CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

211675Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA #:	NDA211675
Drug Name:	Upadacitinib
Indication(s):	Treatment of Rheumatoid Arthritis
Applicant:	AbbVie Inc.
Date(s):	Received: December 18, 2018
	PDUFA: August 18, 2019
Review Priority:	Priority
Biometrics Division:	Division of Biometrics 2
Statistical Reviewer:	William Koh, PhD, Statistical Reviewer
Concurring Reviewers:	Peiling Yang, PhD, Acting Deputy Director and Team Leader
	Mark Rothman, PhD, Director
Medical Division:	Division of Pulmonary, Allergy, and Rheumatology Products
Clinical Team:	Keith Hull, MD, Medical Officer
	Rachel Glaser, MD, Team Leader
	Nikolay Nikolov, MD, Associate Director for Rheumatology
Project Manager:	Nina Ton

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

This review evaluates a new drug application (NDA) by AbbVie for upadacitinib, a Janus Kinase (JAK) inhibitor, as an oral dose of 15 mg once daily (QD), in an extended-release (ER) formulation, with the proposed indication (b) (4) The focus of this review of the efficacy of the proposed dosing regimen of upadacitinib is based on five clinical studies submitted by the applicant. M13-542, M13-549, and M14-465 were randomized, double-blind, placebo-controlled, parallel-group, multi-arm placebo-controlled (cDMARD add-on) studies. M15-555 and M13-545 were randomized, double-blind, active-controlled, parallel-group, multi-arm monotherapy studies. All Phase 3 studies, except M14-465, included upadacitinib 15 mg QD and 30 mg QD. The primary efficacy endpoint in all studies except study M13-545 was American College of Rheumatology 20% improvement from baseline response (ACR20) at Week 12 (Week 14 for study M15-555). The primary efficacy endpoint in study M13-545 was ACR50 at Week 12.

There was statistical evidence of benefit for upadacitinib 15 mg QD with respect to the primary endpoints based on the placebo-controlled and monotherapy studies (3.2.6.1). The estimated probabilities of ACR20 and remaining on randomized treatment at Week 12 were 36.2% (95% CI: 26.2%, 46.2%; p<0.001), 28.1% (95% CI: 19.1%, 37.0%; p<0.001), and 34.1% (95% CI: 29.0%, 39.2%; p<0.001) higher on the absolute scale comparing upadacitinib 15 mg QD to placebo in MTX-add on studies M13-542, M13-549, and M14-465 respectively. In study M15-555, this estimated probability of ACR20 response at Week 14 were 26.5% (95% CI: 17.5%, 35.6%) significantly higher on the absolute scale comparing upadacitinib 15 mg QD monotherapy to MTX monotherapy. The estimated probability of ACR50 response at Week 12 were 23.7% (95% CI: 16.0%, 35.6%) significantly higher on the absolute scale comparing upadacitinib 15 mg QD monotherapy to MTX monotherapy in study M13-545.

Sensitivity analyses were conducted to evaluate the treatment effect on the individual components of ACR20 as well as the impact of missing data on the primary multicomponent ACR20 endpoint at primary efficacy timepoint. Findings remained consistent for these various sensitivity analyses.

Supportive analyses based on all observed data at Week 12/14 (prior to crossover) for these secondary endpoints such as health assessment questionnaire-disability index (HAQ-DI) (3.2.6.2), disease activity score using 28 joint counts based on C-reactive protein or DAS28(CRP) (3.2.6.3), as well as for patient-reported outcomes (3.2.6.4), remained consistent with primary efficacy findings of an effect on the improvement of signs and symptoms of patients and further supported the efficacy of upadacitinib 15 mg QD.

In study M14-465, there was statistical evidence that upadacitinib delayed the rate of radiograph progression compared to placebo (3.2.6.6). In our analyses of study M14-465, the estimated rate of radiographic progression up to Week 26 was 0.78 and 0.15 for placebo and upadacitinib 15 mg QD arm respectively. The estimated difference in the adjusted linear rate of radiographic progression at Week 26 was -0.52 (95% CI: -0.75, -0.28; p-value <0.001) slower in the upadacitinib 15 mg QD arm relative to the placebo arm, indicating a statistically significant treatment effect in mTSS at Week 26. Radiographic results from M13-545 were consistent with the findings in M14-465.

Numerical trends of efficacy were consistent based on subgroup analyses of key demographic and clinical factors.

In this review, we also discussed the concerns with the primary estimand for the radiographic endpoints, limitations of a single study to support comparative claims with adalimumab, limitation of the different study designs in presence of crossover between products in characterizing long term safety, and appropriate pooling of AEs to describe safety findings in the product label.

In summary, there was convincing evidence that upadacitinib 15 mg QD is efficacious based not only on the primary endpoints of ACR20 and HAQ-DI, but also based on components, and patient reported outcomes evaluated at the primary timepoint assessed. There was also convincing evidence of delay in radiograph progression based on M14-465, supported further by study M13-545. Based on the efficacy findings, limited safety analyses conducted in this review, and discussion with clinical, I recommend approval of upadacitinib 15 mg QD as a treatment for adult patients with moderately to severely active RA with limitation of use to RA patients who have had an inadequate response or are intolerant to MTX.

2 INTRODUCTION

2.1 Overview

2.1.1 Drug Class and Indication

Upadacitinib, a small molecule inhibitor of JAK family, is proposed (b) (4) Other approved treatments in this class include tofacitinib and baricitinib. The dosing regimens evaluated in the clinical study is an oral tablet, extended release, 15 mg to be taken once daily. The higher dosing, 30 mg QD, was evaluated in the four of the clinical studies but the applicant did not request for approval of the higher dosing regimen.

