

12. The numbering of appeal points in this document reflects those that were used during the hearing. The text of this document does not represent a verbatim account of the proceedings nor a documentation of the order of events that took place but rather, provides a brief summary of the Lilly and committee submissions for the points that were discussed relevant to the decisions of the panel.
13. Before the appeal panel inquired into the detailed complaints the following made a preliminary statement:
 - a. Emily Pegg on behalf of Lilly; and
 - b. Andy Fox on behalf of NICE.

Appeal point 1(a)1: The distinction between health-related quality of life and quality of life experiences by carers of persons with Alzheimer's disease conflicts with definitions in the Manual and is inadequately explained.

14. Adela Williams introduced this appeal point for Lilly, which related to the assessment of quality of life of those caring for people with Alzheimer's disease.
15. At paragraph 3.19 of the Final Draft Guidance ("FDG"), the committee notes that *"many of the aspects that carers of people with Alzheimer's disease reported to affect their daily quality of life may not be directly related to their health"*.
16. Mental health is a domain within EQ-5D, and although quality of life and health-related quality of life ("HRQoL") may be used interchangeably by patients and carers, NICE's technology appraisal and highly specialised technologies guidance: the manual ("the Manual") provides that HRQoL can be incorporated into quality adjusted life year ("QALY") calculations. Adela Williams suggested

that the committee appears to have relied on that distinction to discount clinical expert input. She described the distinction as being unclear, lacking in transparency and in conflict with the glossary published on NICE's website which defines HRQoL as "*a combination of a person's physical, mental and social well-being; not merely the absence of disease*".

17. Summarising, Adela Williams described the distinction as artificial and inadequately explained and that it appears to have impacted the committee's ability to take other data presented by Lilly and clinical experts into account. Lilly is not aware of this distinction (between HRQoL and quality of life) being drawn in any other NICE appraisal insofar as it relates to caregiver utilities.
18. Will Sullivan, for NICE, described in detail the hierarchy of measures set out at section 4.3.10 of the Manual. First, EQ-5D is preferred to measure HRQoL in adults, but there are circumstances where EQ-5D may not be the most appropriate measure. In those circumstances, the Manual provides for a hierarchy of evidence (section 4.3.12 of the Manual). That hierarchy explains that when appropriate, vignettes should be developed using the Decision Support Unit's ("DSU's") best practice recommendations and evaluated from a sample of the general population using an appropriate preference elicitation technique. The two additional documents of assistance in developing vignettes are the 2020 DSU Report and Technical Support Document 11. Will Sullivan noted that neither of these documents had been referred to in Lilly's evidence in responding to either draft guidance consultation.
19. The panel asked Lilly whether the crux of their appeal point is that the committee was not aware of the broader impacts on carers, or that this broader impact was not taken into account. Adela Williams responded that the FDG was unclear exactly what the committee had

Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

taken into account. Further, Adela Williams noted that the distinction itself is contentious – as all of the aspects related to quality of life are likely to impact overall health – consistent with the definition of HRQoL in both the manual and the World Health Organisation¹.

20. Adela Williams also disagreed with Will Sullivan's emphasis on the 2020 DSU Report and Technical Support Document 11 in developing the vignette on which Lilly sought to rely, as these were merely informative rather than a part of the standard methods.
21. Lilly's view was that some aspects of the committee's assessment of HRQoL in the FDG were unfair, as the distinction drawn between HRQoL and quality of life is inadequately explained and therefore it remains unclear which elements (if any) have been excluded from the committee's consideration.
22. The panel asked whether it made a difference that the definition of HRQoL in NICE's glossary appeared to refer to patients rather than carers. Adela Williams noted that there is no material difference in the measure of utility values in carers vs patients. Ben Rothwell, for Lilly, noted that he interprets the definition as relating to patients and carers.
23. The panel highlighted that Lilly's base case submission was in line with NICE's reference case and, at that stage at least, Lilly appeared comfortable to limit the measure of benefit to HRQoL.
24. Adela Williams disagreed with the premise of the panel's question, noting that the glossary does not distinguish health to the exclusion of social elements of HRQoL. The base case submission measuring HRQoL should not be viewed as Lilly seeking to limit the measured

¹ WHOQOL User Manual (WHO/HIS Rev 2012.03)
Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

benefit strictly to those directly related to health as this was not Lilly's interpretation of HRQoL.

25. Ben Rothwell noted that the distinction only became apparent once it was highlighted by the External Assessment Group (“EAG”) in response to the consultation on the second draft guidance. He explained that each reference to health effects made by Lilly prior to this distinction being drawn related to the glossary definition of HRQoL (albeit without specific reference to that definition).
26. Jacoline Bouvy, for NICE, explained that the glossary defines health in line with the common international definition. She described HRQoL as being a subset of the broader definition of health. The reference case itself is clear that it is HRQoL that is relevant to patients and carers. The section of the FDG which considers the impact of donanemab on carer utility does include impacts beyond HRQoL, for example social functioning, ability to do other things, wellbeing, time spent caring for a person with Alzheimer's disease etc. All manner of social aspects are relevant, but the reference case requires HRQoL to be measured, using EQ-5D (or equivalent in some circumstances) to quantitatively account for patient/carer utility.
27. Jacoline Bouvy added that there is provision in the Manual for consideration of wider societal impact where NICE is so directed by the Secretary of State for Health and Social Care. No such direction was given in the appraisal of donanemab. As a result, this gave the committee clear boundaries as to what it could take into account in its decision making.
28. The panel highlighted section 4.3.4 of the Manual which provides that the *“valuation of HRQoL measured by patients (or their carers) should be based on a valuation of public preferences from a representative sample of the UK population using a choice-based method. This*

valuation leads to the calculation of utility values" (emphasis added).

The panel described this as giving the committee a mandate to consider issues broadly considered relevant by patients and carers. The panel asked why the committee, in this appraisal, considered that it did not have the remit to take a broader interpretation of health as a result.

