2.3 of the submission outlines a regression analysis from the EQ-5D data collected under the IBMS study which found the following statistically significant coefficients (Table 46).

Table 46 Utility regression using HDPM

Variable	Coefficient	P value
Intercept	0.822	$p<0.001$
HDPM	-0.014	$p<0.001$
Age	-0.002	$p<0.001$

The manufacturer uses this to in part justify the use of HDPM as the defining variable for health states. But this also raises the question as to why HDPM were not included in the estimation of utility,

with this instead relying upon the IBMS MSQ data. These coefficients can be applied to the average age of 42 and the week 24 HDPM data contained within the electronic copy of the model[3] and appear to suggest the following utility estimates, though this has not been confirmed with the manufacturer.

Table 47 Utility regression including HDPM implied utilities and net utilities: 24 week

	Placebo		Botox		Net
	HDPM	Utility	HDPM	Utility	impact
HS1: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HS2: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HS3: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HS4: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HS5: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
HS6: On Tx	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disc.	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based upon the number of HDPM, these estimates suggest minimal differences between Botox and placebo in the average utility experienced for a given health state. This further suggests that for there to be noticeable differences between Botox and placebo in the average utility experienced for a given health state, the characteristics of the headaches experienced would have to differ between Botox and placebo. This could arise from differences in: the number of headaches experienced during a headache day; or, the duration of these headaches; or, the proportion of these headaches which were migraines.

Table 6.12 of the submission, as repeated below, provides some justification for this in terms of the reduction in the number of moderate to severe headaches from baseline experienced by those in the six health states. The net impact upon the moderate to severe headaches between the arms at 24 weeks appears to be greater than the net impact upon HDPM at 24 weeks as summarised above[4].

[3] Cells C272:E300 of the *Input Parameters* worksheet. These are described as "Migraine Frequency (as measured in Headache Days per 28 Days)". These also include an average HDPM for those having discontinued therapy which appears to have been assumed to be the same for both arms.

[4] The common 14.4 HDPM for each arm complicates this, but may be an assumed simple average across the health states and arms which the ERG calculates from these figures as 14.2 HDPM, any discrepancy possibly being due to rounding.

Table 48 Mean reduction in moderate to severe headaches from baseline at week 24

Health state	Botox	Placebo	net	net MSQ utility
HS1: 0 - 3	[redacted]	[redacted]	[redacted]	[redacted]
HS2: 4 – 9	[redacted]	[redacted]	[redacted]	[redacted]
HS3: 10 – 14	[redacted]	[redacted]	[redacted]	[redacted]
HS4: 15 – 19	[redacted]	[redacted]	[redacted]	[redacted]
HS5: 20 – 24	[redacted]	[redacted]	[redacted]	[redacted]
HS6: 24 +	[redacted]	[redacted]	[redacted]	[redacted]

The parallel figures for migraine episodes could provide further reassurance along these lines.

As already noted, the submission appears to provide no definite detail as to the proportion of the ITT 1-prior patient population in each arm these estimates are drawn from, and the evolution of these proportions if the MSQ data for the utility estimation is drawn from different time points. This raises concerns around a potential for bias, given the steady evolution of HDPM over time in both arms of the pooled trial data.

- Resource use

Botox administration costs and placebo routine monitoring costs

The hourly cost for medical consultant patient related time cross checks with the PSSRU Unit Costs of Health and Social Care. ERG expert opinion suggests that it would take 30-40 minutes to review a patient and inject Botox, and that this might lengthen if the patient bleeds a lot during administration which sometimes happens. A follow-the-pain approach would add an additional 10-15 minutes. The Botox administration time assumed by the manufacturer may be towards the optimistic end of the spectrum.

ERG expert opinion agrees that a 15 minute appointment would be normal for routine follow up not including Botox, and that this figure is also broadly representative of the average across all neurology OP appointments. The manufacturer expert opinion as provided in answer to ERG clarification question A2 suggests a frequency for these visits of between 3 and 6 months, with an average number of visits prior to discharge of 4. Of the six experts one is an outlier, suggesting a range between 4 and 10 visits prior to discharge. An average of 4 visits in total with 3 to 6 months between them suggests that a rough average for the model might be around £27 per 12 week cycle.

The above also need to be contrasted with the neurology consultant led first OP visit reference cost of £207, and £140 for follow up visits[5].

Other costs

As reviewed in the section cross checking the correspondence between the submission and the references cited above, there are reasons to believe that the costs offsets have been overestimated in the manufacturer submission. The ERG views the estimates of Blumenfeld and colleagues (2010) as a more reasonable source for this data, though even these estimates may be too high for UK practice.

[5] 2009-10 reference costs: Consultant led non-admitted face-to-face: Service code 400

6 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The ERG is of the opinion that a more reasonable baseline would apply[6]:

- The neurology outpatient consultant face to face follow-up reference cost of £140 for Botox administration and placebo follow up
- Placebo routine care costs for those discontinuing therapy
- The resource use estimates specified within Blumenfeld and colleagues (2010)
- An average A&E cost of £77
- The TPMs from the 3-prior patient subgroup

There is also uncertainty around the degree to which a positive stopping rule at one year can or will be applied. If it is applicable, there remains additional uncertainty around the average time to relapse and subsequent treatment. The admittedly small Rothrock study suggests that most patients continue on Botox injections.[79] For this reason the analysis is presented with the one year stopping rule and without it. The ICER rises markedly if the stopping rule is not applied.

Table 49 ERG revisions 3-prior with positive stopping rule: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,657	1.20			
Botox	£2,573	1.29	£916	0.09	£10,257

[6] Implemented in the BOTOXCost worksheet by setting cell D5=£140 and cell D12 equal to half of this; in the *Markov Model -Botox Discontinue* worksheet setting cell BK6=BD6*cCOMP/(1+DISC)^(ROUND(($D6/52)-0.5,0)) with the cells in BK7:BK14 similarly revised and then summing these and adding to the botos arm total administration cost; revising cells BK6 :BK14 in the *Markov Model - SoC Discontinue* worksheet in a similar fashion and summing these and adding to the placebo arm total administration cost; in the *Input Parameters* worksheet by setting cells E356:E358=0.69, E359:E361=2.07, E362:E364=0.03, E365:E367=0.09, E368:E370=0.10 and E371:E373=0.41; and, in the *selected_values* worksheet by setting cell D130=£77.

Table 50 ERG revisions 3-prior without positive stopping rule: deterministic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,657	1.20			
Botox	£3,167	1.29	£1,510	0.09	£17,517

Within the above, the ERG may not have correctly revised the manufacturer model to take into account routine administration and monitoring costs associated with those discontinuing treatment, including those failing on Botox at 24 weeks[7]. Excluding this revision, retaining the one year stopping rule the cost effectiveness estimate is £7,817 per QALY while including increases it to £10,257 per QALY. The parallel changes in the ERG model rebuild cross check cause the cost effectiveness estimate to rise from £7,877 per QALY to £9,910 per QALY.

Due to the apparent non-linearity of the model these results should also be presented probabilistically. Given the concerns around the manufacturer probabilistic modelling, the probabilistic estimates are provided that apply the previously described ERG revisions.

Table 51 ERG revisions 3-prior with positive stopping rule: probabilistic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,457	1.16			
Botox	£2,513	1.22	£1,056	0.07	£16,165

[7] Note that the manufacturer model includes a placeholder for this cost, cNH=L172, which is in the hidden sheet comp_NaturalHx and is set to zero. Setting this equal to cCOMP results in a cost effectiveness estimate of £16,035 per QALY.