Rheumatoid arthritis (RA) is a chronic inflammatory disease that is manifested not only on the joints but also can lead to damage in a wide variety of body systems. This disease affects more than 1.3 million people in the US population and patients with this disease are typically in the 40s, and more likely women. The clinical features of this disease include tender, warm, and swollen joints, and impaired functioning. Patients may experience joint stiffness that is generally worse in the morning and after inactivity. Other clinical signs and symptoms include fatigue, fever, and weight loss. Because this is a chronic disease and inflammation affects joint linings, causing painful swelling that can lead to bone erosion and joint deformity, a key goal of treatment is to delay progression of structural damage to the bones and joints.

Currently, there are several treatment options for RA patients. First-line treatment options include non-steroid anti-inflammatory medications (NSAIDs) and conventional disease-modifying anti-rheumatic-drugs (cDMARD) such as methotrexate (MTX). MTX is the most common first line treatment but has many associated side effects. The more targeted biologic DMARDs (bDMARD) aim to inhibit certain inflammatory pathways. Some of the currently approved biologic therapies include tumor-necrosis factor (TNF) inhibitors (such as Enbrel,

Remicade, Humira, Simponi, and Cimzia), interleukin-1 (IL-1) inhibitors (Kineret), interleukin-6 (IL-6) inhibitors (Actemra and Kevzara), inhibitors of costimulatory pathways of T-cells (Orencia), and anti-CD20 (cluster of differentiation) products (Rituxan). The small molecule JAK inhibitors, namely Xeljanz and Olumiant, are also approved treatment options.

2.1.2 History of Drug Development

The applicant has had several meetings with the Agency over the course of their clinical development initiated under IND114747. They included an End of Phase 2 (EOP2) meeting on October 26, 2015, a Type C meeting on September 6, 2016, Type C written response on April 17, 2017, a pre-NDA meeting on May 1st, 2019. There were additional communications with the applicant on their proposed study protocols and statistical analysis plan (SAP). I included relevant sections important for this review.

The EOP2 package included synopsis from four proposed phase 3 studies, M13-542, M13-549, M14-465, and M15-555. The Agency raised concerns on the proposed study designs that crossed over patients into their investigational product, use of linear extrapolation for radiographic endpoints, use of DAS28(CRP) as a proposed co-secondary endpoint, and the proposed non-inferiority margin with active comparator. The study designs had allowed placebo to crossover to investigational product at early escape. Although study M14-465 included an active comparator to provide longer term safety and efficacy data, the possibility of crossover from adalimumab arm to investigational product in study M14-465 limited the interpretation of results. The Agency did not agree with the use of linear extrapolation as a single imputation approach to impute radiographic scores in patients who escape or withdraw from the study early. Instead, the Agency emphasized that an analysis evaluating the intention-to-treat or de facto estimand based on all Week 24 data in all randomized patients regardless of adherence or use of ancillary therapies would be critical from a regulatory perspective. The Agency specified that DAS28(CRP) may be used as a co-primary endpoint but would expect demonstration of substantial evidence of benefit in the physical function HAQ-DI. The goal of the non-inferiority comparison in Study M14-465 was unclear. The Agency commented that explicit comparative effectiveness claims will generally require positive results from at least two adequate and well-controlled studies designed to compare two products.

The applicant submitted a request for a Type C meeting to discuss the acceptability of existing evidence and validation plans for the Functional Assessment of Chronic Illness Therapy – Fatigue Scale (FACIT-F), (b) (4) and morning stiffness – Numerical Rating Scale (NRS) endpoints to support product labeling. The Agency expressed concerns that it was unclear whether FACIT-F, not originally developed for use in RA patients, can capture fatigue associated with disease activity in RA. It was unclear whether this can translate to clinical benefits distinct from the benefit seen with control of disease activity, already captured by American College of Rheumatology (ACR) response criteria. The Agency did not agree (b) (4) The Agency agreed that duration of morning stiffness is a relevant endpoint although this may overlap with other core outcome measures in RA. The Agency advised the applicant to include the duration of morning stiffness in the statistical hierarchy, and be evaluated in two studies if they planned to include in the label.

After the EOP2 meeting, the applicant submitted their draft statistical analysis plan (SAP) for phase 3 studies M13-542, M13-549, M14-465, and M15-555. The Agency reiterated concerns of crossover between arms in the studies and stated that post-escape results will limit safety interpretations. The Agency recommended the inclusion of two-dimensional tipping point analyses to more comprehensively address the impact of potential departures of the missing-at-random (MAR) assumption for continuous endpoints such as DAS28(CRP) and HAQ-DI. The Agency communicated that tipping point sensitivity analysis be included for the radiograph endpoint based on all data collected without any use of linear extrapolation. Other comments included additional description on how visit windows were derived and descriptive planned analyses for adverse events of special interest.

The applicant submitted a Type C meeting request on April 17, 2017, to discuss the content of their SAP for Integrated Summary of Efficacy (ISE) and Integrated Summary of Safety (ISS), their proposal to define treatment-emergent adverse events, the proposed SAS datasets and SAS programs for the planned NDA based on studies M13-542, M13-549, M14-465, and M15-555. The Agency agreed, via a written correspondence on June 22, 2017, with the applicant's proposal for the datasets for submission. There was concerns on the applicant's proposal to (b) (4) The Agency recommended that integrated safety analyses pooled studies that evaluated similar treatment arms, and that the analysis be restricted to the placebo-controlled or active-controlled period. The objectives of the pooling should be specified, i.e., to compare AEs between upadacitinib vs placebo or MTX. An appropriate summary measure (e.g., risk difference, rate ratio, hazard ratio) should be chosen and confidence intervals (CIs), stratified by study, should be included to quantify the uncertainty in the comparison of the AE rates.