29. Jacoline Bouvy explained that section 4.3.4 of the Manual notes that where there is an observed health status (using EQ-5D, a disease specific tool, other generic measure or a vignette) – that specific method should be translated into a health utility value. There is a standard value set and algorithm already in place for EQ-5D that allows translation of the measure into a utility value. For a vignette study the measured health benefits still need to be translated into a utility value. In the vignette in question here, Lilly did so using a time trade off approach.
30. Jacoline Bouvy explained that the Manual sets out the expectations in relation to the acceptable evidence and what evidence the committee will want to see to justify departing from the EQ-5D. She noted also that the donanemab appraisal was rare insofar as in most appraisals, carer EQ-5D measures are not available and therefore if Lilly does not believe it appropriate to follow that evidence, it should be prepared to explain that view. Will Sullivan explained that one paragraph justifying departing from EQ-5D was not sufficient, and the paragraph provided did not refer to either the 2020 DSU Report or Technical Support Document 11.
31. Ben Rothwell highlighted that this appeal point related to the aspects of HRQoL that the committee had (or had not) taken into account in its decision making and not the rationality or otherwise of using EQ-5D as a health measure. Ross Dent, for NICE, responded on behalf of the committee, explaining that the distinction drawn in the FDG

Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

between HRQoL and quality of life was important as it formed part of the committee's consideration of whether or not to accept a vignette study. Ross Dent highlighted some quality-of-life impacts for carers that would not necessarily be considered HRQoL, for example time to themselves, work related impacts, irritation and frustration and having to do tasks that they had not previously needed to do (such as driving). Some of those aspects considered by people responding to the vignette included such non-health related quality of life features which was an important consideration in the committee's decision that it could not prefer the vignette study.

32. Adela Williams said that the discussion had itself revealed that there was a lack of clarity as to whether or not the committee distinguish between HRQoL and quality of life. However, if there is a distinction drawn, this should be explained in the FDG.
33. Will Sullivan noted that the committee had not sought to exclude any specific aspect of the EQ-5D. Instead, the focus on mental health was because these are the aspects that carers described had most impact on their lives. That does not render the mental health domain of the EQ-5D the only important thing. He recognised that for certain conditions some domains are inherently affected more than others – which is why domains are not excluded.
34. Jacoline Bouvy noted that the glossary to which Lilly refers is a generic glossary on NICE's website, and not a provision of the Manual.
35. Adela Williams said that it would be helpful to understand which elements assessed go beyond HRQoL. It is not, in her view, procedurally acceptable to say that the glossary on NICE's website is not relevant to the Manual.

36. The appeal panel concluded as follows. The panel agreed that the Manual refers to the health of a person and that this included both a patient and/or a carer. They also noted that the Manual states that HRQoL effects assessed by EQ-5D is the preferred methodology to use by NICE in economic modelling.
37. Whilst the panel noted that the Manual was unclear on the definition of what constitutes “health” and how HRQoL should therefore be defined, the glossary on the NICE website was more explicit in its definition which was broad. The panel considered that, whilst the Manual lacked a clear definition, it was implicit that a similar definition as used in the glossary could be inferred for the Manual.
38. The panel recalled that this appeal point under ground 1(a) was that that the distinction between ‘health related quality of life’ and ‘quality of life’ experiences by carers of persons with Alzheimer's disease was inadequately explained, and that there appeared to be a disparity between the definition used in the glossary and that applied by the committee at paragraph 3.19 of the FDG.
39. Section 4.3.6 of the Manual states that the EQ-5D measurement method is preferred to measure health-related quality of life in adults. That being the case, the panel concluded that it was tolerably clear that the components of ‘health-related quality of life in adults’ in both the Manual and as used by the committee, are the factors measured by EQ-5D.
40. The appeal panel therefore dismissed the appeal on this point.

Appeal point 1(a)2: There was inadequate time for consultation following disclosure of NHS England’s “Infusion Costs Estimates” document.

41. Adela Williams introduced this appeal point for Lilly. She explained that as donanemab is a new therapy, NHS England has no direct

Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

experience of donanemab infusion and therefore based its infusion cost estimates on proxies. Lilly was concerned that those proxies were inappropriate and overstated the costs. The committee noted that the most appropriate cost was "likely closer to NHS England's estimates" but noted the uncertainties as to how the costs had been calculated and concluded that it would consider both NHS England's and Lilly's infusion cost estimates in deciding on a plausible Incremental Cost Effectiveness Ratio ("ICER") range.

42. Adela Williams explained that Lilly considered that it had inadequate opportunity to review and analyse NHS England's Infusion Cost Estimates Document. Where substantial new evidence is disclosed, NICE is obliged to allow consultees opportunity to consider and analyse that evidence. No reason was given by the committee as to why Lilly was not afforded the opportunity to respond to NHS England's Infusion Cost Estimates Document.
43. Ben Rothwell, for Lilly, added that, since the third committee meeting, Lilly has now had the opportunity to commission a company familiar with NHS cost coding to analyse the proxies. That work took 58 hours in total to properly understand the HRG codes.² This demonstrated, he explained, that Lilly could not have reviewed NHS England's Infusion Cost Estimates Document within the four days prior to the third committee meeting.
44. Ross Dent, for NICE, explained that NICE had requested further information as to the basis of NHS England's infusion cost estimates at several points during the appraisal, including during draft guidance consultation. He confirmed that NICE had not withheld any information from Lilly in relation to the infusion cost estimates, and

² A Healthcare Resource Group ("HRG") code is a standardised, five character code used to categorise patient activity derived from NHS patient records.
Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

that the information was shared with Lilly as soon as it was provided to NICE.

45. Ross Dent described for the panel's benefit what NHS England's Infusion Cost Estimates Document contained. He explained that the information contained in the document was not new, and was instead an explanatory note of estimates previously provided, alongside a workbook of HRG codes considered. He noted that although NICE would have preferred receiving the document earlier, it cannot compel NHS England to provide information.
46. He noted that the Manual does not provide an opportunity for the company (or any other stakeholder) to respond to each and every piece of evidence, particularly where it is presented outside the draft guidance consultation periods. Similarly, there is no opportunity for one stakeholder to respond directly to another stakeholder's consultation response.
47. Ross Dent noted also that usually the papers would be shared approximately 5 working days before a committee meeting, and therefore he did not consider the truncated period of 4 days with NHS England's Infusion Cost Estimates Document to have been a significant outlier. Further, he noted that Lilly had not made any request to delay the meeting in any way to allow for further time to consider NHS England's Infusion Cost Estimates Document.
48. Ross Dent agreed that the information in question was extensive. In relation to whether or not the information was consequential, he explained that while the estimates themselves are consequential, NHS England's Infusion Cost Estimates Document was merely supporting evidence to explain how NHS England had reached the estimates that it had already proposed earlier in the appraisal process. It would be a very different situation were NHS England

providing entirely new infusion cost estimates four days before the committee meeting.

