Ian Ingram, Managing Editor, MedPage Today

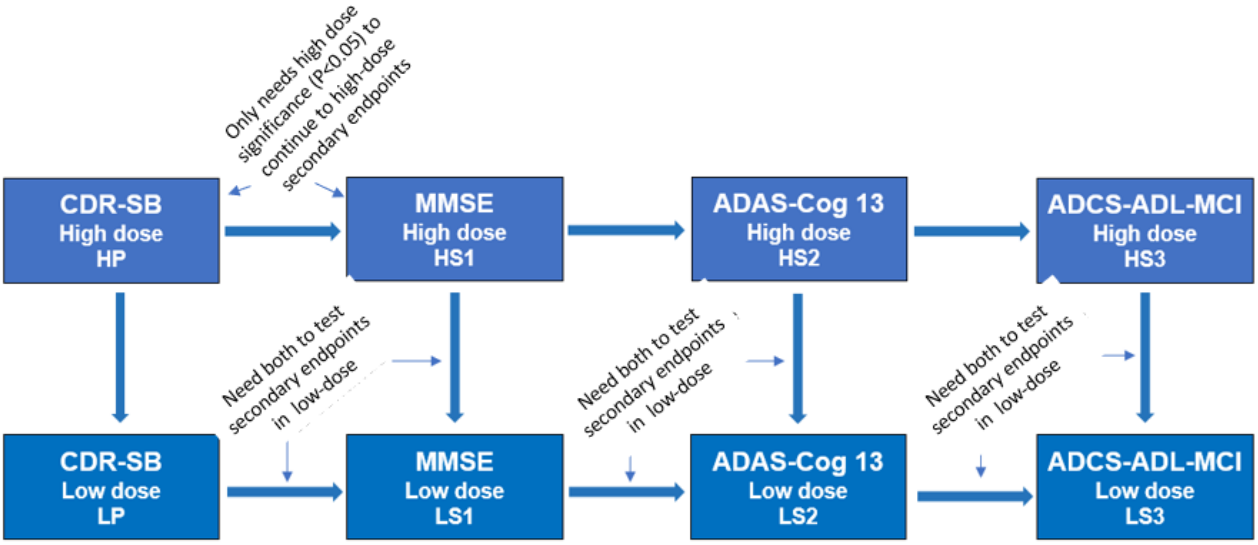
To the editor: As stated in the manuscript published last week in *JPAD*, a sequential testing procedure, prespecified in the study protocols, was used to control type I error rate due to multiple endpoints and multiple comparisons.

According to this sequential testing procedure (Figure 1, below), once the aducanumab high dose for CDR-SB (HP) is significant ($p < 0.05$), testing simultaneously moves to aducanumab high dose MMSE (HS1) and low dose CDR-SB (LP). If high dose MMSE (HS1) is significant ($p < 0.05$), testing continues to aducanumab high dose ADAS-Cog 13 (HS2), and if that is significant ($p < 0.05$), testing continues to aducanumab high dose ADCS-ADL-MCI (HS3). In this procedure, testing for aducanumab high dose of the next ranked endpoint requires significance only on high dose of the previous endpoint and does not require significance on low dose of the previous endpoint (the top row in Figure 1).

Testing for low dose of the next ranked endpoint requires significance on both low dose of the previous endpoint and high dose of this endpoint itself. This is reflected by the bottom row in Figure 1, in which there are 2 arrows pointing to the each of the 3 secondary endpoints. If both low dose CDR-SB (LP) and high dose MMSE (HS1) are significant ($p < 0.05$), testing continues to low dose MMSE (LS1). If both low dose MMSE (LS1) and high dose ADAS-Cog 13 (HS2) are significant ($p < 0.05$), testing continues to low dose ADAS-Cog 13 (LS2). If both low dose ADAS-Cog 13 (LS2) and high dose ADCS-ADL-MCI (HS3) are significant ($p < 0.05$), testing continues to low dose ADCS-ADL-MCI (LS3). The family-wise type I error rate for the primary endpoint is controlled at 0.05. For the 3 secondary endpoints, the family-wise type I error rate is controlled between 0.05 to 0.1. This testing procedure was accepted by the FDA.

The clinical principle underpinning the testing strategy was that high-dose (10 mg/kg) was the target dose. Therefore, failure of low-dose on any endpoint should not preclude testing of the high-dose.

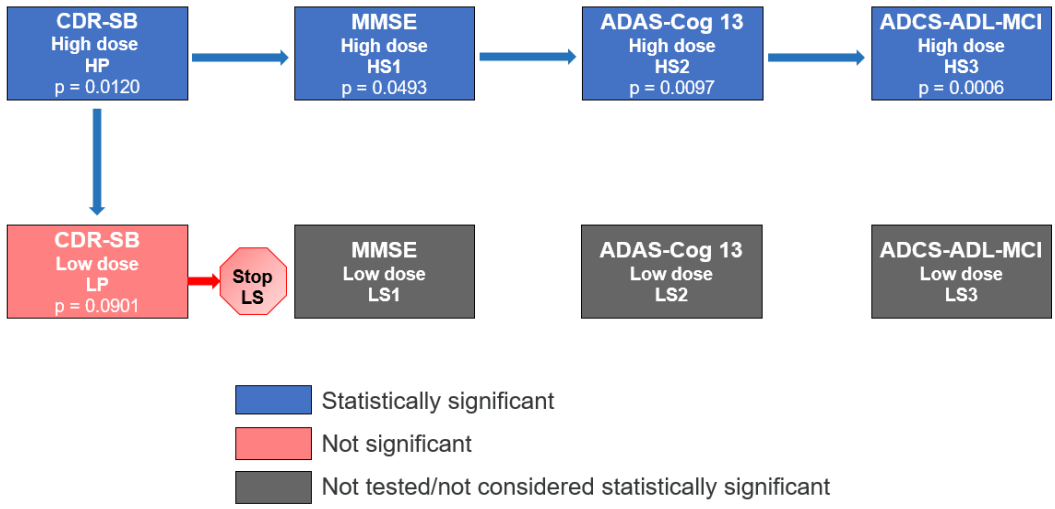
Figure 1



As currently written, the article assumes an endpoint-by-endpoint review, first reviewing the high-dose in CDR-SB (for example) and then the low-dose in CDR-SB, and then followed by secondary endpoints in the high dose. This is simply not the case based on the prespecified SAP.

In EMERGE (Study 302), as shown in Figure 2 below, the testing for high dose continues from HP up to HS3, so the statistical significance is reached for these 4 comparisons according to the pre-specified testing procedure. For low dose, since LP has a p-value > 0.05, the testing stops at LP, and the downstream LS1-LS3 are not tested.

Figure 2: Multiple testing results for Study 302



For ENGAGE (Study 301), as shown in Figure 3 below, since HP is not significant, all downstream testing is not assessed.

Figure 3: Multiple testing results for Study 301

We will continue to provide physicians with efficacy and safety data to help them make the best treatment decisions for patients as we learn from our ongoing trials and real-world evidence.

Samantha Budd Haeberlein, Ph.D., Cambridge, Mass
Head of Neurodegeneration Development Unit, Biogen