CASE AUTH/2270/10/09

# ALLERGAN v MERZ PHARMA

## Promotion of Xeomin

**Allergan complained about the promotion of Xeomin (botulinum neurotoxin type A) by Merz Pharma. The claim at issue 'At least as effective as Botox with a similar safety profile' was referenced to Benecke *et al* (2005) and Roggenkamper *et al* (2006) and appeared on an exhibition panel at the Association of British Neurologists meeting in Liverpool in June 2009. Allergan marketed Botox (botulinum neurotoxin).**

**Allergan alleged that the use of the unqualified claim 'At least as effective as' when based on the results from two non-inferiority studies did not accurately reflect the available evidence and was misleading. A non-inferiority trial was only intended to show that the effect of a new treatment was not worse than that of an active control by more than a specified margin. Therefore, it was possible to claim that Xeomin was no worse than Botox by the pre-specified margins in the studies.**

**Allergan agreed it was true that a product that had been shown to be non-inferior to another product might be equivalent to it, or even superior. However, without evidence supporting equivalence or superiority, all that could be said on the basis of a non-inferiority study was that the product was no worse than the comparator by the pre-specified margins.**

**In order to make the claim 'At least as effective as', further evidence that confirmed equivalent efficacy and clinically relevant superiority would be required. A clinician was likely to interpret the claim at issue as meaning this evidence existed, which it did not.**

**The detailed response from Merz is given below.**

**The Panel considered that there was a difference between showing non-inferiority to showing comparability. The Panel considered on the basis of the data the claim that Xeomin was 'At least as effective as Botox' did not reflect the available evidence. It implied possible superiority of Xeomin as alleged and was misleading. Breaches of the Code were ruled.**

**Upon appeal by Merz the Appeal Board noted that both parties agreed that Benecke *et al* and Roggenkamper *et al* were non-inferiority studies that showed that Xeomin was no worse than Botox by a pre-specified margin (delta) that was clinically acceptable.**

**The Appeal Board noted Merz's submission that it had no data upon which to make the claim that Xeomin was equivalent to Botox. In the Appeal Board's view the claim 'At least as effective' not only implied equivalence but also possible superiority which was misleading. The Appeal Board did not consider that the claim could be substantiated by the available data. The Appeal Board upheld the Panel's ruling of breaches of the Code.**

Allergan Ltd complained about the promotion of Xeomin (botulinum neurotoxin type A) by Merz Pharma UK Ltd. The claim at issue 'At least as effective as Botox with a similar safety profile' was referenced to Benecke *et al* (2005) and Roggenkamper *et al* (2006) and appeared on an exhibition panel at the Association of British Neurologists meeting in Liverpool in June 2009.

Allergan marketed Botox (botulinum neurotoxin).

**COMPLAINT**

Allergan alleged that the use of the unqualified claim 'At least as effective as' when based on the results from two non-inferiority studies did not accurately reflect the available evidence and was misleading. A non-inferiority trial was only intended to show that the effect of a new treatment was not worse than that of an active control by more than a specified margin. Therefore, from Roggenkamper *et al* it was possible to claim that Xeomin was no worse than Botox by the pre-specified margin in the Jankovi Rating Scale (JRS) sum score. From Benecke *et al* it was possible to claim that Xeomin was no worse than Botox by the pre-specified margin in the Toronto Western Spasmodic Torticollis Scale (TWSTRS) severity score.

Allergan agreed it was true that a product that had been shown to be non-inferior to another product might be equivalent to it, or even superior. However, without evidence supporting equivalence or superiority, all that could be said on the basis of a non-inferiority study was that the product was no worse than the comparator by the pre-specified margin. The European Medicines Evaluation Agency (EMEA) Guideline on the Choice of the Non-inferiority Margin (EMEA/CPMA/EWP/2158/99) summarised it as:

'The objective of a non-inferiority trial is sometimes stated as being to demonstrate that the test product is not inferior to the comparator. However, only a superiority trial can demonstrate this. In fact a non-inferiority trial aims to demonstrate that the test product is not worse than the comparator by more than a pre-specified, small amount. This amount is knows as the non-inferiority margin, or delta'.

To make the claim 'At least as effective as', further evidence confirming equivalent efficacy and clinically relevant superiority would be required. A clinician was likely to interpret the claim at issue as meaning this evidence existed, which it did not.