The applicant submitted a revised SAP for their ISE and ISS, SAP for study M14-465, a fifth study, M13-545, in their pooling strategy, and requested for feedback on December 15, 2018. In a written response sent on April 5, 2019, the Agency recommended that an independent adjudication committee be included, if none had been set up, to retrospectively evaluate any potential thrombosis events. The Agency emphasized a list of AEs that should be included in each study's summary report and the ISS report. The Agency provided mock safety tables to the applicant for their consideration as they planned their study report for study M14-46, ISS analysis for doses of upadacitinib during placebo-controlled period (respectively MTX monotherapy) with appropriate censoring when patients were escaped to investigational product. The Agency pointed out that the objectives of the long-term analyses using "Any phase 3 upadacitinib 15 mg/Any Upadacitinib 15 mg and 30 mg/Any RA Upadacitinib" remained unclear and reiterated that the presence of crossover provided limited controlled safety interpretation.

During the pre-NDA meeting on May 1st, 2019, the details of the submission package and comments on the ISS and SAP previously sent to the applicant were discussed. In general, the applicant planned to submit five clinical studies for review and use a Priority Review Voucher for the submission. The Agency reminded the applicant to provide programs and macros for efficacy, disposition, and safety results reported in the CSR, and further noted that the applicant would submit post-hoc tipping point analyses. The Agency stated that study M13-545 was not

previously discussed and the study would be a review issue. The Agency agreed that the proposed definitions by the applicant for cardiovascular and thromboembolic events were reasonable.

During the meeting, the applicant requested clarifications on the Agency's comments on their revised ISS SAP. The Agency stated that the applicant's definition of total exposure, when calculating incidence rates for an AE, should reflect the time at risk for an event rather than total patient time on study drug. The placebo-controlled and methotrexate-controlled safety comparisons would provide appropriate short-term comparison of AE rates with placebo or the MTX active comparator. The Agency stated that the long-term integrated safety analyses would be limited in terms of interpretation because placebo patients were crossover to upadacitinib for all studies and adalimumab patients from study M14-465 were crossover to upadacitinib. The utility of long-term integrated safety analyses, accounting for differences in exposure across studies, would be reasonable to compare AE rates between doses of upadacitinib but would not provide any placebo/active controlled comparison with upadacitinib.

The applicant asked whether the safety database, with different database lock cut-off dates across studies, would be considered acceptable. The Agency raised concerns on the amount of safety data the applicant would have at the time of filing, given the complexity of the crossover for some studies and that there may be limited safety from patients beyond one-year. Specifically, the Agency was concerned on assessment of rare events without adequate long-term controlled comparisons, particularly for events that were not dose dependent or had long latency. The Agency commented that AEs rates up to Week 48/52 follow-up would be of interest for review purposes even though some patients could have been exposed for longer than a year. The applicant stated that data up to Week 48 from study M13-545 will be included in the safety update. In a post meeting comment, the Agency requested that the applicant include Week 48 data from study M13-545 in the submission.

2.2 Specific Studies Reviewed

The applicant had submitted the results of five Phase 3 studies (M13-542, M13-549, M14-465, M15-555, and M13-545) to support the safety and effectiveness of upadacitinib with the proposed dosing of 15 mg once daily (QD) for treatment of adult patients with moderately to severely active rheumatoid arthritis (RA). They are described in the following order for the remainder of this review: M13-542, M13-549, M14-465, followed by monotherapy studies M15-555 and M13-545. Radiographic progression was evaluated in studies M14-465 and M13-545. The submission included two Phase 2 studies (M13-550 and M13-537) using a different formulation and a Phase 2b/3 study (M14-663). The studies are not included in this review.

2.3 Data Sources

Data were submitted by the applicant to the CDER electronic data room in study data tabulation model (SDTM) and analysis data model (ADaM) format. Protocols, correspondence, data listings, program code, and study reports were accessed under the network path \\cdsesub1\evsprod\NDA211675\nda211675.enx.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

In general, the quality of the submitted datasets was acceptable and datasets were carefully documented. The relevant flags used for the primary and secondary efficacy analyses were described sufficiently in the analysis data reviewer's guide (ADRG). I was able to reproduce the primary and key secondary analyses without noticeable deviations. For the integrated summary of safety results, I was able to reproduce the important results.

In general, the study conduct appeared adequate. The five studies were monitored by a common Data Monitoring Committee (DMC). Based on the meeting minutes submitted, the study conduct appeared adequate. The randomization schemes were consistent with the protocol description.

Subsequent information requests (IRs) related to the statistical review were summarized in Table 1.

Table 1 Data links, Dates, and Summary of Information Requests Made to the Applicant.

Information request links	Date of Response	Summary of the response to IR
0012	March 28, 2019	Macros were submitted for tables generated in the ISS documents. Additional disposition tables for studies M13-545 and M14-465 were submitted for review. The applicant submitted only the worse-case scenario for the primary efficacy endpoint and did not conduct a two-dimensional tipping point as requested. [2] Multiple imputation for HAQ-DI, SF36, and DAS28(CRP) at primary week of analysis was submitted. Radiograph disposition and additional analysis was submitted for the mTSS endpoint.
0022	May 20, 2019	Follow-up to the statistical analysis conducted in 0012 for the radiographic endpoint was submitted to the Agency for review. Data presentation was submitted for patients who had DAS28(CRP) < 2.6 and had active joint counts at primary week of analysis.