49. The panel pointed out that the committee could not reach a conclusion on their preferred infusion cost estimates at the second committee meeting, but in the third committee meeting, consequent upon receiving NHS England's Infusion Cost Estimates Document, the committee considered that it preferred NHS England's data.
50. Ross Dent recalled that this was discussed in great detail at the third committee meeting, and that NHS England's National Clinical Director for Dementia was present and was questioned rigorously by the committee. He recalled that Lilly had some opportunity to raise issues about NHS England's Infusion Cost Estimates Document which informed the committee's questioning of the National Clinical Director.
51. The panel asked whether the change in the initial estimate from c.£500 to £432 was a result of NHS England's Infusion Cost Estimates Document. Ross Dent explained that a different approach was submitted in the first committee meeting which was subsequently updated in the second committee meeting. But a consistent methodology was used from the second committee onwards (albeit updating the HRG codes to account for the new financial year during the course of the appraisal).
52. Ross Dent said that the information was made available to other stakeholders including the EAG, who were invited to comment (they did not make comments). He also added this was not the EAG's area of expertise, in comparison to NHS England's pricing team. The committee had no reason to doubt that the information provided was accurate.

53. Adela Williams described that as irrelevant, and said that the appeal point focusses on whether or not Lilly had sufficient time to scrutinise the data before the final committee meeting. If not, she said, that is procedurally unfair.
54. Adela Williams confirmed Lilly had the opportunity to contribute at the third committee meeting, but only to the limited extent possible given the time that it had to prepare for that meeting.
55. The panel asked whether Lilly had asked for more time to consider the information. Ben Rothwell confirmed that it had not, and did not consider that it was a matter for Lilly to delay the committee meeting – instead, he said, the committee itself should be the arbiter of fairness of opportunity to consider new evidence.
56. The panel recalled Ross Dent's explanation that this was supporting evidence, rather than new data. The panel put to Lilly whether this evidence was sufficiently important to delay the committee meeting given that it would not have changed the outcome (as this was not new evidence).
57. Ben Rothwell disagreed with the characterisation of NHS England's Infusion Cost Estimates Document's being supportive evidence. Lilly did not previously know the derivation of the costs, and it was impossible to properly scrutinise the background to those costs without the information eventually provided.
58. Adela Williams highlighted the significant impact that infusion cost estimates had on the ICER, and that it is therefore of critical importance that where proxy costing is used to estimate that cost, all stakeholders have the opportunity to properly scrutinise those estimates. She added that work to analyse the information after the third committee meeting has revealed many errors.

59. The panel asked whether it was fair for the committee to have accepted NHS England's Infusion Cost Estimates Document at a much later stage in the process, after the usual deadline for papers. In light of that, the panel asked whether the committee considered that Lilly had enough opportunity to look at the data.
60. Ross Dent explained that the committee had to balance the fact that the information was provided late in the day, against the detriment of not sharing that information with stakeholders particularly as everyone had been asking for this document. He noted that the information was shared as soon as it was provided to NICE. Not being asked by Lilly for additional time also fed into the committee's assessment of whether or not to delay the third committee meeting. Ross Dent noted that Lilly's appeal letter was the first time that it was made clear to him that Lilly did not consider it had sufficient time to consider NHS England's Infusion Cost Estimates Document. Ben Rothwell responded that the true scale of the disadvantage caused to Lilly as a result only became apparent on proper scrutiny of NHS England's Infusion Cost Estimates Document, which was only possible after the third committee meeting.
61. Adela Williams added that the fact that Lilly had an opportunity to comment on the estimates in the absence of NHS England's Infusion Cost Estimates Document at second draft guidance was no answer at all, as it was not sufficiently transparent to enable proper scrutiny.
62. Ben Rothwell said that there were several hundreds or thousands of HRG codes in the document and Lilly had sought to test the HRG codes to see whether they were accurate and relevant proxies to the costs of donanemab infusion. Fundamentally, he explained, regardless of whether or not Lilly now considers that information to be correct or incorrect, there was not enough time to test whether they were accurate during the appraisal process. He explained that the

scrutiny now completed has, in Lilly's view, revealed shortcomings, including that some codes are inappropriate and that there is double counting of inflation, but those were factors that Lilly would have benefitted from understanding before the third committee meeting.

63. The panel noted that 5 days was the usual timeframe stipulated by NICE to respond to new data and therefore 4 days was shorter than usual, albeit this would not of itself be unfair.
64. The appeal panel concluded as follows.
65. The panel noted that Lilly had not requested for the third committee meeting to be delayed, despite being able to submit only a preliminary response by email in the allotted time. The panel also noted that subsequent analysis done by Lilly took 58 hours to complete and that Lilly had reported that the analysis detected several errors in the HRG codes used to inform the cost estimate.
66. The panel reasoned that for a decision of this significance, given the high impact the infusion cost had on the final ICER, Lilly should have been given more time and opportunity to scrutinise the data and respond fully, rather than the limited time offered with the consequential partial response provided.
67. Whilst the panel acknowledged that rescheduling the meeting would have been logistically difficult, they were of the opinion that the administrative difficulty did not outweigh fairness to allow Lilly time to respond fully to these crucial data.
68. The panel also did not accept NICE's argument that the infusion cost data provided by NHS England were merely explanatory rather than influential data; the infusion cost data were highly influential, as the panel noted at the second committee meeting the committee was

undecided until they received these data. Receiving the NHS infusion cost data therefore influenced the committee's view decisively.

69. The panel therefore concluded that while they recognise the pressure to proceed efficiently, it was unfair not to give Lilly adequate time to respond to such influential data.

70. Therefore the appeal panel upheld the appeal on this point.

Appeal point 1(a)3: The committee's reasons for rejecting the infusion costs estimates by Lilly are inadequately explained and the committee does not appear to have taken into account the evidence submitted by Lilly in response to the second Draft Guidance.