Allergan alleged that the claim 'At least as effective as Botox with a similar side effect profile' without appropriate context and qualification was in breach of Clauses 7.2 and 7.3 of the Code. This claim would be interpreted to mean not only equivalence but also possible superior efficacy, and this data was not available.

**RESPONSE**

Merz submitted that the claim 'At least as effective as Botox' complied with the Code, however, Allergan had not previously mentioned an allegation of a breach of Clause 7.3 or that the claim might be misleading. Merz therefore did not believe that the requirement for inter-company dialogue had been fulfilled in these regards. Equally, the part of the claim on the safety profile had not been explored between the companies or raised as an issue by Allergan.

The two studies in question (Benecke *et al* and Roggenkamper *et al*) were used as part of the regulatory submission for the Xeomin marketing authorization and as such the methodology and the 'non-inferiority margin' had been accepted by European and other regulators.

The Xeomin claim 'At least as effective as Botox' was internally approved following thorough research into its appropriateness especially with reference to non-inferiority studies.

Firstly, Merz reviewed previous cases. Only one case in which the Panel commented on the interpretation of a non-inferiority trial was found involving this specific wording. In Case AUTH/1667/12/04: a clinical trial was cited where Cancidas was shown to be non-inferior to AmBisome. The Panel commented twice upon this non-inferiority clinical trial. It was clear that the Panel's views of non-inferiority results was 'at least as effective as'. This was an important factor in approving this claim.

Secondly, a literature search was conducted to ascertain the statisticians' view of the non-inferiority result. An article published as an extension to the CONSORT (Consolidated Standards of Reporting Trials) statement published in JAMA (Piaggio *et al*, 2006) stated 'Non-inferiority trials are intended to show whether a new treatment has at least as much efficacy as the standard or is worse by an amount less than [delta]'. With respect to the delta it stated 'A prestated margin of n is often chosen as the smallest value that would be a clinically important effect'.

Thus, if the difference between the two products, defined by the confidence interval, was less than the 'delta' (or non-inferiority margin) the difference between the products was clinically unimportant. The products could then be described as 'non-inferior' or 'no worse than' each other. This therefore left only two possibilities that the new treatment (in this case Xeomin) was as good as or better than the comparator (Botox). That was to say it was 'at least as good as' Botox.

As it was an established principle of the Code that all claims referred to the clinical situation, to suggest that Xeomin might be inferior to Botox by an amount that was not clinically relevant would be misleading. As a difference less than this 'delta' would be clinically unimportant it could be stated that, clinically, Xeomin had at least as much efficacy as Botox (by adapting the CONSORT statement above).

The EMEA guideline stated that the 'delta' was chosen '…to show that there is no important loss of efficacy if the test product is used instead of the reference'. It was later stated that this was 'supported by evidence of what is considered an unimportant difference in the particular disease area'.

The EMEA guideline further proved that the delta was clinically unimportant reinforcing the message that Xeomin was no worse than Botox leaving only that it may be the same or better – or at least as good as – Botox.

Allergan's allegation that equivalence and possible superiority were not proven by a non-inferiority trial was directly contradicted by Laster and Johnson (2003) who stated that 'The terminology 'at least as good as' or equivalently, non-inferiority, may be interpreted as either literal equivalence or superiority'. This peer reviewed paper in a statistics journal was at odds with Allergan's view. It seemed rational to accept the peer reviewed paper authors' view as they did not have a vested interest in a particular viewpoint, unlike the unreferenced statement from Allergan. Whilst Merz had no interest in promoting 'equivalence' or 'superiority' for Xeomin over Botox without specific evidence of such, this paper clearly demonstrated that Allergan's arguments were fundamentally flawed.

Merz submitted that the published evidence fully supported the claim 'At least as [good] as Botox'. Merz was firmly convinced that directly employing words previously used by the Panel in describing the exact same type of study meant that the claim could not be in breach of Clause 7.2.

**PANEL RULING**

The Panel noted Merz's comments that Allergan had not alleged that the claim was misleading in inter-company dialogue and therefore that aspect of the complaint should not proceed. The Panel noted that in inter-company dialogue Allergan had referred to Clause 7.2 which included a requirement that material should not mislead. Clause 7.3 also

included a requirement that comparisons should not be misleading. Whilst good practice it was not necessary to cite each clause at issue in inter-company dialogue. The substance of the complaint should be clearly identified and discussed and the clauses subsequently cited relevant to that discussion. Therefore the Director decided that in the circumstances the Panel could consider the alleged breach of Clause 7.3.