1: Data for the 90-day safety update was submitted together in this sequence by the applicant.
2: The statistical reviewer did not consider the applicant's tipping point analysis to be complete and conducted his own tipping point analysis due to time constraints with this application.
Abbreviations: DMC=Data Monitoring Committee; IWRS=interactive web response system; IVRS=interactive voice response system; HAQ-DI=health assessment questionnaire-disability index; SF-36=short form-36; mTSS=modified total sharp score; CRP=C-reactive protein; DAS28(CRP)=disease activity score using 28 joints based on CRP
[Source: Statistical Reviewer]

3.2 Evaluation of Efficacy

3.2.1 Study Design

For the remainder of this review, studies M13-542, M13-549, and M14-465 are collectively described as cDMARD add-on studies. I note that majority of the patients in these studies were on MTX. Likewise, studies M15-555 and M13-545 are collectively described as monotherapy

studies. For cDMARD add-on studies, the reference arm is placebo. For monotherapy studies, the reference arm is the MTX monotherapy.

For all studies, the controlled period for cDMARD add-on studies and monotherapy studies was at minimal 12 weeks (Table 2). In all studies except M14-465, both doses of upadacitinib were evaluated. There were differences in the screened patient population, randomization ratios, doses evaluated, duration of controlled period that were considered critical in this review and were limiting factors in critical safety evaluation.

Study M13-542 (NCT02706847) was a 24-Week cDMARD add-on trial in which 499 screened patients with moderate to severe active rheumatoid arthritis, dosed at 153 study sites located in 26 countries, who had an inadequate response to a bDMARD were randomized in a 1:1:1 ratio to receive received placebo, upadacitinib 15 mg QD or upadacitinib 30 mg QD. Placebo patients were pre-randomized in a 1:1 ratio to switch, in a blinded manner, to upadacitinib 15 or 30 mg QD at Week 12.

Study M13-549 (NCT02675426) was a 12-Week cDMARD add-on trial in which 661 screened patients with moderate to severe active rheumatoid arthritis, dosed at 150 study sites located in 35 countries, who had an inadequate response to cDMARD were randomized in a 1:1:1 ratio to receive received placebo, upadacitinib 15 mg QD or upadacitinib 30 mg QD. Placebo patients were pre-randomized in a 1:1 ratio to switch, in a blinded manner, to upadacitinib 15 or 30 mg QD at Week 12.

Study M14-465 (NCT02629159) was a 26-Week MTX add-on trial in which 1629 screened patients with moderate to severe active rheumatoid arthritis, dosed at 286 study sites located in 41 countries, who were inadequate response to MTX were randomized in a 2:2:1 ratio to receive placebo, upadacitinib 15 mg QD or adalimumab 40 mg every other week (EOW). At either Week 14, 18, or 22, patients with inadequate efficacy response, as measured by no more than 20% improvement in both tender joint counts (TJC) and swollen joint counts (SJC) relative to baseline, were offered escape to other treatment arms. Patients on adalimumab arm meeting the criteria were crossed over to receive upadacitinib 15 mg QD. Patients on upadacitinib arm meeting the criteria were crossed over to receive adalimumab 40 mg EOW. Placebo patients meeting the criteria were crossed over to receive upadacitinib 15 mg QD.

In Study M14-465, after Week 26, all placebo patients remaining in the study, regardless of efficacy response, were crossover to upadacitinib 15 mg QD until Week 48. Patients randomized to adalimumab (respectively upadacitinib) and were not early escaped prior to Week 26 and had a clinical disease activity index (CDAI) score ≤ 10, remained on respective treatment. Patients randomized to adalimumab who were not escaped prior to week 26 and did not achieve a CDAI score ≤ 10 at Week 26 were switched to blinded upadacitinib arm. Patients randomized to upadacitinib who were not escaped prior to Week 26 but did not achieve a CDAI score ≤ 10 at Week 26 were switched to blinded adalimumab arm. Because of the multiple crossovers, the lack of placebo patients, the interpretation of safety results or efficacy data after week 26 is limited.

Table 2 Summary of Study Designs

Study	**M13-542** (N=499) [1]	**M13-549** (N=661)	**M14-465** (N=1629)	**M15-555** (N=648)	**M13-545** (N=947) [2]
Patient Population	bDMARD-IR hsCRP ≥ 3 mg/L	cDMARD-IR hsCRP ≥ 3 mg/L	MTX-IR hsCRP ≥ 5 mg/L	MTX-IR hsCRP ≥ 3 mg/L	MTX naïve hsCRP ≥ 5 mg/L
Design	MC, R, DB, PG, PC	MC, R, DB, PG, PC	MC, R, DBDD, PG, PC (Active comparator)	MC, R, DB, PG, AC	MC, R, DB, PG, AC
TP1	24 Weeks	12 Weeks	48 Weeks	14 Week	48 Weeks
TP2	Blinded LTE up to 216 weeks	Blinded LTE up to 5 years	LTE up to 5 years [3]	Blinded LTE (226 Weeks)	192-week LTE [3]
# Sites / Countries	153 / 26	150 / 35	286 / 41	138 / 24	286/ 41
# of Patients Per Arm	PBO (n=169) UPA 15 mg (n=165) UPA 30 mg (n=165)	PBO (n=221) UPA 15 mg (n=221) UPA 30 mg (n=219)	PBO (n=651) UPA 15 mg (n=651) ADA 40 mg [4] (n=327)	MTX (n=216) UPA 15 mg (n=217) UPA 30 mg (n=215)	MTX (n=315) UPA 15 mg (n=317) UPA 30 mg (n=315)
Stratification Factor	Geographic region Prior bDMARD use [5]	Geographic region Prior exp. to bDMARD	Geographic region Prior bDMARD exposure	Geographic Region	Geographic Region
Primary Timepoint	12	12	12	14	12
Rescue Strategy After primary Timepoint (Before W24/26)	PBO patients were pre-randomized in a 1:1 ratio to cross over to UPA 15 mg or UPA 30 mg after W12		Based on <20% improvement in either SJC or TJC evaluated at weeks 14, 18, and 22. PBO -> UPA UPA -> ADA ADA -> UPA	MTX patients were pre-randomized to cross over to UPA 15 mg or UPA 30 mg after W14	Based on <20% improvement in either SJC or TJC at two consecutive efficacy assessments Initiate or increase background RA medications