71. Adela Williams introduced this appeal point for Lilly. Reflecting on the previous appeal point, she noted that it had been established that the infusion cost estimates preferred by the committee exerted a substantial influence on the ICER. Lilly had submitted an infusion cost estimate based on simple parenteral chemotherapy, of £207.39. NHS England's infusion cost estimate was c.£500.

72. Adela Williams suggested that Lilly's figure was rejected by the committee based on NHS England's advice against using simple parenteral chemotherapy as a proxy. This was for three reasons: that donanemab was more complex to prepare, had greater potential for adverse reactions, and that the patients receiving donanemab had greater needs than those receiving chemotherapy.

73. That conclusion, Adela Williams said, was not supported by evidence. First, the incidence of infusion reactions is not greater in donanemab (a comparison undertaken between trastuzumab and donanemab showed similar adverse reaction rates). Secondly, patients with mild Alzheimer's disease are physically well, and there is no basis for concluding that they have more complex needs than those with

malignancies. Finally, the suggestion that multiple sclerosis is an appropriate comparator simply because it is another neurological disease is not valid. In summary, she explained, the FDG has not adequately explained why the costs proposed by Lilly should be rejected. Instead, the FDG appears to accept NHS England's costs purely by virtue of them being NHS England's proposals despite the fact that neither Lilly, EAG or NICE itself had sufficient time to scrutinise them.

74. Jacoline Bouvy, for NICE, agreed with Adela Williams that cost of infusion is a key issue in the appraisal of donanemab. The committee was faced with uncertainty – this is a new class of treatment not currently administered in the NHS. It is delivered by way of IV infusion every four weeks. Jacoline Bouvy walked the panel through the various estimates considered before the committee concluded. Lilly had based its costs on simple parenteral chemotherapy infusion throughout the appraisal. The draft guidance makes clear Lilly's view on NHS England's estimate. The second draft guidance concludes that neither cost estimate had been sufficiently explained and that the committee needed a further breakdown of costs by both NHS England and Lilly.
75. At the second committee meeting, the committee considered the updated submissions by NHS England and Lilly, which required a second period of consultation.
76. Jacoline Bouvy explained that ultimately, the reference case requires that costs relate to real NHS/Personal Social Services (“PSS”) costs. As a result, when the national commissioner states what it will pay commissioners for the cost of donanemab infusion, it is difficult to discard that evidence on the basis of lack of face validity. It is well documented in the FDG that there were shortcomings in the information available, which is why the committee had asked for more

Donanemab for treating mild cognitive impairment or mild dementia caused by Alzheimer's disease [ID6222]

information. Jacqueline Bouvy did not accept the suggestion that NHS England's estimate was unsupported by clinical evidence, as the committee had the opportunity to hear from NHS England's National Clinical Director who is a neurologist with real world experience of treating those with Alzheimer's disease. The National Clinical Director supported the cost estimate derived from NHS England's pricing team.

77. Ross Dent, for NICE, confirmed that a large part of the discussion in the third committee meeting focused on the committee's requests to NHS England to support their assumptions.
78. Nick Fox, for Lilly, added that it is reasonable for NICE to be concerned about the uncertainties in evidence, but that it was notable at the third committee meeting that the discussion still centred around the characteristics of a patient who would be receiving donanemab – a factor which is key to the appraisal. He explained that in his clinical experience of administering donanemab, there is no additional complexity in preparing it or in the needs of the patient receiving it.
79. The appeal panel concluded as follows. The panel acknowledged that unlike Lilly's infusion cost estimate, the NHS England infusion costs were not itemised and therefore could not be fully validated by the EAG, committee or Lilly.
80. Whilst it may be argued that the NHS England data were possibly inaccurate and did not reflect "real-world", the panel were reminded that NHS England was the main payor and that the therapy was not available on the NHS and therefore any infusion cost calculation is an estimate and liable to inaccuracy. Furthermore, the Manual stipulates that, when available, NHS prices should be used.

81. The panel reminded itself that this ground 1 appeal point was concerned with the adequacy of reasons provided by the committee, and not the reasonableness of the committee's conclusion. The panel considered that the reasons provided by the committee in the FDG for declining to follow Lilly's estimate of infusion costs, and adopting the NHS England estimate, were sufficient for the reader to understand the drivers of the committee's conclusion. The committee adequately explained how it had reached its decision, on the basis of the evidence available to it.
82. The panel concluded that the process whereby the committee had reached its decision was not unfair, but suggested NICE may wish to modify the FDG to emphasise the uncertainty inherent in reaching their decision.
83. The appeal panel therefore dismissed the appeal on this point.

Appeal point 1(a)4: The committee has relied on the EAG's criticisms of the long-term data for donanemab, even though Lilly was given no opportunity to respond to or clarify these issues.

84. Nuha Brookfield introduced this appeal point for Lilly. She explained that Lilly's position is that when foundational evidence is not fully examined, the conclusions reached are not secure. Lilly had submitted long term extension data before the third committee meeting. The EAG had reviewed these data and had two areas of concern which were relied on by the committee. Nuha Brookfield described that Lilly had been explicitly prohibited from responding to or clarifying those areas of concern. Long term data are crucial in this appraisal as they show the disease modifying potential of donanemab. The impact on the ICER is large, and therefore a considered discussion would have resulted in more evidenced based conclusions.

85. Ross Dent, for NICE, responded that there is no opportunity for stakeholders to comment on other stakeholders' evidence at that late stage in the appraisal process. There was however an opportunity to make factual accuracy submissions and discussion at the committee meeting. Secondly, he noted that paragraph 3.11 of the FDG concludes that the results were uncertain, which is a factual statement rather than indicating an absolute reliance on the EAG's concern.
86. Nuha Brookfield added that failing to consider the long-term data undermined the efficacy of donanemab. She stated that Lilly were informed that they could not discuss this matter at the third appraisal committee meeting. Had Lilly been given the opportunity to do so, she explained, Lilly would have explained all the mitigations it put in place to compensate for the lack of a randomised control trial, and the analysis of those stopping treatment at 18 months. The body of those data speaks to the clinical meaningfulness of donanemab.
87. Ross Dent noted that if the committee did have outstanding queries about the long-term data by the third appraisal committee, they would have been put to Lilly at the meeting. Ross Dent could not recall Lilly requesting to submit additional analysis. Nuha Brookfield confirmed that Lilly did not ask to submit additional scenario analysis and instead considered that the committee would itself have been minded to ask for the data given its clinical significance.
88. The panel highlighted that the FDG considers this matter in detail. Nuha Brookfield said that although the data itself might be considered, the context around that data was not considered. She described the difference between Lilly's modelling approach, where long-term effect is based on amyloid re-accumulation, whereas the application of the EAG's position relates to treatment waning effect. The long-term data were highly relevant to those positions and could have materially changed the outcome.