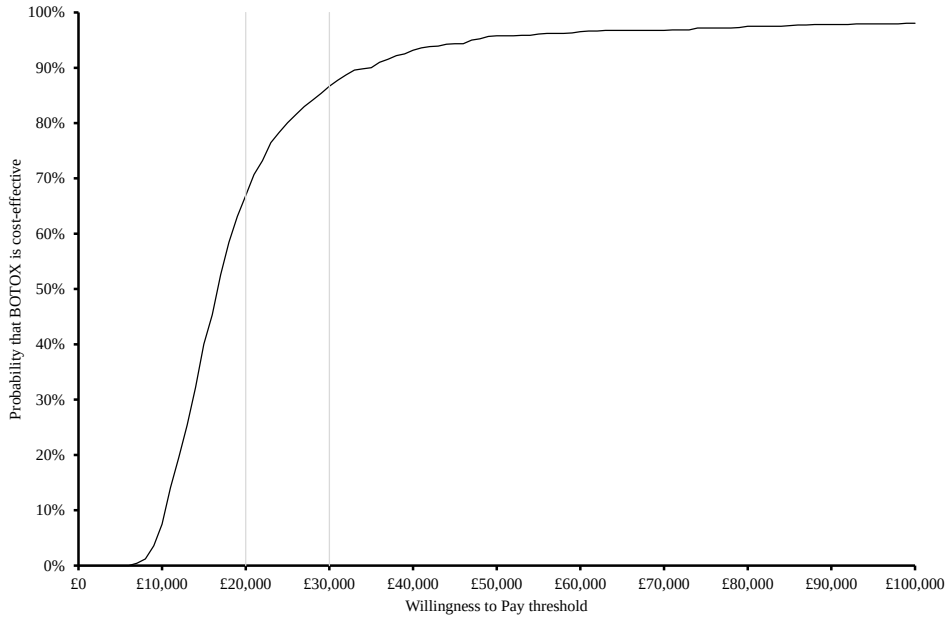


Figure 19 ERG revisions 3-prior with positive stopping rule CEAC

Table 52 ERG revisions 3-prior with positive stopping rule cost effectiveness likelihood

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	0%
£10,000	7%
£20,000	67%
£30,000	87%
£40,000	93%
£100,000	98%

Table 53 ERG revisions 3-prior without positive stopping rule: probabilistic modelling

	Costs	QALYs	Δ Costs	Δ QALYs	ICER
Placebo	£1,466	1.16			
Botox	£3,114	1.22	£1,646	0.06	£26,494

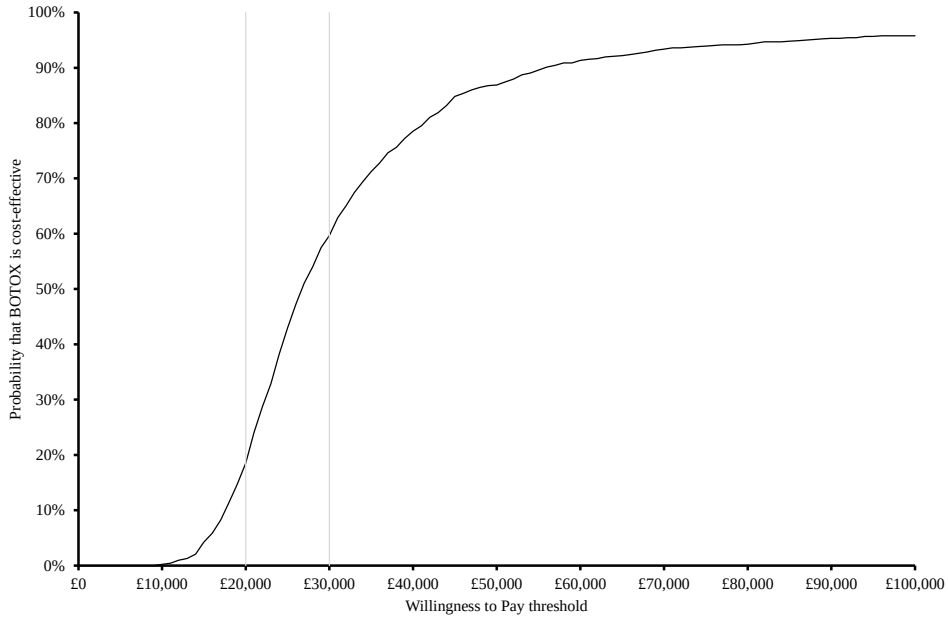


Figure 20 ERG revisions 3-prior without positive stopping rule CEAC

Table 54 ERG revisions 3-prior without positive stopping rule cost effectiveness likelihood

WTP per QALY	Likelihood Botox cost effective
£0	0%
£5,000	0%
£10,000	0%
£20,000	19%
£30,000	60%
£40,000	79%
£100,000	96%

7 End of life

The end of life modifiers are not applicable.

8 Overall conclusions on the manufacturer's modelling

The general model structure employed by the manufacturer is reasonable. Its implementation is difficult to cross check and the ERG has been forced to fall back on a model rebuild. The outputs of this show a high degree of correspondence with the manufacturer model.

The main concerns of the ERG are:

- Not presenting the analysis for the 3-prior group as the base case given the NICE scope
- The Botox TPMs derived from the open label phase have been estimated in a peculiar manner and are not obviously justified.
- The model is strongly non-linear and the manufacturer does not particularly consider the central estimates of the probabilistic modelling.
- The validation data presented in table 6.22 of the submission appears to suggest that the model may overestimate the net impact of Botox upon the mean HDPM by up to 50%.
- The administration and monitoring costs may be better estimated through standard neurology outpatient reference costs.
- The manufacturer estimates of the resource use associated with a given health state are too high.
- It is currently unknown whether for the 3-prior group the utility estimates for a given health state would show as large a benefit for Botox over placebo as those used within the modelling.
- A lack of detail about the proportions of the ITT patients in each arm from whom utility estimates were made, and the evolution of these over time

9 Discussion

9.1 Other comparators

A full review of all possible other treatments is outwith the scope of this ERG report, but we have conducted a quick search for recent reviews.

A review, published in 2000, of physical and behavioural treatments for migraine by the Agency for Health care Policy and Research (AHCPR) is reported in a guideline from the US Headache Consortium summarised below.[84] It is noted that these interventions are mainly for prophylaxis. The findings for various treatments are as follow.

- Relaxation training. Five trials when pooled in a meta-analysis showed a reduction in headache index or frequency of 41%, with an effect size in the meta-analysis of 0.55 in favour of relaxation therapies (95% CI 0.14 to 0.96).
- Hypnotherapy. One trial compared hypnotherapy with prochlorperazine, and another with thermal biofeedback plus relaxation therapy. Both found significant reductions in headache frequency before and after, but no difference between arm
- Thermal biofeedback training. Five trials had an average improvement of 37% in headache activity, but the three that could be meta-analysed showed an effect size of only 0.38, which was not statistically significant (95% CI – 0.18 to 0.94)
- Thermal biofeedback plus relaxation training. Eight studies could be entered into a meta-analysis. Their average improvement in headache activity was 33%, and the effect size was 0.40, which was statistically significant (95% CI 0.01 to 0.79)
- Electromyographic biofeedback gave a 40% reduction in headache activity, but the effect size (based on 3 trials) of 0.77 was not statistically significant (CI 0.24 to 1.3)
- Cognitive behaviour therapy. Seven trials reported an average 49% improvement in headache activity. Five trials could be entered into a meta-analysis, effect size 0.54 (0.13 to 0.94)
- Acupuncture. Mixed results were reported in seven small trials, with some reporting benefit. Publication bias may be a problem.
- Transcutaneous electrical nerve stimulation (TENS). Two trials showed no benefit but at least one was for acute migraine, not prophylaxis
- Cervical manipulation did not appear effective.

The main recommendation in the guideline was:

"Relaxation training, thermal biofeedback combined with relaxation training, EMG biofeedback and cognitive behavioural therapy may be considered as treatment options for prevention of migraine."

Several Cochrane reviews assessed the effects of non-pharmacological prophylaxis. Some of these are now several years old, and many of the trials they include predated the new definition of CM.

One on acupuncture for migraine prophylaxis concluded that acupuncture was better than no treatment but the effect size was small, and could have been due to a placebo effect.[85] When acupuncture was compared with sham acupuncture, no difference was seen. The review does not specify whether patients had chronic migraine as now defined.

The second review was on non-invasive physical treatments for chronic and recurrent headaches.[86] These treatments included spinal manipulation, and transcutaneous electrical nerve stimulation (TENS). The five patient groups included those with other types of headaches such as "chronic tension-type headache". Some trials included an amitriptyline arm, so that patient groups may not be relevant to those who have failed on 3 or more oral prophylactics, as in the NICE decision problem. Spinal manipulation was no better in the short-term than amitriptyline. There was heterogeneity amongst the trials. There was weak evidence of benefit from TENS.

One problem in reviews is that authors do not distinguish between acute and chronic migraine.

A review of oxygen treatment (including both hyperbaric and normal pressure) reported one trial of hyperbaric oxygen therapy (HBOT) versus sham HBOT for prevention of migraine, that found no significant difference in the mean number of headache days.[87] No trials of normobaric oxygen therapy for prevention were found.

A recent review of spinal manipulation for migraine found one trial that reported on frequency of migraine, but they concluded that there was no evidence to support spinal manipulation.[88]

Hence the evidence for non-pharmacological interventions is mixed, but some seem to have some effect.