1: One patient from 15 mg arm was randomized but did not receive study drug.
2: One patient from MTX arm and one from 30 mg arm was randomized but did not receive study drug.
3: Long term extension is unblinded when the last patient completed TP1.
4: The dosing regimen for adalimumab is 40 mg EOW.
5: The categories are failed one or two biologics with same mechanism of action (MOA) vs failed more than three biologics with same MOA and/or multiple MOA)
Abbreviations: MC=multi-center; R=randomized; PG=parallel group; DB=double-blind; DBDD=double-blind double dummy DB; PC=placebo-controlled; AC=active controlled; QD=once daily; EOW=every other week; IR=inadequate response; PBO=placebo; UPA=upadacitinib; MTX=methotrexate; ADA=adalimumab; hsCRP=high sensitivity C-reactive protein; b/cDMARD=biologic/conventional disease modifying anti-rheumatic drugs; LTE=long-term extension; ACR=American College of Rheumatology; CFB=change from baseline; CRP=C-reactive protein; DAS28=disease activity score using 28 joints based on CRP; mTSS=modified Total Sharp score; HAQ-DI=health assessment questionnaire-disability index; W=week; PCS=physical component score; FACIT-F=functional assessment of chronic illness therapy-fatigue; ADA=adalimumab; Sup=superiority; NI=non-inferiority
[Source: Statistical Reviewer]

Study M15-555 (NCT02706951) was a 14-Week MTX monotherapy active-comparator trial in which 648 screened patients with moderate to severe active rheumatoid arthritis who had an inadequate response to MTX were randomized in a 1:1:1 ratio to receive MTX monotherapy, upadacitinib 15 mg QD monotherapy or upadacitinib 30 mg QD monotherapy. At Week 14, all patients randomized to MTX monotherapy were advanced in a blinded fashion to a second predetermined treatment of upadacitinib 15 or 30 mg QD in a 1:1 ratio.

Study M13-545 (NCT02706873) was a 48-Week trial with 24-week active-controlled period in which 947 screened patients with moderate to severe active rheumatoid arthritis who were naïve to MTX were randomized in a 1:1:1 ratio to receive MTX monotherapy, upadacitinib 15 mg QD monotherapy or upadacitinib 30 mg QD monotherapy. After Week 12, patients who did not achieve ≥ 20% improvement in both TJC and SJC compared with baseline at two consecutive visits could optimize (initiate or increase) their background RA medications such as NSAIDs, corticosteroids (oral ≤ 10 mg/day prednisone or prednisone ≤ 0.5 mg/kg/day for 3 consecutive days) and or low-potency analgesics.

In study M13-545, after Week 26, if patients did not achieve ≥ 20% improvement in both TJC and SJC relative to baseline and had CDAI score > 2.8, patients on MTX arm were re-randomized in a 1:1 ratio to receive upadacitinib 15 mg QD or 30 mg QD on top of randomized MTX treatment in a blinded manner for the remainder of Period 1 or until the study as unblinded. Patients on upadacitinib monotherapy arm would additionally receive MTX 10 mg/week throughout the remainder of Period 1 or until the study is unblinded. Patients who achieved more than 20% improvement in both TJC and SJC relative to baseline but have CDAI score > 2.8 continued blinded randomized treatment but could optimize (initiate or increase) background RA medications throughout the remainder of Period 1.

Radiograph evaluation, read via a central imaging laboratory, was assessed in M14-465 and M13-545. In M14-465, radiographs were collected at baseline, Week 26, and 48. Study M13-545 collected radiographs at baseline, Week 24, and 48. In this review, radiographs collected after Week 24/26 were not of interest in this review since there were no prospective planned comparison at Week 48. In Study M14-465, there were no placebo patients remaining on randomized treatment after Week 26.

An x-ray was taken at Week 14 and Week 12 for study M14-465 and M13-545 respectively under certain conditions. In study M14-465, patients with inadequate efficacy response, as measured by no more than 20% improvement in both TJC and SJC relative to baseline at Week 14, were crossed over to upadacitinib had an evaluation prior to Day 141. In study M13-545, patients who discontinued randomized treatment between Week 12 and Week 16 but continued to participate in the study to collect efficacy assessments would have a Week 12 x-rays taken if they discontinued; No new x-rays were taken if they discontinued randomized treatment prior to Week 12.

3.2.2 Study Endpoints

The primary analysis timepoint for non-radiographic endpoints for all studies (except M15-555) is at Week 12. The primary analysis timepoint for non-radiographic endpoints is Week 14 in study M14-465.