89. The panel asked whether the committee considered it fair that Lilly did not have an opportunity to comment on this at the third committee meeting. Ross Dent explained that primarily, Lilly is there to answer the committee's questions but there is an opportunity for Lilly to raise a hand and ask a question.
90. The panel noted that there appeared to be an impasse – on the one hand, the committee stated that Lilly could have asked a question should they want to do so, whereas Lilly say that they were explicitly prevented from doing so.
91. Ross Dent explained that he thought that Lilly had probably assumed that the committee would raise this matter for discussion themselves, hence had not asked for the opportunity to submit additional scenario analyses. He said that it was at the committee chair's discretion to invite Lilly to comment, and as far as he (Ross) was aware, Lilly had not asked to speak. Nuha Brookfield disagreed, saying that Lilly were explicitly prohibited from addressing long term data, and that she personally had raised her hand to ask the question numerous times during the meeting.
92. Jacoline Bouvy, for NICE, noted that there is always the opportunity to make factual accuracy submissions. In this instance, the committee did not feel that it required further discussion on the point.
93. Adela Williams, for Lilly, said that it is reasonable to expect that Lilly would be invited to comment/respond in the committee meeting to those criticisms made by the EAG as to Lilly's long-term data.
94. Jacoline Bouvy, for NICE, explained that it is difficult to have a discussion about confidential data in the public part of the committee meeting, which may explain the chair's reticence to discuss the data in that open forum.

95. The panel asked whether the committee had seen the data referred to by Nuha Brookfield. Ross Dent confirmed that the committee had considered those data, and that it is in the committee meeting slides (albeit redacted in the publicly available version of those slides).
96. The appeal panel concluded as follows. The panel acknowledged that the EAG had been critical of Lilly's long-term data between the second and third committee meetings, but that Lilly had not been able to respond to these criticisms at the committee meeting.
97. The panel noted that the long-term data were materially important and had a potentially significant impact on the calculated ICER.
98. The panel were reminded that the Manual states that stakeholders should be invited to contribute and highlight factual inaccuracies and it was the opinion of the panel that Lilly had been prevented from contributing to discussions.
99. The panel considered that the chair of the committee was duty bound to invite stakeholders to comment on matters of factual inaccuracy and in failing to do so at the third committee meeting, this rendered the process unfair.
100. The appeal panel therefore upheld the appeal on this point.

Appeal point 1(a)7: The committee has failed to consider scenario analyses which include unpaid care costs.

101. Ben Rothwell introduced this appeal point for Lilly. He described Lilly's position that the committee's approach to unpaid care costs was inconsistent with the Manual and failed to reflect the true burden of care for those with Alzheimer's disease. Half of the total cost of caring for those with dementia is unpaid care cost and is estimated to

amount to nearly £21 billion each year in the UK, with 1/3 carers reporting that they spend 100 hours a week caring for their loved one.

102. Ben Rothwell referred to a report by NICE's HTA Lab on unpaid care costs noting that these should be analysed as a non-reference case. This aligns with section 4.2.4 of the Manual, that it may be appropriate to consider the costs of care that would otherwise be provided by the NHS/PSS.
103. Lilly's modelled scenario has a meaningful reduction per QALY. Ben Rothwell explained that the FDG makes no reference to NICE's obligation to consider non-reference cases, no mention of the impact on the ICER nor how this impacted the committee's decision making, if at all.
104. Ross Dent, for NICE explained that the initial Lilly submission was based on Wittenberg et al³ which included an estimate of the unpaid care costs in the UK. In response to the second draft guidance Lilly updated its base case to match the committee's preferences and also included the scenario with unpaid care costs. He explained that it is correct that Lilly had asked the committee to consider non-reference case analysis per section 4.4.24 of the Manual. However, the committee's interpretation of that section of the Manual is not that it is intended to capture all informal care, but rather it is intended to capture only that care that the NHS/PSS would normally provide but is instead being provided informally. In the first and second draft guidance, the committee accepted that this was outside the reference case and there was no additional information on the matter at that stage. In hindsight, Ross Dent explained, that could be better

³ Wittenberg R, Knapp M, Hu B et al. (2019) The costs of dementia in England. *Int J Geriatr Psychiatry*, 34:1095

described in the FDG. At the third committee meeting, the base case matched the committee's preference on care costs, but he recognised that the committee may have wrongly considered Lilly's position to be more aligned to the committee's position than was actually the case in practice.

105. The panel noted and agreed that there appeared to have been alignment on this matter at the third committee meeting and that it is only subsequently that the issue arose of whether that approach was appropriate or not.
106. Ben Rothwell explained that Lilly agrees with the reference case element, but that where relevant, those other costs falling outside the reference case ought to have been considered using non-reference case analysis. Lilly had aligned with the reference case on the understanding that the non-reference case unpaid care costs would also be considered given their importance, which does not appear to have happened.
107. The panel asked the committee whether it was precluded from considering unpaid care costs. Ross Dent confirmed that there are circumstances where the committee would consider those costs as non-reference case analysis, but specifically only care that would otherwise be provided by the NHS/PSS and not all informal care.
108. The panel asked whether the committee had received any submissions from carers as to the very substantial burden of costs associated with caring for a family member with Alzheimer's disease that would not usually be reimbursed by the NHS/PSS.
109. Ross Dent confirmed that the committee had received a substantial amount of input about the care burden throughout the appraisal. In response, the panel asked why the committee did not feel as though it

could account for that under section 4.4.24 of the Manual. Ross Dent responded that it is to be taken into account where the cost would otherwise be borne by the NHS/PSS which does not, in the committee's view, amount to the large amount of informal care that is otherwise being met. The panel asked, as an example, that if PSS should be providing care but was not, and that this was borne by the family, whether or not that would be taken into account. Jacqueline Bouvy, for NICE, confirmed that it would.