Occipital nerve interventions

Askenazi and colleagues carried out a systematic review of peripheral nerve blocks in the management of various types of headache, including migraine.[89] They found only four studies, of which only one was clearly in chronic (presumably – the paper refers to refractory) migraine. None of these studies were RCTs. One had a follow-up of 30 minutes so was presumably in acute migraine. Two studies using repeated greater occipital nerve blocks reported improvements for up to 6 months. The authors conclude that better quality research is needed.

Another review by Tobin and Flitman also reviewed ONB at different types of headache. [90]They reported two studies not included in the Ashkenazi review, one being their own unpublished study, the other being a study by Afridi et al of a single ONB with a short (30 days) follow-up.[91]

A recent series was published too recently to be included in the Ashkenazi review. Weibelt and colleagues carried out occipital nerve blocks (ONBs) in a subgroup of patients with chronic migraine thought to originate from the base of the skull – called "cervicogenic" headache. 150 patients were treated with unilateral or bilateral ONBs and were reviewed a month later.[92] Just over half reported a reduction in headache days of 50% or more. However, with no control group, it is not possible to say how much of the change was due to placebo effect

The ERG conclusion from this brief review of the papers on ONB is that the evidence base is too limited to say whether ONB is of value or not. We need controlled trials with longer follow-up.

One implication is that ONB should not be seen as a cost to the NHS if Botox is not available, since its efficacy is unproven.

A newer intervention is occipital nerve stimulation, with leads implanted subcutaneously in the back of the neck, with the stimulating device implanted in abdomen or buttocks. Saper et al carried out a feasibility/exploratory RCT, sponsored and analysed by Medtronic, with two control groups, one having medical management, the other having the device implanted but not activated (it operated for one minute per day).[93] The active group had stimulation adjusted by the patients.

The intervention group did much better with a reduction in headache days of 6.7 days per month, compared to only 1.5 days in the "implanted placebo" group. The lack of placebo effect in the latter group is surprising, given the marked placebo effect in the Botox trials.

The ERG suggests that the Health Technology Assessment Programme should commission a systematic review of all therapies for chronic migraine.

7.2 Other reviews of Botox in CM

The Australian Public Assessment Record (AusPAR) for the Therapeutic Goods Administration makes some useful points.[42] The Clinical Evaluation Report recommended rejection of Botox on a number of grounds;

- Those on Botox often knew they were. In the phase 2 trials, 85-86% of those on Botox guessed that they were, whereas only 58% of those on placebo guessed they were (AusPAR table 12).

- A second point is on dosage. In the early part of the report, they comment on the lack of rationale for the dosage used in the PREEMPT trial; *"No adequate dose-finding results under-pinning the pivotal studies" "Dosage used in the pivotal studies appeared to be arbitrary".*
- The population in the pivotal trials did not entirely match the ICHD-2 definition of chronic migraine. Firstly, the baseline observation period was one month, rather than the 3 months in the ICHD-2 definition. Secondly, medication over-use was common in the PREEMPT trials. Patients were included with medication overuse, without first determining if their headaches resolved with medication withdrawal. So some may have had medication overuse headache rather than, or in addition to, migraine. This means that the observed effect in the trial could be due to an effect on MOH rather than migraine.
- There was a lack of evidence of continued efficacy

As regards this last point, there are few long-term follow-up studies. One small one was reported by Rothrock and colleagues.[79] They followed 100 consecutive patients, who had responded well to Botox treatment as per the PREEMPT protocol. After 2 years,

- 8 had relapsed despite treatment and were back in CM
- 24 had done well and come off treatment, remaining free of headaches for 6 months or more. They had had an average of 5 injections (range 2 to 8)
- 68 had responded but had required on-going maintenance Botox, roughly every 3 months.

The AusPAR clinical evaluation report concluded that "*evidence of efficacy in a well defined population is inadequate"*. There were no concerns about safety.

Allergan submitted some supplementary data, including arguments for the being "within the range of > 150 and <200 units". These do not appear convincing to the ERG.

Allergan presented the results as a responder analysis, which did not appear to have been done in the first submission, and made the case that the proportion of patients who were good responders was greater in the Botox arm. This seem to have carried more weight with the AusPAR evaluator, though it was noted that the number needed to treat to get a 50% reduction in headache days was 8.3. The report notes that this NNT is higher than most NNTs in migraine prophylaxis, which are in the range 2 to 5 (AuspAR report table 32). However the NNT argument would not be relevant to patients in the NICE scope group, who have tried and failed on oral prophylactics.

Allergan also provided a number of arguments regarding unblinding. These did not remove the doubts of the AusPAR evaluator. However the final verdict was that the submission be approved. The AusPAR review noted some advantages of Botox therapy;

- Speed of onset of effect

- Duration of effect
- Convenience to patients, and assurance of adherence (compared to taking oral prophylactics)
- No problems with interactions with other medications (bearing in mind co-morbitidies)

The **Scottish Medicines Consortium** issued guidance on Botox for prophylaxis in April 2011.[94] Botox was not recommended. The guidance refers to weaknesses in the clinical effectiveness data. The section on clinical effectiveness makes a number of points, including;

- Uncertainty about the effectiveness of the comparator, which in the model was based on the placebo arms of the PREEMPT trials. The SMC noted that if changes were made to the base case effectiveness for the comparator (best supportive care - BSC) the QALY gain could be reduced by 25% and the cost per QALY rose to around £25,000.
- The assumptions on resource use were deemed to be wrong. Increasing the duration of an OP appointment to 45 minutes and reducing the OP visits in the BSC comparator by 30% increased the ICER to over £25k
- Uncertainty about continuing efficacy
- Alternative approaches to mapping of utilities gave much higher ICERs

We note that in the SMC document, there is mention of later injections of Botox being given at 18 week intervals. If that was effective, the cost would be reduced.

MHRA

A thorough review of an Allergan submission is provided by the MHTA, with a range of useful comments by the MHRA assessor.[30] Selected points are;

- There were issues around definitions, and inclusions of people with medication over-use, but the assessor noted that the IHS Clinical Trials Subcommittee has recommended that medication over-users should be included in trials as long as randomisation is stratified – as was done in the PREEMPT trials. The assessor concluded that the trial patients were representative of people with chronic migraine.
- The trials were conducted at a time when definitions were changing, and guidelines for trials were being developed (the guidelines were published after the trials had finished).
- The assessor noted that the combination of fewer headaches days but not fewer episodes, could imply that duration of episodes had been reduced
- The switch of primary endpoint in PREEMPT 2 was considered justified
- Because of the marked placebo effect, the benefits of Botox were modest, even when statistically significant

- The assessor note the lack of evidence in men, who made up only 14% of the PREEMPT recruits. The difference of 0.7 headache days was not statistically significant, and there was no difference in the number of episodes.
- Botox was accepted as safe.

7.3 Providing placebo?

There is an issue over what would be provided if Botox were not. There was clearly a large placebo effect in the trials, and indeed most (or perhaps even all) of the benefit came from the placebo effect. If Botox was not recommended for use, the biggest loss, going from the trials, would be the placebo effect. How could the NHS provide this? Would it be ethical to provide a placebo, which would have to involve injections (acupuncture?).

7.4 Marginal costs and benefits

Dosages – why 155 Units?

Silberstein 2005 reported the results of a 4-armed trial of three doses of Botox (225 U, 150 U, 75 U) and placebo injections in chronic daily headache. It was sponsored by Allergan.[65] The Botox was given into 20 sites across 7 muscle areas.

The paper states that headaches could include any combination of migraines including episodic and chronic tension-type headaches, but recruits experienced more than 15 days in a 30-day baseline screening period. Of the 69% who had headaches classified, 77% were classed as having transformed migraine.

About half the patients were taking prophylactic medications, and about 40% were classed as over-using analgesics for acute headaches.

Randomisation was stratified according to whether patients had been responders or non-responders to placebo in the placebo run-in phase, which lasted 30 days. The phases of the trial included;

- A 30-day screening period to assess eligibility
- The 30 day placebo run-in period
- A 9-month treatment period with three 3-month cycles.

The primary outcome was the number of headache-free days in days 151 to 180 in the placebo-non-responder group, which comprised 77% of recruits. The secondary outcome was the proportion of

patients with a decrease of at least 50% in the frequency of headache days over the same period, again in the placebo non-responder group.

For the primary outcome, the reductions in headache days were 6 for the 225 Unit group, 7.9 for the 150 U, 7.9 for the 75 U, and 8 for placebo, so the primary outcome was not met. There were very large confidence intervals; p =0.44. Curiously, at day 30, the group that did best was the 75 U one (an increase of 5 headache free days, whereas all the other groups had increases of around 3; p = 0.01).