The Panel noted that both parties agreed that Benecke *et al* and Roggenkamper *et al* were non-inferiority studies. It was also agreed that Xeomin could be described as no worse than Botox. The Panel considered that there was a difference between showing non-inferiority to showing comparability. The Panel considered on the basis of the data the claim that Xeomin was 'At least as effective as Botox' did not reflect the available evidence. It implied possible superiority of Xeomin as alleged and was misleading. A breach of Clauses 7.2 and 7.3 was ruled.

**APPEAL BY MERZ**

Merz submitted that the claim 'At least as effective as Botox' complied with the Code based upon case precedent including wording used directly by the Panel, peer reviewed statistical publications and an EMEA guideline. In its ruling the Panel accepted that Xeomin was no worse than Botox. However the Panel also stated that the two studies could not be used to state that the two products were comparable and that the claim 'At least as effective as Botox' implied possible superiority and that this was misleading. The Panel had not commented in its ruling on Merz's defense of the claim, despite quoting Allergan's allegation in full. This disadvantaged Merz in its appeal.

**Previous case precedent**

Merz submitted that the Panel had previously ruled positively on the description of a 'non- inferiority' trial in Case AUTH/1667/12/04, in which the claim used was 'Cancidas is at least as effective as Ambisome …'. In the clinical trial in question Cancidas was shown to be non-inferior to AmBisome. The complaint was specifically 'Although the bullet point following the claim stated that 'Cancidas was at least as effective as Ambisome …' the reader was left with the distinct impression that Cancidas was better than Ambisome'. This was essentially the identical claim at issue in Case AUTH/2270/10/09. In its ruling the Panel had commented twice upon this non-inferiority clinical trial. The quotations were 'The Panel considered that the claim implied that Cancidas was an alternative first line antifungal therapy to Ambisome and noted the two prominent bullet points set out the reasons why that was so, ie **the two were at least as effective as each other** but Cancidas was significantly better tolerated', the Panel went on to state it '**did not accept that the claim […] implied greater efficacy for Cancidas** …' and 'The claim summarized the data which had already been presented [that the study showed non-inferiority with respect to efficacy but better tolerability] ie that Cancidas **was at least as effective as** Ambisome but better tolerated' (emphasis added by Merz).

Merz noted that the Panel ruled no breach of Clause 7.2 and that it was clear that the Panel's view of non-inferiority results was 'at least as effective as'. This was an important factor in approving this claim.

Merz noted that the Panel in the current case (Case AUTH/2270/10/09) had ruled that the claim at issue had 'implied superiority' and that this was misleading. This contradicted its previous position (as in the first quotation above) where the Panel stated that it 'did not accept that the claim implied greater efficacy …'. Merz was surprised and concerned that the ruling did not include an explanation as to why the Panel had gone against this precedent. It would be a disturbing precedent itself if previous Panel rulings and wording used by the Panel could not be used with assurance by companies to approve material. It would question the rationale of publishing the case reports and go against the principle of natural justice where each company would be treated equitably.

**Appropriate use of statistical terminology**

Merz noted that it had submitted two statistical papers published in peer reviewed journals and further information from the EMEA guideline partially quoted by Allergan. The first paper was an extension to the CONSORT statement published in JAMA (Piaggio *et al*) which stated that 'Non-inferiority trials are intended to show whether a new treatment has at least as much efficacy as the standard or is worse by an amount less than [delta]'. With respect to the delta it stated that 'A prestated margin of noninferiority is often chosen as the smallest value that would be a clinically important effect'.

Thus, Merz submitted that if the difference between the two products, defined by the confidence interval, was less than the 'delta' (or non-inferiority margin) the difference between the products was clinically unimportant. The products could then be described as 'non-inferior' or 'no worse than' each other. This therefore left only two possibilities that the new treatment (in this case Xeomin) was as good as or better than the comparator (Botox) ie it was 'at least as good as' Botox.