110. Jacqueline Bouvy added that non-NHS non-PSS costs can be taken into account if NICE is specifically directed by the Department of Health and Social Care ("DHSC") to do so, but in this case NICE had not received such a direction.
111. Ben Rothwell noted that the appeal point related to the committee's failure to take into consideration the non-reference case analysis. Secondly the distinction between informal unpaid care and unpaid care otherwise provided by the NHS/PSS was not clear to Lilly as the two concepts appear difficult to untangle.
112. Jacqueline Bouvy added that this was considered as part of the uncaptured aspects of the appraisal.
113. The appeal panel concluded as follows. The panel were reminded that the Manual states that committees may consider unpaid care costs when these costs would have otherwise been funded by the NHS/PSS.
114. The panel discussed whether the committee was restricted by the absence of DHSC direction to consider broader unpaid care costs (i.e. those that would not otherwise have been funded by the NHS/PSS). They concluded that in these circumstances the Manual allows for non-reference case scenario analysis.

115. The panel agreed that committees have the discretion to use different case scenarios in assessing costs and should use non-reference case scenario analysis when considering costs of providing care, when base case and reference case scenario analysis may be inadequate.
116. The panel reflected that including a non-reference case scenario analysis was important to consider in the economic model and would have had a significant impact on the ICER.
117. The panel considered whether the EAG and committee had adequately considered unpaid care costs in a non-reference case scenario. The panel was of the opinion that they had not seen enough evidence provided in either the FDG or at the hearing that convinced them that the EAG and committee had considered including a non-reference case scenario analysis, irrespective if they had subsequently chosen not to incorporate it into the economic model. They viewed this omission to be a process failure and procedurally unfair, and as a result upheld the appeal on this point.

Appeal point 2.1 - The committee's conclusions on caregiver utilities do not reflect the evidence available and are therefore unreasonable.

118. Nick Fox introduced this appeal point for Lilly. He expressed the view that the committee's conclusions on caregiver utilities do not reflect the evidence and are unreasonable. They are at odds with both clinical experience and the lived experience of patients and those who care for them. That experience shows that there is substantial, progressive and devastating impact of Alzheimer's disease, which impacts mental and physical health, affecting sleep and daily functioning.
119. He said that such is the disparity between the reported lived experience and the carer utility values preferred by the committee,

that they lack face validity and are therefore unreasonable. Crucially, the utility values used materially underestimate the progression of the disease, and the resultant loss in caregiver utility as the disease progresses through mild to severe dementia.

120. He described the difference in a person with a clinical dementia rating ("CDR") of 1 vs CDR3. A person at CDR1 would have some impairment to daily living. Memory loss for recent events, but daily function often intact, able to watch grandkids, drive and independently attend infusion appointments. A person at CDR3 would have very significant functional impairment, severe memory loss, would not be able to hold meaningful conversations or attend to the basic self-care. Incontinence, sleep disturbance, distressing delusions and paranoia are also features in some patients, whereas others may be rendered entirely mute and bed bound.
121. The impact is profound, on patient and carer, as the disease progresses. The utility values preferred by the committee modelled caregiver utility of 0.80 in mild Alzheimer's disease, and 0.76 in severe Alzheimer's disease. A decrement of 0.04 in carer utility is not, in Lilly's view, plausible or in line with any lived experience.
122. Nick Fox explained that part of the reason that the carer utility value is not reflective of real-world practice is because EQ-5D was used to measure carer utility. The domains measured by EQ-5D necessarily limit the ability to measure the full decrement on carer utility, and even those that can capture the utility values, such as the Anxiety/Depression domain, are subject to ceiling effects.
123. The vignette study developed by Lilly showed a consistent drop in carer utility from mild to severe Alzheimer's disease. The EAG's scenario based on the Lilly vignettes showed a similar reduction. In both the Lilly and EAG scenario modelling there was an approx. 30%

decrement. As a result, Nick Fox stated, it is incomprehensible why the committee chose to prefer carer utility values that showed only a 0.04% decrement between mild and severe Alzheimer's disease based on the evidence before it.

124. Will Sullivan, for NICE, acknowledged, the profound impact that Alzheimer's disease has on patients and their carers. The committee was, however, bound to appraise cost effectiveness within the boundaries of NICE's methods. There is provision to depart from EQ-5D in certain circumstances, but there are steps that Lilly has to take to demonstrate that EQ-5D is not the appropriate tool. Those steps were not met here, he said.
125. He explained that Lilly were given a choice between the EAG's preferred EQ-5D sources derived from GERAS⁴, other published EQ-5D data sources giving similar results or Lilly's vignettes. Nick Fox added that the Dementia Research Centre also submitted independent vignettes which supported the consistent drop in carer utility demonstrated in Lilly's vignette.
126. Jacoline Bouvy, for NICE, highlighted that the carer utility values (between 0.76 and 0.68) in the EAG's base case were akin to those seen in metastatic cancer undergoing immunotherapy. She explained that the committee considered the progress from mild to severe disease in terms of utility values (for both patient and carer) in the broader context of the appraisal.
127. The panel asked whether the committee accepted that the preferred sources for carer utility were those with a utility decrement of 0.04% from mild to severe Alzheimer's disease. That did not appear to be in

⁴ Reed C, Barrett A, Lebec J et al. How useful is the EQ-5D in assessing the impact of caring for people with Alzheimer's disease? *Health & Quality of Life Outcomes* (2017) 15:16

dispute by the committee – instead, both Jacoline Bouvy and Will Sullivan explained that EQ-5D is a necessarily blunt tool, and that the committee recognised that the impact on carers went beyond those aspects of health impacted.