There was no significant difference in the primary outcome in the placebo-responder group.

Hence for the key outcome, there was no difference at 180 days amongst the 225, 150 Unit and placebo groups.

There appeared to be no interaction between the placebo responder/non-responder status and outcomes, so the groups were pooled for further analysis.

The authors commented that the response rate in the placebo group was greater than expected, noting that a high placebo response rate is common in headache studies. No Botox group was consistently better than placebo.

Patients were allowed to use analgesics as required, and as time went on, the placebo group used more than the three Botox groups. At day 180, the percentages over-using analgesics were about 20% in the placebo group, and about 10% in the Botox 150 U and 225 U groups (from graph – Figure 4 of paper).

Much higher adverse event rates were seen in the Botox groups, especially in the 225 U group. The commonest AEs were muscular weakness, neck pain and neck rigidity – respectively 26%, 22% and 8% in the 150u group, versus 0.6%, 2% and 0.6% in the placebo group. But few patients discontinued the study because of AEs

The authors note the results of the trial by Mathew et al that used the "follow-the-pain" (FTP) approach with a Botox dose of 190 U. They then suggest that "an effective Botox dose may be in the range 150 to 225u".[64]

Mathew et al carried out a 2-arm RCT of Botox versus placebo in patients with chronic daily headaches, very similar in some ways to the Silberstein trial, probably because both were sponsored by Allergan. The main difference was in the dosing – a single Botox arm in which patients got injections at baseline, 90 and 180 days, of 105 to 260 Units, into a minimum of six muscle areas, with 23 to 58 injection sites. Dosage was decided by the physician based on a FTP approach.

At the end of the placebo run-in period, participants were classified as responders to placebo or non-responders and randomisation was stratified by these groups.

Patients were asked what treatment they thought they had received. Their guesses were correct 70% after the first treatment, 65% after the second and 69% after the third. A 50% correct rate would have been expected by chance. Surprisingly, the authors report that even patients who correctly guessed that they had received placebo injections, had "dramatic" improvements.

The primary outcome was headache-free days. In those classed as placebo non-responders after the run-in, the mean increases in the numbers of headache free days at day 180 were 6.7 days for the Botox group and 5.2 days for the placebo group. In those who were placebo-responders, the mean increases were 12.1 days for Botox and 10.5 for placebo. The differences were not statistically significant.

The mean dose received was around 191 units, with a range of 105 to 260 units. No data are given in the paper on dose-response curves.

Blumenfeld and colleagues review the evolution of Botox therapy for CM.[95] They place most emphasis on the Silberstein and Mathew trials, noting that,

"Overall, these exploratory phase 2 studies provided guidance and shaped the study design and the injection paradigm of the phase 3 PREEMPT clinical program."

They comment that in the Silberstein trial, the 225 U and 150 U arms were more effective, but that the 225 arm had more adverse effects. The "more effective" comment was presumably based on secondary endpoints such as number of headaches and use of analgesics, since there was no difference in the primary outcome. They go on to say that,

"*Therefore it was determined that the optimal total dose to maximise efficacy and tolerability was within the range of >150 U and <200 U."*

However it is not clear why the 150 was converted to >150 U.

The importance of this point relates to the costs. Vial costs (Table 1.10 of the Allergan submission) are;

50 Units £77.50

100 Units £138.20

200 Units £276.40

Hence adopting a minimum dose of 155 U means that a 200 U vial must be used, whereas were a dose of 150 U to be used, the cost would be £215.70. The marginal cost of the extra 5 Units is £60.70, but the marginal benefit is not known. The marginal cost is 28%, the marginal dose is 3%.

The ERG has seen no good evidence for the move from the 150 Units in the Silberstein trial to 155 or more in the PREEMPT trials, and has doubts about the cost-effectiveness of the 155 U dose compared to the 150 U one.

The ERG notes that in a small uncontrolled study, Oterino et al used a starting dose of 100 Units, with fewer injections sites (20) and only increased it in non-responders. This study had only 35 patients, and most did not respond.[96] But some did, so this might be another approach if confirmed by a larger trial.

Adding the "follow-the-pain" dosage

In the PREEMPT trials, the basic dosage was 155 Units, but the treating physicians could administer up to 40 more units if they thought that was appropriate.

As shown in Table 55, out of a total of 1384 PREEMPT subjects, 49% (331 Botox and 339 placebo) were allocated to FSFD only and 51% (354 Botox and 355 placebo) received additional FTP dosages during the RCT phase.

Table 55 Proportion of patients in PREEMPT trials receiving the FSFD and FSFD+ FTP doses

	Botox	Placebo	Total number (%)
FSFD only	331	339	670 (48.6%)
FSFD + FTP	354	355	709 (51.4%)
	685	694	1379

Table 56 Change in mean frequency of headache days

FSFD only sub-group	Botox (n=331)	Placebo (n=339)	Botox-placebo	
Decrease from baseline in mean frequency of headache days	9.0	6.7	2.3	<0.001
All PREEMPT participants	Botox (n=688)	Placebo n=696)		
Decrease from baseline in mean frequency of headache days	8.4	6.6	1.8	<0.001

As shown in Table 56, the FSFD only sub-group appears to do better on 155 U compared to whole group on the primary outcome of decrease from baseline in mean frequency of headache days.

Health related quality of life

Results are provided in Table 57 below for the change from baseline FSFD group and for the whole group, as shown below for MSQ change scores from baseline.

Table 57 Change in MSQ scores in group receiving only FSFD injections and all PREEMPT participants

FSFD only sub-group	Botox (n=331)	Placebo (n=339)	Botox-placebo	P value
MSQ role restrictive (RFR)	17.6	8.8	8.8	<0.001
MSQ role preventative (RFP)	13.8	6.1	7.7	<0.001
MSQ emotional functioning (EF)	18.4	8.9	9.5	<0.001
All PREEMPT participants	Botox (n=688)	Placebo (n=696)		
MSQ role restrictive (RFR)	17.0	8.6	8.4	<0.001
MSQ role preventative (RFP)	13.1	6.4	6.7	<0.001
MSQ emotional functioning (EF)	17.9	9.5	8.4	<0.001

The implication of these results is that the whole group does no better than the FSFD sub-group, implying that the FSDF+ FTP group does not appear to do any better in terms of quality of life than the FSFD group.

As shown in Table 58 for HIT-6 scores, no differences were seen between the FSFD only sub-group and all PREEMPT patients.

Table 58 Change in HIT-6 scores in group receiving only FSFD injections and all PREEMPT participants

FSFD only sub-group	Botox (n=331)	Placebo (n=339)	
Proportion with 5-point decrease* from baseline in total HIT-6 scores	40.0%	25.4%	significantly higher - p value not given
All PREEMPT participants			
Proportion of patients with a ≥5 point decrease in total HIT-6 score at 24 weeks	40.8%	25.3%	0.001

**A ≥5 point decrease in total HIT-6 is a meaningful change for an individual patient.*

These results might suggest that adding the extra FTP injections does not confer additional benefit. However, an alternative explanation is that the patients who received FTP were more severely affected. They were not randomly allocated to FTP, because the PREEMPT trials were not designed to evaluate the marginal benefits of adding FTP doses to FSFD. To do that, it would have been necessary to randomise suitable patients to FSFD or to FSFD + FTP.

Table 59 below shows that the FSFD + FTP group tended to do less well than the FSFD group, in most variables, both on absolute improvements and relative to placebo. However it is noteworthy that there was little difference in the absolute improvements on placebo in the two groups, which may go against the hypothesis that they were more severely affected.

Pending an RCT of FSFD versus FSFD + FTP, we cannot say whether there are significant marginal benefits from adding FTP. The marginal costs will be only in time and adverse effects if the 155u FSFD dose is used, because no extra vial would be opened.

Table 59 Efficacy of FSFD versus FSFD+FTP

	FSFD				**FSFD + FTP**			
Efficacy variable (per 28 days)	Botox n=331	Placebo n=339	Botox-Placebo	p value	Botox n=354	Placebo n=355	Botox-Placebo	p value
Frequency of headache days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Frequency of headache episodes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Frequency of migraine days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Frequency of moderate/severe headache days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total cumulative hours of headache on headache days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Freqeuency of migraine episodes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Frequency of AHPM intakes	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Frequency of AHPM days	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 60 shows slightly more AEs with the FSFD + FTP dosing.

Table 60 Adverse events in FSFD versus FSFD+FTP groups

	FSFD			FSFD + FTP		
	Botox (n=332) n (%)	Placebo (n=338) n (%)	Total (n=670) n (%)	Botox (n=355) n (%)	Placebo (n=354) n (%)	Total (n=709) n (%)
All adverse events* (AEs)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-related AEs+	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Serious AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment related AEs+	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Discontinuation related AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Deaths	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*All AEs include all reported events, regardless of relationship to treatment.