Merz submitted that as it was an established principle of the Code that all claims referred to the clinical situation; to suggest that Xeomin might be inferior to Botox by an amount that was not clinically relevant would be misleading. As a difference less that this 'delta' would be clinically unimportant it could be stated that, clinically, Xeomin had at least as much efficacy as Botox (by adapting the CONSORT statement above).
Merz submitted that Laster and Johnson refuted Allergan's allegation, and went against the Panel

ruling, that equivalence and possible superiority were not proven by a non-inferiority trial. This paper stated that 'The terminology "at least as good as" or equivalently, non-inferiority, may be interpreted as either literal equivalence or superiority'. It seemed rational to accept the peer reviewed paper authors' view as they did not have a vested interest in a particular viewpoint, unlike Allergan's unreferenced statement. Whilst Merz had no interest in promoting 'equivalence' or 'superiority' for Xeomin over Botox without specific evidence of such, this paper clearly demonstrated that Allergan's arguments and the Panel's ruling were fundamentally flawed.

Merz submitted that the statement provided by Allergan from the EMEA guideline was a direct 'cut and paste' however it represented a narrow view of that guideline and failed to capture the full context of the document with respect to the clinical situation. The guideline went on to state that the 'delta' was chosen '…to show that there is no important loss of efficacy if the test product is used instead of the reference' and that this was 'supported by evidence of what is considered an unimportant difference in the particular disease area'.

Merz submitted that the EMEA guideline quoted by Allergan further established that the delta was clinically unimportant reinforcing the message that Xeomin was no worse than Botox leaving only that it might be the same or better – or at least as good as – Botox. Merz submitted that the published evidence fully supported the claim 'At least as effective as Botox'. The Panel's ruling of a breach of Clauses 7.2 and 7.3 should be overturned as:

- it directly contradicted a previous ruling for these types of trials thus going against the principles of precedent and natural justice;
- it went against the only two peer reviewed published papers on the subject of 'non-inferiority' trials and;
- contradicted the Code principle that claims referred to the clinical situation by narrowly interpreting a short section of an EMEA guideline out of context and without reference to its clinical relevance.

**RESPONSE FROM ALLERGAN**

Allergan was surprised by Merz's stance in this case and its appeal of the Panel's ruling. Allergan disagreed that Merz had been disadvantaged in any way. The claim at issue 'At least as effective as Botox with a similar side effect profile' was not supported by case precedent, or peer reviewed publications, as suggested by Merz.

Allergan stated that in essence the matter was very simple. There was a single claim at issue which was on a promotional stand at the Association of British Neurologists meeting in June 2009. Allergan understood from inter-company correspondence that this claim was also included in other Xeomin promotional materials. Allergan alleged the use of the unqualified claim 'At least as effective as', based on the results from two non-inferiority studies, did not accurately reflect the available evidence and was misleading. Without appropriate context and qualification readers would interpret the claim to mean equivalence but also possible superior efficacy, and this data was not available. A non-inferiority trial was only intended to show that the effect of a new treatment was not worse than that of an active control by more than a specified margin. Therefore, from Roggenkamper *et al* it was possible to claim that Xeomin was no worse than Botox by the pre-specified margin in the JRS sum score and from Benecke *et al* it was possible to claim that Xeomin was no worse than Botox by the pre-specified margin in the TWSTRS score. It was true that a product that had been shown to be non-inferior to another product might actually be equivalent to it, or even superior. However, without evidence supporting equivalence or superiority, all that could be said on the basis of a non-inferiority study was that the product was not worse than the comparator by the pre-specified margin in the study.

**Case precedent**

Allergan disagreed with Merz's interpretation of Case AUTH/1667/12/04 that the Panel found that use of a non-inferiority study supported the claim that Cancidas was 'at least as effective as' Ambisone. The claim upon which the Panel had ruled was a different one - whether Cancidas could be used as an alternative to Ambisone. The Panel considered that the supporting evidence, of which the non-inferiority study was only a part, did support this claim. The Panel was not asked to rule on, and did not comment on, whether the claim 'at least as effective as' could be made based on a non-inferiority study. Therefore, Allergan strongly disagreed with Merz's conclusion that the Panel had gone against a previous case precedent and was very concerned by Merz's suggestion that it was not being treated equitably.

Allergan alleged that there was a much more relevant case, Case AUTH/2131/6/08, in which the Panel ruled that the claims 'Versatis is comparable to pregabalin in reducing pain intensity at 4 weeks' and 'Statistically shown to be at least comparable in efficacy to pregabalin' were not supported by a non-inferiority study. The Panel clearly stated in its ruling that it 'considered there was a difference between showing non-inferiority to showing comparability'. Whilst case precedent was a helpful guide each case must be ruled on its own merits.