128. The panel asked the committee how it sought to balance the fact that the GERAS utility values showed the 0.04% decrement, in light of the evidence from Lilly, academic and clinical experts, patient groups and the EAG's scenario modelling based on Lilly vignette which all suggested that 0.04% was an unreasonable decrement.
129. Jacoline Bouvy said that the committee considers the quality of the evidence, and then uses the judgement of the committee to consider whether it lacks face validity. She recognised that the GERAS study itself nuanced in how it considers EQ-5D to measure carer utility. The committee's interpretation of that conclusion is that it acknowledges that the impact goes beyond physical health, and EQ-5D is a tool primarily concerned with measuring physical health.
130. To mitigate, Jacoline Bouvy explained that the committee multiplied the carers impacted by 1.8. She added that the alternatives put forward by Lilly also had limitations.
131. Will Sullivan noted that the committee is perfectly open to alternatives if there is good reason to use the alternative and that it stands up to scrutiny.
132. The panel asked whether the same level of scrutiny was applied to utility values measured by EQ-5D. If the EQ-5D values lacked face validity, would the committee have applied the same considerations as it did for the vignette. In short, what evidence would the committee have needed to enable it to prefer Lilly's evidence.

133. First, Will Sullivan explained, Lilly would have had to follow the Manual in terms of departing from EQ-5D, and properly justify that departure. Demonstrable alignment to the 2020 DSU Report and Technical Support Document 11 would have strengthened the Lilly's evidence base as it would have allowed Lilly, and then the committee, to reflect on the strength of the vignette.
134. Second, even though the committee was acutely aware of the broader evidence (particularly lived experience), there was a limited subset of data available to it that it could actually use to measure the utility values.
135. Third, the EQ-5D is already validated and scrutinised, and so scrutiny of the tool itself is not necessary.
136. Will Sullivan summarised the committee's position, that if the EQ-5D is inappropriate, there is a way in which that departure can be justified and supported by evidence. That was not the case here, he said.
137. Ben Rothwell, for Lilly, disagreed with Will Sullivan's view that Lilly had not adequately justified its departure from EQ-5D. He drew the panel's attention to Lilly's response to the second draft guidance where Lilly explained the content validity, supported by evidence on construct validity and the responsiveness of the carer/patient population. He described the construct validity of the various measures open to the committee in its decision making and explained that, in reference to the very small utility decrement shown in the GERAS utility values for carers between mild and severe Alzheimer's disease this was contrary to clinical evidence and market research. He highlighted a graph in the appeal papers (page 21 of Committee papers published on the NICE website on 19 June 2025 – Figure 7: Caregiver EQ-5D UK Utilities by AD severity from the overall EU GERAS Study) which showed that in mild Alzheimer's disease, 37.8%

reported a utility of 1 (i.e. perfect health), whereas in severe disease 38.6% reported a utility of 1. This effectively reports no change, and even a mild increase, in carer utility between mild and severe disease states which lacked face validity.

138. Ben Rothwell explained that Lilly had sought to comply with the flowchart depicting the hierarchy of evidence (Figure 4.1 in the Manual) throughout the appraisal. He expressed the view that there was little scrutiny of the vignettes in practice as the committee appeared to be stuck on EQ-5D throughout the appraisal. Further, Ben Rothwell explained that there was a great deal of evidence supporting the vignettes which were developed using a time trade off approach, in line with the DSU recommendations in circumstances where EQ-5D were inappropriate.
139. Will Sullivan explained that initial evidence submission is the best opportunity to raise a concern with the suitability of EQ-5D. In his view, the best thing Lilly could have done is to document how it aligned its approach with the supporting material. The committee can suggest a deviation from EQ-5D of its own volition early in the process. Jacqueline Bouvy added that the committee must ensure that the evidence and assumptions are reasonable, and have a responsibility to scrutinise and test the evidence before it suggests any deviation.
140. Will Sullivan also added that there were other, generic instruments, that could have been considered by Lilly before moving straight to a vignette study which was further down in the hierarchy of evidence.
141. Fundamentally, Ben Rothwell explained, the EQ-5D is designed for patients, not carers, and other preferred generic instruments suffer from the same shortcomings. It is Lilly's view that no other generic measures would have sufficiently reflected the impact on carers.

142. Adela Williams, for Lilly, highlighted that the evidence supporting the use of vignettes was not limited to Lilly, and that academic submission (including the Dementia Research Centres) also supported these data.
143. Nick Fox noted that the committee had not described the 0.04% utility decrement as reasonable.
144. Jacoline Bouvy explained that, for the avoidance of doubt, the committee did consider preferring the GERAS utilities to have been a reasonable assumption to make. That is not to say that the evidence was perfect, and the committee acknowledged that it may underestimate the impact on carers which is why the model assumed that 1.8 carers would be impacted. On balance, the committee considered this to have been a reasonable approach while recognising that there were limitations present.
145. Ben Rothwell highlighted that the committee did not appear particularly satisfied that EQ-5D was plausible and that Lilly had asked the committee to consider the middle ground between Company vignette and the GERAS values – i.e. the EAG adjusted company vignettes. That was not included or discussed, said Ben Rothwell.
146. Jacoline Bouvy highlighted that the committee's rationale on this matter is described in detail in the FDG. Ben Rothwell expressed the view that the large section in the FDG was because of the large impact on the ICER. He said that it was doubtful whether or not the face validity of the GERAS values had been properly scrutinised.
147. Will Sullivan explained that multiplying x1.8 was effectively superimposing the primary carer burden on other carers which is a likely overestimation of the burden. Will Sullivan conceded that

multiplying by 1.8 did not address the uncertainties completely, but that the committee sought to be reasonable by not viewing uncertainties in isolation (hence multiplying by 1.8).

148. Jacoline Bouvy added that it is clear in the papers that there was a substantial difference in the utility values between GERAS and the Lilly vignette. The committee felt on balance that the GERAS utilities were a reasonable source. If the consequence of that is underestimating the impact on carers, then multiplying by 1.8 seeks to mitigate the impact of the estimate in a manner that is reasonable but nonetheless conservative.
149. Concluding this appeal point, Adela Williams drew the panel's attention to the significant difference between Lilly's utility values and the GERAS utility values preferred by the committee. Applying a 1.8 multiplier to a disutility figure that is derisory and not representative does not create a representative figure, she said.
150. The appeal panel concluded as follows. The panel agreed with Lilly's position that the EQ-5D tool was inadequate to assess carer utility values in these circumstances and that it was illogical for the committee to continue to use the EQ-5D, despite it being the preferred model in the Manual, when it was aware that the carer utility values calculated were grossly under-estimated and remote from other (vignette) models provided by both the EAG and Lilly and were also completely at odds with expert opinion.
151. The panel understood that the committee knew the important impact Alzheimer's disease had on carers and that this had been the reason to use a quantitative rather than qualitative approach to this issue, notwithstanding the inherent issues of using EQ-5D.