†Treatment-related AEs are those that in the investigator's opinion may have been caused by the study medication with reasonable possibility. The one treatment-related serious AE was migraine requiring hospitalization

7.5 Conclusions

The ERG considers that the Allergan modelling follows a logical sequence, summarised in Figure 21.

However at various stages, the ERG has some concerns, also summarised in Figure 21. In brief;

- The mix of patients in the trials, with 65% over-users of acute medication, does not match the NICE decision problem group, in which patients are expected to have failed on three or more prophylactic drugs, and have medication over-use appropriately managed. There may be issues around definitions here. "Failure" should be defined either as inability to tolerate side-effects of prophylactics, or when they are ineffective after a reasonable trial – say a minimum of 3 months. It should be noted that clinicians have to rely completely on patients' reported symptoms. This may also be relevant to stopping rules – if patients want to continue Botox treatment, they can simply report still having more than 15 headache days.
- The rationale for the standard FSFD of 155 units is not clear. We see no good evidence that 155 units is cost-effective compared to 150 units – the extra 3% dose costs an extra 28% for drug alone. Nor is there evidence for the cost-effectiveness of adding the FTP doses.
- The effect of unblinding amongst the Botox group is uncertain, but it is possible that some of the relatively small observed difference between Botox and placebo (1.8 days) could be due to differences in the size of the placebo effect.
- We wondered about how well the MSQ and HIT-6 applied to chronic migraine, since they were developed for episodic migraine. Is it possible that they under-estimate the impact of chronic migraine? 93% of patients were in the severely impacted group at baseline (HIT-6 of 60 or more) and perhaps it would be worth another sub-division into very severe with HIT-6 of 65 or over.
- One of our main reservations was about the cost estimates used in the modelling, and in particular the inclusion of hospital admission and A and E costs. We think the costs of migraine are over-estimated.
- However we also think that the cost of Botox could be reduced, by having a dose not exceeding 150 units (perhaps using the Oterino et al approach of trying 100 units first), by not adding the FTP dosing, and by extending the interval between maintenance doses, perhaps to 18 weeks. If NICE approves Botox, further research into dosages would be worthwhile.

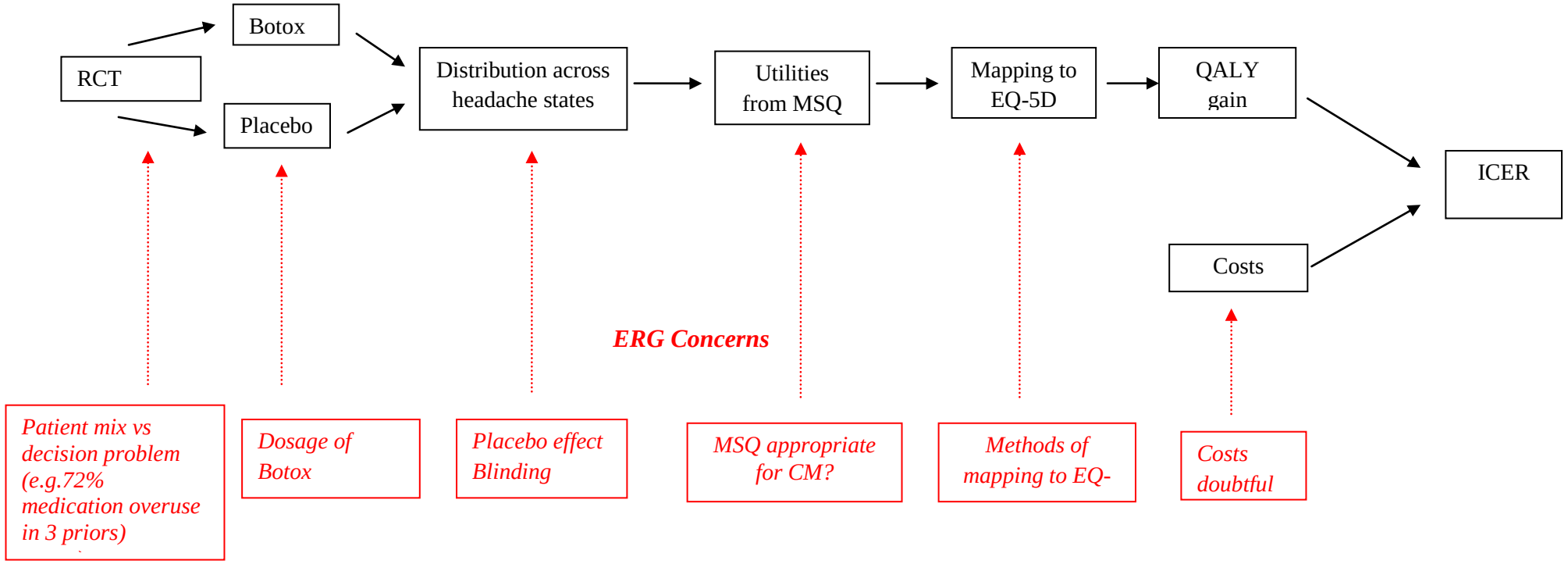


Figure 21 Summary of derivation of ICERs: ERG concerns

7.6 Research needs

If Botox were to be approved by NICE, we suggest that further trials be done, including;

- A range of doses, comparing the PREEMPT 155 unit FSFD regimen against doses below 150 units. Would a lower dose be equally effective but at lower cost?
- An RCT of FSFD versus FSFD plus the FTP extra dosage. Is the extra FTP dosage cost-effective?

One problem that would need to be addressed is how to overcome the problem of unblinding in the Botox arm. One option might be to use an active control, but with a very short duration of action. One possibility would be low dose pancuronium, a muscle relaxant.

References

1. Katsarava Z, Manack A, Yoon MS, Obermann M, Becker H, Dommes P *et al*. Chronic migraine: classification and comparisons. *Cephalalgia* 2011;31:520-9.

2. Lima MM, Padula NA, Santos LC, Oliveira LD, Agapejev S, Padovani C. Critical analysis of the international classification of headache disorders diagnostic criteria (ICHD I-1988) and (ICHD II-2004), for migraine in children and adolescents. *Cephalalgia* 2005;25:1042-7.

3. Olesen J. *Headache classification and epidemiology*. New York: Raven Presss, 1994.

4. Lipton RB. Chronic migraine, classification, differential diagnosis, and epidemiology. *Headache* 2011;51:77-83.

5. Manzoni GC, Granella F, Sandrini G, Cavallini A, Zanferrari C, Nappi G. Classification of chronic daily headache by International Headache Society criteria: limits and new proposals. *Cephalalgia* 1995;15:37-43.

6. Classification Committee of the International Headache Society. The International Classification of Headache Disorders: 2nd edition. *Cephalalgia* 2004;24 Suppl 1:9-160.

7. Manack A, Turkel C, Silberstein S. The evolution of chronic migraine: classification and nomenclature. *Headache* 2009;49:1206-13.

8. Bigal ME, Tepper SJ, Sheftell FD, Rapoport AM, Lipton RB. Field testing alternative criteria for chronic migraine. *Cephalalgia* 2006;26:477-82.

9. Olesen J, Bousser MG, Diener HC, Dodick D, First M, Goadsby PJ *et al*. New appendix criteria open for a broader concept of chronic migraine. *Cephalalgia* 2006;26:742-6.

10. Manzoni GC, Bonavita V, Bussone G, Cortelli P, Narbone MC, Cevoli S *et al*. Chronic migraine classification: current knowledge and future perspectives. *J Headache Pain* 2011;12:585-95.

11. Sun-Edelstein C, Bigal ME, Rapoport AM. Chronic migraine and medication overuse headache: clarifying the current International Headache Society classification criteria. [Review] [23 refs]. *Cephalalgia* 2009;29:445-52.

12. Silberstein S, Tfelt-Hansen P, Dodick DW, Limmroth V, Lipton RB, Pascual J *et al*. Guidelines for controlled trials of prophylactic treatment of chronic migraine in adults. *Cephalalgia* 2008;28:484-95.

13. Burstein R, Cutrer FM, Yarnitsky D. The development of cutaneous allodynia during a migraine attack: clinical evidence for the sequential recruitment of spinal and supraspinla nociceptive neurons in migraine. *Brain* 2000;123:1703-9.