In all studies except M13-545, the primary endpoint was the proportion of randomized patients with ACR 20% improvement (ACR20) response at the primary timepoint (Table 3). Study M13-545 used ACR50 response as the primary efficacy endpoint. Treatment comparisons were made for each dose of upadacitinib with the reference arm when the study evaluated both doses of upadacitinib. In study M14-465, treatment comparisons were made with respect to placebo arm.

The primary analysis timepoint for radiographic in study M14-465 is Week 26.

The primary analysis timepoint for radiographic in study M13-545 is Week 24.

Table 3 Primary Efficacy Endpoint, Key Secondary Endpoints, and Other Endpoints

Study	M13-542	M13-549	M14-465	M15-555	M13-545
Primary Endpoint/ Week	ACR20/ Week 12	ACR20/ Week 12	ACR20/ Week 12	ACR20/ Week 14	ACR50/ Week 12
Key Secondary Endpoints (Ranked in the order based on multiplicity)	CFB DAS28(CRP) CFB HAQ-DI DAS28 ≤ 3.2 CFB SF-36 PCS	CFB DAS28(CRP) CFB HAQ-DI CFB SF-36 PCS DAS28≤3.2 DAS28<2.6 CDAI ≤10 CFB morning stiffness CFB FACIT-F	CFB DAS28(CRP) CFB mTSS (w26) CFB HAQ-DI ACR50 (NI vs ADA) CFB SF-36 PCS DAS28≤3.2 DAS28<2.6 CDAI ≤10 CFB morning stiffness CFB FACIT-Fatigue ACR50 (Sup vs ADA) CFB Pain (Sup vs ADA) CFB HAQ-DI (Sup vs ADA)	CFB DAS28(CRP) CFB HAQ-DI CFB SF-36 PCS DAS28(CRP) ≤3.2 DAS28(CRP) <2.6 CFB morning stiffness	CFB DAS28(CRP) CFB HAQ-DI ACR20 CFB mTSS @ w24 ACR70 DAS28 ≤3.2 @W24 DAS28 <2.6 @W24 CFB SF-36 PCS
Other Secondary Endpoints [1]	ACR50 ACR70 ACR20 (W1)	ACR50 ACR70 ACR20 (W1)	ACR50 ACR70 No radiographic progression (W26)	ACR50 ACR70	ACR20 ACR70 No radiographic progression (W24)

1: Other key secondary endpoints were not included in the graphical hierarchy procedure.
Abbreviations: CRP= C-reactive protein; b/cDMARD=biologic/conventional disease modifying anti-rheumatic drugs; ACR=American College of Rheumatology; CFB=change from baseline; DAS28=disease activity score using 28 joints based on CRP; mTSS=modified Total Sharp score; HAQ-DI=health assessment questionnaire-disability index; W=week; PCS=physical component score; FACIT-F=functional assessment of chronic illness therapy-fatigue; ADA=adalimumab; Sup=superiority; NI=non-inferiority
[Source: Statistical Reviewer]

In all studies, the applicant prespecified a list of ranked secondary endpoints, generally similar but not necessary in the same order across studies, evaluated at the primary timepoint. In all studies except M14-465, the applicant pre-specified a graphical testing procedure to control across the testing of multiple endpoints and doses (Figure 9, Figure 10, Figure 11, Figure 12) using a family-wise 2-sided type 1 error rate of α=0.05. In all studies, the applicant listed other key secondary endpoints (ACR at other thresholds at the primary efficacy timepoint) which were not included in the graphical procedure.

In study M14-465, the applicant pre-specified the sequential testing procedure to control the family-wise 2-sided type 1 error rate of α=0.05 across the ranked endpoints and comparisons. The applicant specified a non-inferiority margin of 10% using ACR50 comparing upadacitinib and adalimumab. It is unclear from the protocol the objective of this comparison in presence of a placebo arm, which allows us to directly compare the efficacy of upadacitinib with placebo. Further, the protocol had not justified why such a margin evaluated at Week 12 would be beneficial to patients and prescribers.

3.2.3 Statistical Methodologies

3.2.3.1 Primary Statistical Analysis for Primary and Key Secondary Endpoints

The statistical analysis plans for the non-radiographic endpoints for the studies M13-542, M13-549, M14-465, and M15-555 were similar. In study M13-545, the SAP was not submitted for review and the statistical analysis for some continuous secondary endpoints differed from the four studies.

In each study, the dataset used for the primary statistical analysis was based on all randomized subjects who received at least a dose of study treatment. Descriptive statistics were reported by the applicant. Continuous variables were summarized and reported using the minimum, 25th percentile, mean, median, 75th percentile, maximum, and standard deviation. Binary or categorical variables were summarized and reported as counts and percentages.

In each study, the primary efficacy endpoint, the proportion of ACR20 responders (or ACR50 responders for study M13-545), and key secondary responder type endpoints, were analyzed using the difference in binomial proportions by comparing each dose of upadacitinib arm with the reference arm. The 95% CI was based on normal approximation to the difference in binomial proportions. P-values were calculated based on the Cochran Mantel-Haenszel (CMH) test adjusting for the stratification factor listed in Table 4. In studies M13-542, M13-549, and M14-465, geographical region was noted to be an additional stratification variable for randomization but was not included as a variable in the statistical analysis.