**Appropriate use of statistical terminology**

Whilst there had been significant discussion around statistical terminology, it was important to remember the context of the claim. It was a stand alone, unqualified claim, on an exhibition panel. Allergan alleged that the claim 'At least as effective as Botox with a similar side effect profile' without

appropriate context and qualification would be interpreted to mean equivalence but also possible superior efficacy. Data to support equivalence or superiority was not available. Merz had referenced to two statistical papers and further information from the EMEA guideline provided by Allergan.

Regarding Piaggio *et al*, Allergan disagreed with Merz's 'adaptation' of the section quoted from the CONSORT statement. If products could be described as 'non-inferior' or 'no worse than' each other it did not follow that there were only two possibilities as suggested by Merz – that the new product (Xeomin) was (1) as good as or (2) better than the comparator (Botox). The new product could be worse than the comparator but by less than the delta. Whilst Merz alleged it would be 'misleading' to suggest that Xeomin might be inferior to Botox by an amount less than the delta, it was equally misleading to suggest that Xeomin might be superior (better than) Botox.

With regard to Laster and Johnson, Allergan noted that it was interesting to read beyond the limited quotation provided by Merz. 'The terminology "at least as good as" or equivalently, non-inferiority, may be interpreted as literal equivalence or superiority. Since the statistical demonstration of literal equivalence is fruitless (that is, proving the null hypothesis of no difference), an operational definition must be considered which allows experimental therapy to be inferior to standard therapy by a clinically tolerable amount'. Therefore, this paper seemed to support the view that a non-inferiority trial allowed experimental therapy to be 'inferior to standard therapy by a clinically tolerable amount', not just 'as good as' or 'better'. With reference to the EMEA guideline on the choice of the non-inferiority margin, the additional sections selected by Merz had not supported its case for the use of the claim 'At least as effective as'. The guideline clearly stated that 'The objective of a non-inferiority trial is sometimes stated as being to demonstrate that the test product is not inferior to the comparator. However, only a superiority trial can demonstrate this. In fact a non-inferiority trial aims to demonstrate that the test product is not worse than the comparator by more than a pre-specified, small amount. This amount is known as the non-inferiority margin, or delta'. In order to claim 'At least as effective as', further evidence that confirmed equivalent efficacy and clinically relevant superiority would be required. A clinician was likely to interpret the claim at issue as meaning this evidence existed, which it did not.

Allergan disagreed with Merz's interpretation of case precedent and published evidence. Allergan did not believe it had been selective or narrow in its interpretation of the EMEA guideline presented. Whilst there had been significant discussion around statistical terminology, it was important to remember the context of this claim and how it would be interpreted by the reader.
Therefore, as previously stated, Allergan alleged that the claim 'At least as effective as Botox with a similar side effect profile' without appropriate context and qualification was in breach of Clauses 7.2 and 7.3 of the Code. The claim would be interpreted to mean equivalence but also possible superior efficacy, and this data was not available.

**APPEAL BOARD RULING**

The Appeal Board noted that both parties agreed that Benecke *et al* and Roggenkamper *et al* were non-inferiority studies that showed that Xeomin was no worse than Botox by a pre-specified margin (delta) that was clinically acceptable.

The Appeal Board noted that Merz had cited Case AUTH/1667/12/04 in support of the use of the claim at issue because it submitted that in that case the Panel had ruled that the claim 'Cancidas was at least as effective as AmBisome' was acceptable based upon a non-inferiority trial. However, the Appeal Board noted that this claim had appeared as a bullet point in support of the actual claim at issue which was about the use of Cancidas instead of AmBisome as first line empirical treatment. The Panel had not been called upon to consider the claim '…at least as effective as …' *per se*. In any event each case was judged on its merits. The context in which claims were used was important. The Appeal Board was concerned that Merz had selected a part of the Panel ruling in Case AUTH/1667/12/04 to support its case.

The Appeal Board noted that non-inferiority studies showed that even if one product was worse than another it was only worse within clinically unimportant limits. The Appeal Board noted Merz's submission at the appeal that it had no data upon which to make the claim that Xeomin was equivalent to Botox. In the Appeal Board's view the claim 'At least as effective' not only implied equivalence but also possible superiority which was misleading. The Appeal Board did not consider that the claim could be substantiated by the available data. The Appeal Board upheld the Panel's ruling of breaches of Clauses 7.2 and 7.3. The appeal was unsuccessful.

**Complaint received** **5 October 2009**

**Case completed** **4 January 2010**