14. Pocock SJ. *Clinical trials : a practical approach*. Chichester: Wiley, 1983.

15. International Headache Society. ICHD-II classification. Appendix A1.5.1 - Chronic Migraine. 2011 http://ihs-classification.org/en/02_klassifikation/05_anhang/01.05.01_anhang.html (accessed 8 December 2011)

16. Steiner TJ, Scher AI, Stewart WF, Kolodner K, Liberman J, Lipton RB. The prevalence and disability burden of adult migraine in England and their relationships to age, gender and ethnicity. *Cephalalgia* 2003;23:519-27.

17. World Health Organization. Headache disorders Fact sheet N°277. 2004, March http://www.who.int/mediacentre/factsheets/fs277/en/ (accessed 8 December 2011)

18. Natoli JL, Manack A, Dean BB. Chronic migraine criteria and their effects on global prevalence estimates: A systematic review. *Value in Health* 2009;Conference: ISPOR 14th Annual International Meeting Orlando, FL United States. Conference Start: 20090516 Conference End: 20090520. Conference Publication:3-A190.

19. Stovner L, Hagen K, Jensen R, Katsarava Z, Lipton R, Scher A *et al*. The global burden of headache: a documentation of headache prevalence and disability worldwide. *Cephalalgia* 2007;27:193-210.

20. Straube A, Pfaffenrath V, Ladwig KH, Meisinger C, Hoffmann W, Fendrich K *et al*. Prevalence of chronic migraine and medication overuse headache in Germany--the German DMKG headache study. *Cephalalgia* 2010;30:207-13.

21. Cassidy EM, Tomkins E, Hardiman O, O'Keane V. Factors associated with burden of primary headache in a specialty clinic. *Headache* 2003;43:638-44.

22. Buse DC, Manack A, Serrano D, Turkel C, Lipton RB. Sociodemographic and comorbidity profiles of chronic migraine and episodic migraine sufferers. *J Neurol Neurosurg Psychiatry* 2010;81:428-32.

23. Blumenfeld AM, Varon SF, Wilcox TK, Buse DC, Kawata AK, Manack A *et al*. Disability, HRQoL and resource use among chronic and episodic migraineurs: results from the International Burden of Migraine Study (IBMS). *Cephalalgia* 2011;31:301-15.

24. Lanteri-Minet M, Duru G, Mudge M, Cottrell S. Quality of life impairment, disability and economic burden associated with chronic daily headache, focusing on chronic migraine with or without medication overuse: a systematic review. *Cephalalgia* 2011;31:837-50.

25. British Association for the Study of Headache. Guidelines for All Healthcare Professionals in the Diagnosis and Management of Migraine, Tension-Type Headache, Cluster Headache, Medication-Overuse Headache. 3rd edtion. 2010 http://217.174.249.183/upload/NS_BASH/2010_BASH_Guidelines.pdf (accessed 8 December 2011)

26. Bigal ME, Serrano D, Reed M, Lipton RB. Chronic migraine in the population. *Neurology* 2008;71:559-66.

27. Cady RK, Schreiber CP, Porter JA, Blumenfeld AM, Farmer KU. A multi-center double-blind pilot comparison of onabotulinumtoxinA and topiramate for the prophylactic treatment of chronic migraine. *Headache* 2011;51:21-32.

28. Mathew NT, Jaffri SF. A double-blind comparison of onabotulinumtoxina (BOTOX) and topiramate (TOPAMAX) for the prophylactic treatment of chronic migraine: a pilot study. *Headache* 2009;49:1466-78.

29. NHS. Clinical Knowledge Summaries. 2011 http://www.cks.nhs.uk/home (accessed 9 December 2011)

30. Medicines and Healthcare products Regulatory Agency. Botox (Botulinum toxin type A): UKPAR. 2010, September http://www.mhra.gov.uk/home/groups/par/documents/websiteresources/con108643.pdf (accessed 9 December 2011)

31. SIGN. Diagnosis and management of headache in adults: a national clinical guideline. Guideline No. 107. 2008Scottish Intercollegiate Guidelines Network. http://www.sign.ac.uk/guidelines/fulltext/107/index.html (accessed 9 December 2011)

32. Freitag FG, Diamond S, Diamond M, Urban G. Botulinum Toxin Type A in the treatment of chronic migraine without medication overuse. *Headache* 2008;48:201-9.

33. Dodick DW, Turkel CC, DeGryse RE, Aurora SK, Silberstein SD, Lipton RB *et al*. OnabotulinumtoxinA for treatment of chronic migraine: pooled results from the double-blind, randomized, placebo-controlled phases of the PREEMPT clinical program. *Headache* 2010;50:921-36.

34. Aurora SK, Dodick DW, Turkel CC, DeGryse RE, Silberstein SD, Lipton RB *et al*. OnabotulinumtoxinA for treatment of chronic migraine: results from the double-blind, randomized, placebo-controlled phase of the PREEMPT 1 trial. *Cephalalgia* 2010;30:793-803.

35. Aurora SK, Winner P, Freeman MC, Spierings EL, Heiring JO, DeGryse RE *et al*. OnabotulinumtoxinA for Treatment of Chronic Migraine: Pooled Analyses of the 56-Week PREEMPT Clinical Program. *Headache* 2011;51:1358-73.

36. Diener HC, Dodick DW, Aurora SK, Turkel CC, DeGryse RE, Lipton RB *et al*. OnabotulinumtoxinA for treatment of chronic migraine: results from the double-blind, randomized, placebo-controlled phase of the PREEMPT 2 trial. *Cephalalgia* 2010;30:804-14.

37. Markert RJ, Solomon GD. OnabotulinumtoxinA for Treatment of Chronic Migraine. *Headache* 2011;51:1002-3.

38. Dodick DW, Turkel CC, DeGryse RE, Aurora SK, Silberstein SD, Lipton RB *et al*. OnabotulinumtoxinA for Treatment of Chronic Migraine: A Response. *Headache* 2011;51:1003-8.

39. U.S.Department of Health and Human Services F. Guidance for Industry Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. 1998 http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078749.pdf (accessed 21 November 2011)

40. Aurora SK, Dodick DW, Varon SF, DeGryse RE, Turkel CC. OnabotulinumtoxinA for Chronic Migraine: Effect of Fixed-Site, Fixed-Dose Injection Paradigm on Health-Related Quality of Life in PREEMPT Clinical Program. *Neurology* 2011;76:A443.

41. Diener HC, Dodick DW, DeGryse RE, Turkel CC. OnabotulinumtoxinA for Chronic Migraine: Safety and Efficacy of Fixed-Site, Fixed-Dose Injection Paradigm in PREEMPT Clinical Program. *Neurology* 2011;76:A267.

42. Therapeutic Goods Administration DoHaAAG. Australian Public Assessment Report for Botulinum toxin Type A. 2011, June http://www.tga.gov.au/pdf/auspar/auspar-botox.pdf (accessed 9 December 2011)

43. Rahimtoola H, Buurma H, Tijssen CC, Leufkens HG, Egberts AC. Migraine prophylactic medication usage patterns in The Netherlands. *Cephalalgia* 2003;23:293-301.

44. Rains JC, Lipchik GL, Penzien DB. Behavioral facilitation of medical treatment for headache--part I: Review of headache treatment compliance. *Headache* 2006;46:1387-94.

45. Yaldo AZ, Wertz DA, Rupnow MF, Quimbo RM. Persistence with migraine prophylactic treatment and acute migraine medication utilization in the managed care setting. *Clin Ther* 2008;30:2452-60.

46. Hedenrud T, Jonsson P, Linde M. Beliefs about medicines and adherence among Swedish migraineurs. *Ann Pharmacother* 2008;42:39-45.

47. Evans RW, Linde M. Expert opinion: adherence to prophylactic migraine medication. *Headache* 2009;49:1054-8.

48. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P *et al*. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain* 2009;146:238-44.

49. Dodick D, Diener H-C, Turkel C, DeGryse R, Brin MF. Onabotulinumtoxin a for treatment of chronic migraine: 75% responder analysis results from the double-blind, randomized, placebo-controlled phase of the preempt clinical program. *Cephalalgia* 2011;31:87.

50. Yang M, Rendas-Baum R, Varon SF, Kosinski M. Validation of the Headache Impact Test (HIT-6) across episodic and chronic migraine. *Cephalalgia* 2011;31:357-67.

51. Coeytaux RR, Kaufman JS, Chao R, Mann JD, Devellis RF. Four methods of estimating the minimal important difference score were compared to establish a clinically significant change in Headache Impact Test. *J Clin Epidemiol* 2006;59:374-80.