Table 4 Stratification Factor used for the Primary Statistical Analysis by Study

Study	**M13-542**	**M13-549**	**M14-465**	**M15-555**	**M13-545**
Stratification factor used for statistical analysis	Prior bDMARD [1]	Prior exposure to bDMARD (Yes or No)	Prior exposure to bDMARD (Yes or No)	Geographical Region (North America, South/Central, America, Eastern Europe, Asia-Japan, or Other)	Geographical Region (North America, South/Central, America, Eastern Europe, Asia/Other)

1: The categories are failed 1 or 2 biologics with the same MOA, or failed ≥ 3 biologics with the same MOA and/or multiple MOA
Abbreviations: bDMARD=biological disease modifying anti-rheumatic drugs; MOA=mechanism of action
[Source: Statistical Reviewer]

In each study, the change from baseline in HAQ-DI and change from baseline in DAS28(CRP) at the week of analysis were fit using a linear regression, assuming homoskedasticity, adjusting for treatment, continuous baseline value, and stratification factor(s) listed in Table 4. The difference in adjusted means comparing each dose of upadacitinib arm relative to the reference arm, respective 95% CI, and Wald-based p-values from the regression model were reported.

In all studies except M13-545, the remaining continuous secondary endpoints was analyzed using the mixed model repeated measurement (MMRM). MMRM using the method of restricted maximum likelihood was fit to the change from baseline adjusting for treatment, visit as categorical variable, interaction of visit and treatment, continuous baseline value, and key stratification factor. An unstructured variance-covariance matrix was used to model the within-

patient errors. The difference in the adjusted means at the week of analysis comparing each dose of upadacitinib with the reference arm, respective 95% CI with the method of Kenward-Roger approximation used to estimate the denominator degrees of freedom, and Wald-based p-values were reported.

In study M13-545, the change from baseline in the continuous secondary endpoints at the week of analysis, was fit using a linear regression, assuming homoskedasticity, adjusting for treatment, continuous baseline value, and the stratification factor listed in Table 4. The difference in adjusted mean change from baseline comparing each dose of upadacitinib monotherapy with MTX monotherapy, respective 95% CI, and Wald-based p-values from the regression model were reported.

In study M14-465, the change from baseline in mTSS at Week 26 was fit using a linear regression, assuming homoscedasticity, adjusting for baseline mTSS, stratification factor, and treatment variable. The difference in the adjusted mean change from baseline in mTSS at Week 26 compared upadacitinib 15 mg QD with reference arm, respectively 95% CI, and Wald-based p-value from the regression model were reported. In study M13-545, similar regression model was fit to the change from baseline in mTSS at Week 24.

Stratified subgroup analyses were conducted based on the primary efficacy endpoint ACR20 (or ACR50 for study M13-545). Point estimates comparing dose of upadacitinib with the reference arm was presented using difference in proportions for each subcategory within the subgroups. The respective 95% CI using normal approximation to the difference in binomial proportions was reported. In addition, I reported subgroup analysis by US vs non-US patients.

In all studies, these subgroups had the same sub-categories:

- Age (< 40, 40 to 64, ≥ 65)
- Sex (male or female)
- Weight (< 60 kg or ≥ 60 kg)
- BMI (< 25 or ≥ 25)
- Race (white, non-white)
- Baseline RF status (positive or negative)
- Baseline anti-CCP antibody status (positive or negative)
- Baseline both RF positive and anti-CCP positive (both positive versus at least 1 negative)
- Baseline both RF negative and anti-CCP negative (both negative versus at least 1 positive)
- Baseline DAS28 (hsCRP) (≤ 5.1 or > 5.1)

The subgroups that were different in the subcategories are as follows:

- The duration of RA diagnosis was categorized as < 5 years or ≥ 5 years in studies M13-549, M14-465, and M15-555. In study M13-545, the duration of RA diagnosis was categorized as < 6 months or ≥ 6 months and in study M13-542, the categories were < 10 years vs ≥ 10 years.

- The categories for geographic region were similar for studies M13-542 and M13-545 (North America, South/Central America, Western Europe, Eastern Europe, Asia/Other). In study M13-549, the region categories were North America, Western Europe, Eastern Europe, and Other. In study M14-465, the region categories were North America, South/Central America, Eastern Europe, and Other. In study M15-555, the region categories were North America, South/Central America, Eastern Europe, Japan, and Other.
- Study M13-542 included prior failed bDMARD and failed anti-IL6 as additional subgroups of interest. Study M13-549 included prior bDMARD use as additional subgroups of interest. Other studies did not include/collect bDMARD as a subgroup of interest.

In this review, I reported the efficacy findings by endpoints.

3.2.3.2 Handling of Missing Data for Primary Statistical Analysis and Post-hoc Sensitivity Analyses

In each study, the applicant had defined the following missing data approaches: non-responder imputation (NRI), observed cases (OC) analysis, and as observed (AO) analysis. The NRI approach imputes patients with missing value for binary endpoints at a specific visit as a non-responder. Patients who discontinued randomized study treatment prior to the end of the double-blind period had their outcomes imputed as non-response for the remainder of the study regardless of whether follow-up efficacy assessments were conducted. In the OC approach, patients with missing evaluations at a visit, or have discontinued randomized study treatment prior to end of double-blind period were excluded from the analysis. In the AO approach, only patients with missing evaluations at a visit were excluded from the analysis. All patients whose follow-up data were collected regardless of early discontinuation of randomized treatment, or crossover of treatment arms are included in the statistical analysis.

In the five studies, the applicant applied NRI for the primary efficacy endpoint and other responder-type endpoints at the primary analysis timepoint. When some components of the multi-component endpoints, such as ACR20, were missing for a patient at any visit, the patient's response was determined from the remaining non-missing components. When the responder status cannot be determined from the non-missing components, the patient's outcome was imputed as a non-response.