152. Whilst the panel acknowledged the committee's attempt to compensate for the inadequacy of EQ-5D by applying a 1.8 multiplier they were not persuaded that this had provided adequate redress.
153. The panel reasoned that the committee had the discretion to use other methods when EQ-5D is inadequate and to not do so in this case was unreasonable.
154. Therefore the appeal panel upheld the appeal on this point.

Appeal point 2.3: The committee's assessment of the additional aspects of donanemab treatment is unreasonable and arbitrary.

155. Adela Williams introduced this appeal point for Lilly. She drew the panel's attention to paragraph 3.28 of the FDG, which includes "uncaptured harms". She explained that there is evidence to support each instance of uncaptured benefit listed in the FDG. The list of uncaptured harms, on the other hand, is in Lilly's view unbalanced, discriminatory and unreasonable. The committee lists as uncaptured harms "false hope", burden associated with treatment, and increased demand on services. Having listed these, the committee concludes that there is not sufficient evidence on the impact of these uncaptured aspects on the plausible ICER so as to change the conclusions on cost effectiveness.
156. Adela Williams went on to address each of the uncaptured harms in turn. "False hope" is a feature of any therapy, and those who are ineligible will simply not commence treatment (which will be addressed through counselling). Infrastructure costs fall outside NICE's remit. Difficulties with diagnosing Alzheimer's disease is also not an issue specific to donanemab, but is as a result of the current state of dementia care in England. She expressed the view that all of

the purported uncaptured harms listed are unsupported by evidence and are irrelevant.

157. Adela Williams expressed Lilly's view that, as a result, the conclusion reached in relation to uncaptured aspects is unsustainable.
158. Victoria Kelly, for NICE, explained that the uncaptured aspects section of the FDG addressed the comments received from all stakeholders, including patient and professional groups. She explained that the committee received a number of comments which were supportive of the negative draft guidance, highlighting the risks and harms of donanemab, which was very unusual. The committee considered it important to highlight the uncaptured harms as well as the benefits to ensure that stakeholder and public opinion was captured in a balanced manner in the FDG. Section 6.3.2 of the Manual provides that the FDG should address risks/harms – and therefore the committee considered it reasonable to address risk/harm in addition to benefit.
159. Andy Fox, for NICE, explained that the "uncaptured aspects" section of the FDG requires the committee to balance the uncertainties. As was clear in the appeal hearing, this is a complex area and the committee was required to analyse a number of qualitative aspects in addition to quantitative ones. What the committee cannot do is dismiss stakeholder comments out of hand – therefore the committee considered it important, in light of those comments, to reflect the harms in addition to the benefits.
160. Andy Fox further explained that part of considering uncaptured aspects is assessing the opportunity cost of the recommendation. Unless there is clear evidence-based reasoning to move the cost effectiveness threshold, which is often discussed at great length by the committee, the committee will be reticent to move the threshold as

it is required to consider the health of others that would be impacted as a result of a decision error.

161. Ross Dent, for NICE, explained that significant work has been done by NHS England to quantify the cost of the impact on primary and secondary care, to consider what investment is needed in infrastructure, service redesign and training. But the committee was aware that NICE was appraising two very similar drugs (lecanemab and donanemab) in parallel and that therefore it would prejudice either drug should the full cost of necessary service redesign be attributed to one or the other. As a result, the committee considered it prudent to remove those costs quantitatively from both appraisals, though this was not explained in the FDG.
162. Jacoline Bouvy, for NICE, reiterated previous comments that it was important to strike a balance between the uncaptured aspects highlighted by all the stakeholders. Adela Williams disagreed with that position, noting that it is unfair to include uncaptured aspects that have not properly been assessed for their validity. It cannot be assumed that the uncaptured aspects have a neutral effect on the ICER where that impact has not been adequately assessed or considered. She emphasised that Lilly's appeal point relates to what is written in the FDG, and the extent to which it was reasonable for those aspects to have been included as uncaptured harms without proper analysis of the validity of those aspects.
163. The appeal panel concluded as follows. The panel noted that the "uncaptured harms" whilst not often considered were included to reflect stakeholder feedback and therefore it was not unreasonable of the committee to include them to reassure stakeholders that their concerns had been considered.

164. The panel was also reminded that it explicitly states in the Manual that the impact of any treatment on NHS resources should be considered.
165. The panel acknowledged that uncaptured benefits and harms are by their very nature qualitative data and attributing anything more than an estimate on their impact is implicit. Furthermore, the panel noted that irrespective of the weighting of uncaptured benefit to harm, this would not have made a material difference to the recommendation made by the committee.
166. The panel were persuaded that the committee considered these uncaptured outcomes carefully and that it was not unreasonable for them to conclude that their impact may increase or decrease the ICER threshold, particularly when they couldn't be measured, and therefore it was reasonable to not change the ICER boundary.
167. The panel therefore dismissed the appeal on this point.

Conclusion and effect of the appeal panel's decision

168. The appeal panel therefore upheld the appeal on appeal points 1(a)2, 1(a)4, 1(a)7 and 2.1, and dismissed all other appeal points.
169. The appraisal of this technology is remitted to the appraisal committee in order to allow further consideration of caregiver utilities and unpaid care costs, and to ensure that adequate opportunity is afforded to stakeholders to respond to the EAG's critique on long-term data of donanemab and consider NHS England's Infusion Cost Estimates document.
170. Additionally, the panel suggested that in respect of appeal point 1(a)3, NICE may wish to modify the FDG to emphasise the uncertainty inherent in reaching their decision regarding infusion cost estimates.

171. There is no possibility of further appeal against this decision of the appeal panel. However, this decision and NICE's decision to issue the final guidance may be challenged by applying to the High Court for permission to apply for a judicial review. Any such application must be made within three months of NICE publishing the final guidance.