52. Lipton RB, Varon SF, Grosberg B, McAllister PJ, Freitag F, Aurora SK *et al*. OnabotulinumtoxinA improves quality of life and reduces impact of chronic migraine. *Neurology* 2011;77:1465-72.

53. Bagley CL, Rendas-Baum R, Maglinte GA, Yang M, Varon SF, Lee J *et al*. Validating Migraine-Specific Quality of Life Questionnaire v2.1 in Episodic and Chronic Migraine. *Headache* 2011.

54. Cole JC, Lin P, Rupnow MF. Minimal important differences in the Migraine-Specific Quality of Life Questionnaire (MSQ) version. *Cephalalgia* 2009;29:1180-7.

55. Dodick DW, Silberstein S, Saper J, Freitag FG, Cady RK, Rapoport AM *et al*. The impact of topiramate on health-related quality of life indicators in chronic migraine. *Headache* 2007;47:1398-408.

56. FDA. BOTOX (onabotulinumtoxinA) for injection, for intramuscular, intradetrusor, or intradermal use. HIGHLIGHTS OF PRESCRIBING INFORMATION . 2011 http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/103000s5232lbl.pdf (accessed 9 December 2011)

57. Rossi RP, Strax TE, Di RA. Severe Dysphagia after botulinum toxin B injection to the lower limbs and lumbar paraspinal muscles. *American Journal of Physical Medicine & Rehabilitation* 2006;85:1011-3.

58. Thobois S, Broussolle E, Toureille L, Vial C. Severe dysphagia after botulinum toxin injection for cervical dystonia in multiple system atrophy. *Movement Disorders* 2001;16:764-5.

59. Tuite PJ, Lang AE. Severe and prolonged dysphagia complicating botulinum toxin A injections for dystonia in Machado-Joseph disease. *Neurology* 1996;46:846.

60. Kwek AB, Tan EK, Luman W. Dysphagia as a side effect of botulinum toxin injection. *Medical Journal of Malaysia* 2004;59:544-6.

61. Comella CL, Tanner CM, DeFoor-Hill L, Smith C. Dysphagia after botulinum toxin injections for spasmodic torticollis: clinical and radiologic findings. *Neurology* 1992;42:1307-10.

62. Allergan. BOTOX® Prophylaxis in Patients With Chronic Migraine. 2011 http://clinicaltrials.gov/ct2/show/NCT01432379?term=NCT01432379&rank=1

63. Chapman MA, Barron R, Tanis DC, Gill CE, Charles PD. Comparison of botulinum neurotoxin preparations for the treatment of cervical dystonia. *Clin Ther* 2007;29:1325-37.

64. Mathew NT, Frishberg BM, Gawel M, Dimitrova R, Gibson J, Turkel C *et al*. Botulinum toxin type A (BOTOX) for the prophylactic treatment of chronic daily headache: a randomized, double-blind, placebo-controlled trial. *Headache* 2005;45:293-307.

65. Silberstein SD, Stark SR, Lucas SM, Christie SN, DeGryse RE, Turkel CC *et al*. Botulinum toxin type A for the prophylactic treatment of chronic daily headache: a randomized, double-blind, placebo-controlled trial. *Mayo Clinic Proceedings* 2005;80:1126-37.

66. de Craen AJ, Tijssen JG, de GJ, Kleijnen J. Placebo effect in the acute treatment of migraine: subcutaneous placebos are better than oral placebos. *J Neurol* 2000;247:183-8.

67. van der Kuy PH, Lohman JJ. A quantification of the placebo response in migraine prophylaxis. *Cephalalgia* 2002;22:265-70.

68. Macedo A, Banos JE, Farre M. Placebo response in the prophylaxis of migraine: a meta-analysis. *European Journal of Pain: Ejp* 2008;12:68-75.

69. Macedo A, Farre M, Banos JE. A meta-analysis of the placebo response in acute migraine and how this response may be influenced by some of the characteristics of clinical trials. *Eur J Clin Pharmacol* 2006;62:161-72.

70. Kaptchuk TJ, Stason WB, Davis RB, Legedza AR, Schnyer RN, Kerr CE *et al*. Sham device v inert pill: randomised controlled trial of two placebo treatments. *BMJ* 2006;332:391-7.

71. Diener HC, Schorn CF, Bingel U, Dodick DW. The importance of placebo in headache research. *Cephalalgia* 2008;28:1003-11.

72. Speciali JG, Peres M, Bigal ME. Migraine treatment and placebo effect. *Expert Rev Neurother* 2010;10:413-9.

73. Schwedt TJ, Hentz JG, Dodick DW. Factors associated with the prophylactic effect of placebo injections in subjects enrolled in a study of botulinum toxin for migraine. *Cephalalgia* 2007;27:528-34.

74. Solomon S. Botulinum toxin for the treatment of chronic migraine: the placebo effect. *Headache* 2011;51:980-4.

75. Olesen J, Tfelt-Hansen P. Licence for Botox in so-called chronic migraine. *The Lancet* 2010;376:1825-6.

76. Payne KA, Varon SF, Kawata AK, Yeomans K, Wilcox TK, Manack A *et al*. The International Burden of Migraine Study (IBMS): study design, methodology, and baseline cohort characteristics. *Cephalalgia* 2011;31:1116-30.

77. Stokes M, Becker WJ, Lipton RB, Sullivan SD, Wilcox TK, Wells L *et al*. Cost of health care among patients with chronic and episodic migraine in Canada and the USA: results from the International Burden of Migraine Study (IBMS). *Headache* 2011;51:1058-77.

78. Munakata J, Hazard E, Serrano D, Klingman D, Rupnow MF, Tierce J *et al*. Economic burden of transformed migraine: results from the American Migraine Prevalence and Prevention (AMPP) Study. *Headache* 2009;49:498-508.

79. Rothrock JF, Andress-Rothrock D, Scanlon C, Weibelt S. Onabotulinumtoxina for the treatment of chronic migraine: Long-term outcome. *Headache* 2011;51:60.

80. Schurks M, Rist PM, Shapiro RE, Kurth T. Migraine and mortality: a systematic review and meta-analysis. *Cephalalgia* 2011;31:1301-14.

81. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010;11:215-25.

82. Brown JS, Papadopoulos G, Neumann PJ, Friedman M, Miller JD, Menzin J. Cost-effectiveness of topiramate in migraine prevention: results from a pharmacoeconomic model of topiramate treatment. *Headache* 2005;45:1012-22.

83. Bloudek LM, Hansen RN, Liu L, Batty AJ, Varon SF, Lipton RB *et al*. Health resource utilization and costs for migraneurs in Scotland. *Value in Health* 2011;14:A207-A208.

84. Campbell JC, Penzien DB, Wall EM. Evidenced-Based Guidelines For Migraine Headache: Behavioral and Physical Treatments. 2000 http://www.aan.com/professionals/practice/pdfs/gl0089.pdf (accessed 30 November 2011)

85. Linde K, Allais G, Brinkhaus B, Manheimer E, Vickers A, White AR. Acupuncture for migraine prophylaxis. *Cochrane Database Syst Rev* 2009;CD001218.

86. Bronfort G, Nilsson N, Haas M, Evans R, Goldsmith CH, Assendelft WJ *et al*. Non-invasive physical treatments for chronic/recurrent headache. *Cochrane Database of Systematic Reviews* 2004;CD001878.

87. Bennett MH, French C, Schnabel A, Wasiak J, Kranke P. Normobaric and hyperbaric oxygen therapy for migraine and cluster headache. *Cochrane Database of Systematic Reviews* 2008;CD005219.

88. Astin JA, Ernst E. The effectiveness of spinal manipulation for the treatment of headache disorders: a systematic review of randomized clinical trials. [Review] [22 refs]. *Cephalalgia* 2002;22:617-23.

89. Ashkenazi A, Matro R, Shaw JW, Abbas MA, Silberstein SD. Greater occipital nerve block using local anaesthetics alone or with triamcinolone for transformed migraine: a randomised comparative study. *J Neurol Neurosurg Psychiatry* 2008;79:415-7.

90. Tobin J, Flitman S. Occipital nerve blocks: when and what to inject? *Headache* 2009;49:1521-33.

91. Afridi SK, Shields KG, Bhola R, Goadsby PJ. Greater occipital nerve injection in primary headache syndromes--prolonged effects from a single injection. *Pain* 2006;122:126-9.

92. Weibelt S, Andress-Rothrock D, King W, Rothrock J. Suboccipital nerve blocks for suppression of chronic migraine: safety, efficacy, and predictors of outcome. *Headache* 2010;50:1041-4.

93. Saper JR, Dodick DW, Silberstein SD, McCarville S, Sun M, Goadsby PJ. Occipital nerve stimulation for the treatment of intractable chronic migraine headache: ONSTIM feasibility study. *Cephalalgia* 2011;31:271-85.

94. Scottish Medicines Consortium. botulinum toxin type a (Botox). 2011 http://www.scottishmedicines.org.uk/SMC_Advice/Advice/692_11_botulinum_toxin_type_a_BOTOX/689_11_botulinum_toxin_type_a_Botox (accessed 9 December 2011)

95. Blumenfeld AM, Silberstein SD, Dodick DW, Aurora SK, Turkel CC, Binder WJ. OnabotulinumtoxinA for Chronic Migraine: PREEMPT Trials Establish a Safe and Effective Dose and Injection Paradigm. *Neurology* 2010;74:A573.

96. Oterino A, Ramon C, Pascual J. Experience with onabotulinumtoxinA (BOTOX) in chronic refractory migraine: focus on severe attacks. *Journal of Headache & Pain* 2011;12:235-8.

Appendices

Appendix 1 The Freitag 2008 study

The aim of this pilot study by Freitag et al[32] was to assess the efficacy and safety of Botox (Botulinum Toxin Type A) compared with placebo in the treatment of chronic migraine without medication overuse. (Note that at time of this trial, termination of the offending agent was required to prove the diagnosis of medication overuse. This definition existed until the most recent revision of the classification of MOH).

The study design was a double-blind placebo-controlled randomized trial. The patients were aged 18-65 and were required to have a 6-month history, prior to baseline, of CM. In addition, they were required to have migraine episodes meeting the criteria 1.1 or 1.2 of the ICHD-I and 15 headache days during the prospective baseline phase. If they were on preventive medications, then they must have been on stable doses of preventive medications for 60 days prior to study entry and be willing to remain on them at those same doses for the duration of the study
Patients entered a 28 day screening phase, and if they met inclusion/exclusion criteria were randomised to their respective treatment groups for baseline diary phase.

They received treatment with 100 units of Botox administered in a fixed dose and site paradigm or placebo injection. They were injected in 5 different sites, with a total of 22 injections, and seen at 4 week intervals for the remaining 16 weeks of the study.

Of the 86 patients enrolled, 26 discontinued prior to randomisation (failure to meet inclusion criteria). Of the 60 remaining patients, 41 patients were treated with the study medication or placebo (19 failed to randomise because of medication overuse during baseline phase). Therefore, a total of 41 patients were randomised and received treatment with either Botox or placebo; 41 were included in ITT analysis. Five patients failed to complete the study, which lasted 4 months after the study medication was injected. Therefore, only 18 in each group were analysed

The mean age was 42 years and overall 73% were female). The primary endpoint was the reduction in migraine headache episodes per month. The migraine attack frequency in the Botox arm declined from 13.8 to 10.1 attacks per month, compared to the placebo arm, which rose from 14.6 at baseline to 15.4 at the end of the study (P=0.001 between groups).

Six patients on Botox compared to 3 patients on placebo had at least a 50% reduction in their migraine episodes. None of the secondary endpoints, including total headache days, headache index, and quality of life measures, showed a statistically significant difference between the groups Acute medication use declined in the Botox group, but not in placebo group, but this difference was not significant. Adverse events were rare and similar in both treatment groups.

Appendix 2 Baseline characteristics of PREEMPT 1 and PREEMPT 2

Table 61 Baseline disease characteristics for primary and secondary efficacy variables in PREEMPT 1 and PREEMPT 2 Studies

	PREEMPT 1			PREEMPT 2		
Characteristics	**Botox (n=341)**	**Placebo (n=338)**	**p-value**	**Botox (n=347)**	**Placebo n=358**	**p-value**
Mean age (yrs)	41.2	42.1	0.317	41	40.9	0.849
Mean time since onset of chronic migraine (yrs)	20.3	20.6	0.839	18.5	17.6	0.279
Women (%)	89.1	85.8	0.187	86.2	84.6	0.565
Caucasian (%)	89.4	91.4	0.381	89.9	89.7	0.913
Mean body mass index (kg/m2)	26.7	27.3	0.147	26.7	26.1	0.305
Mean headache episodes during baseline (SD)	12.3 (5.23)	13.4 (5.71)	0.023	12.0 (5.27)	12.7 (5.29)	0.067
Mean headache days during baseline (SD)	20.0 (3.73)	19.8 (3.71)	0.571	19.9 (3.63)	19.7 (3.65)	0.682
Mean migraine days during baseline (SD)	19.1 (4.04)	19.1 (4.05)	0.978	19.2 (3.94)	18.7 (4.05)	0.156
Mean migraine episodes during baseline (SD)	11.5 (5.06)	12.7 (5.72)	0.006	11.3 (4.99)	11.7 (5.08)	0.067
Mean moderate/severe headache days during baseline (SD)	18.1 (4.22)	18.3 (4.23)	0.674	18.1 (4.03)	17.7 (4.26)	0.333

Cumulative headache hours occurring on headache days during baseline (SD)	295.7 (116.81)	274.9 (110.90)	0.022	296.2 (121.04)	287.2 (118.09)	0.311
Patients who overused acute headache pain medications during baseline (%)	66.3	69.8	0.322	63.4	62.6	0.819
Patients who had previously used ≥1 headache prophylaxis medication (%)	59.5	64.2	0.21	64	66.2	0.536
Patients with severe (≥60) HIT-6 score	94.4	94.7	0.888	92.5	90.8	0.408

At baseline, patients in PREEMPT 1 receiving Botox had significantly fewer headache episodes (12.3 Botox vs. 13.4 placebo; p=0.023) and migraine episodes (11.5 Botox vs. 12.7 placebo; p=0.006) than patients receiving placebo, and significantly more cumulative hours of headache occurring on headache days (p=0.022) (see Appendix). PREEMPT 2 had no significant baseline differences

Most patients overused acute pain medications during the 28-day baseline (65.5% (62.6%-69.8% of the total population of the pooled analyses) and had previously tried at least one prophylactic medication (63.5% [95% CI 59.5%-66.2]) of the total population of the pooled analyses), highlighting the severity of their suffering.

Appendix 3 Week 24 results for PREEMPT 1 and PREEMPT 2

Table 62 Mean change from baseline at week 24 in PREEMPT 1 and PREEMPT 2

Efficacy Variable (per 28 days)	PREEMPT 1				PREEMPT 2			
	Botox (N= 341)	placebo (N=338)	Botox-placebo	P-value	Botox (N = 347)	placebo (N=358)	Botox-placebo	P-value
Frequency of headache days [a]	-7.8	-6.4	-1.4	0.006	-9	-6.7	-2.3	<0.001
Frequency of migraine/probable migraine headache days	-7.6	-6.1	-1.5	0.002	-8.7	-6.3	-2.4	<0.001
Frequency of moderate/severe headache days	-7.2	-5.8	-1.4	0.004	-8.3	-5.8	-2.5	<0.001
Total cumulative hours of headache on headache days	-106.7	-70.4	-36.3	0.003	-132.41	-90.01	-42.4	<0.001
Proportion of patients with severe HIT-6 category scores, %	68.9%	79.9%	-11.0%	0.001	66.3%	76.5%	-10.2%	0.003
Frequency of headache episodes[b]	-5.2	-5.3	0.1	0.344	-5.3	-4.6	-0.7	0.003
Frequency of migraine/probable migraine headache episodes	-4.8	-4.9	0.1	0.206	-4.9	-4.2	-0.7	0.003
Frequency of acute headache pain medication intakes	-10.3	-10.4	0.1	0.795	-9.9	-8.4	-1.5	0.132
Frequency of acute headache pain medication days	-5.7	-5.8	0.1	0.996	-6.4	-4.8	-1.6	<0.001
Total HIT-6 scores	-4.7	-2.4	-2.3	<0.001	-4.9	-2.4	-2.5	<0.001

Proportion of patients with 50% reduction in headache days	43.5%	36.0%	7.50%	0.082	50.5%	34.4%	0.161	<0.001
Proportion of patients with 50% reduction in headache episodes	46.9%	47.5%	-0.6%	0.905	50.2%	39.1%	11.1%	0.008
MSQ RFR scores	-16.8	-8.8	-8	<0.001	-17.2	-8.4	-8.8	<0.001
MSQ RFP scores	-12.6	-7.6	-5	0.005	-13.5	-5.4	-8.1	<0.001
MSQ EF scores	-16.9	-10	-6.9	0.001	-19	-9.1	-9.9	<0.001

a = primary efficacy variable in PREEMPT 2, b = primary efficacy variable in PREEMPT 1