It is reasonable to use NRI when patients discontinued randomized treatment because this targets a composite estimand whereby having early escape or switching to rescue is considered a lack of efficacy. For example, the primary ACR20 endpoint at the primary analysis timepoint is a composite measure that includes the following components: (1) remaining in the study, on treatment, and without protocol-prohibited medication changes through 12 weeks; and (2) achieving an ACR20 response at the primary week of analysis. The use of such a composite outcome combines the effects of treatment on adherence and signs and symptoms, so it is also important to evaluate the treatment effect on each of the components of the composite outcome in supportive analyses.

To evaluate the components of the composite estimand, the applicant had included supportive analyses for the primary efficacy endpoint based on the AO approach. Because it is likely that there is residual missing data even when follow-up data are included in the AO approach, evaluation of the robustness of the results to potential violations of missing data assumptions is critical. The applicant had not prespecified how missing data would be evaluated in the final versions of the SAP, based on the Agency's recommendations. The applicant included post-hoc "tipping point analysis" for ACR in the CSR to address missing data issues. However, this analysis was not considered a tipping point analysis because the applicant only considered the extreme scenario where all missing data from the reference arm was imputed as responders while missing data from the upadacitinib dose group were imputed as non-responders. Further, the approach was a single imputation procedure that did not consider the uncertainty of the imputation procedure.

The primary endpoint ACR20 is a multi-component endpoint, evaluated at primary analysis timepoint, that includes direct measures of how patients feel and function in daily life, such as the tender and swollen joint counts and HAQ-DI physical ability score, as well as acute phase reactants such as CRP. Hence, significant results based on ACR20 response should be supported by trends of improvements with magnitudes that are considered clinically relevant in the components of ACR20. The applicant had included additional supportive analysis to summarize each of the components based on similar linear regression analysis, assuming homoskedasticity, used for the continuous endpoints HAQ-DI and DAS28(CRP) at visit weeks prior and up to the primary week of analysis.

Handling of missing data for continuous secondary endpoints was handled differently depending on the study. For all continuous endpoints, efficacy assessments collected after patients had discontinued randomized treatment were treated as missing. In study M13-545, the multiple imputation approach was used to impute missing data for all continuous non-radiographic endpoints specified in the multiplicity hierarchy. In the remaining studies, multiple imputation was used only for HAQ-DI and DAS28(CRP). The procedure is as follows. Last observation carried forward (LOCF) was used to augment the dataset to have a monotone missing pattern. The multiple imputation approach then sequentially imputed the missing outcome at each subsequent post-baseline visit by regressing upon outcomes at previous visits, baseline value of the outcome, treatment group, demographics, and other key baseline characteristics. A total of five MI datasets was generated, and the primary statistical analysis model was applied to each of the imputed datasets. Rubin's rule was used to combine the estimates and standard errors.

For the remaining continuous non-radiographic endpoints, there was no imputation of data after patients had discontinued randomized treatment or were missing data at the visits. Instead, the applicant stated that the MMRM model will implicitly handle the missing data based on the missing at random (MAR) assumption. Under this assumption, the probability of a missing data point depends only on the patients' observed data and, conditional on these observed data, is not related to the value at the missing data point. In other words, conditional on the observed values and baseline covariates included in the model, the distribution of outcomes after a patient discontinues from the randomized study drug has a similar distribution like patients who had not discontinue randomized study drug and remained in the study. To justify the MAR assumption, the applicant stated that the observed placebo response rate was high and that the proportion of

patients discontinuing in RA studies was between 10-15% and that discontinuation rates due to lack of efficacy was < 5%.

In several communications to the applicant, the Agency had expressed concerns on the validity of the assumptions of the MMRM analysis. Therefore, the Agency had requested that the applicant prospectively include two-dimensional tipping point sensitivity analyses to assess the departures of MAR assumption on the key conclusions. Despite that, in the final version of the SAP, for all studies, the applicant specified, with limited details, that a missing not at random (MNAR) model that "varies assumptions for the missing data in active treatment groups and placebo (respectively MTX monotherapy) group may be used as a sensitivity analysis for key continuous endpoints to account for potential deviation from the MAR assumption." The applicant included post-hoc tipping point analysis focused on the scenarios where missing outcomes on upadacitinib were worse than the observed data on upadacitinib, while missing outcomes on reference arm were better than the observed data on reference arm. The missing values were imputed using the above described MI model. A shift parameter was applied to the imputed values to create the tipping point grid. The minimum or the maximum was substituted to the imputed values if the shifted values were outside the range of plausible values for the endpoint of interest. The same primary statistical analysis was applied to analyze each imputed dataset for each pair of shift parameters. Rubin's rule was used to combine the results for each pair of shift parameters.

3.2.3.3 Estimands for Radiographic Endpoints

To facilitate describing the planned statistical analysis, it is useful to describe the estimand of interest for radiograph progression. There are two key estimands of interest. The first, typically used in reporting of radiograph progression in RA product insert, is defined as the difference in radiographic progression based on all randomized patients if the possibility of escape or rescue to alternative therapies were not available or discontinuation of randomized study treatment was not possible and patients without visit week measurements experienced the same rate of radiograph progression had an evaluation been made at or prior to the time of escape or discontinuation of randomized treatment. This estimand corresponds to the "hypothetical strategy" described in the ICH-E9(R1) addendum. The second, the treatment policy estimand, is defined as the difference in radiographic progression based on all randomized subjects regardless of any escape or discontinuation from study treatment. The goal of the treatment policy estimand seeks to evaluate the potential effectiveness of upadacitinib under the real-world conditions where placebo patients have opportunities to receive other treatments.

In studies M13-545 and M14-465, the primary estimand of interest for the radiographic endpoint is the hypothetical estimand. The applicant had included supplementary analysis to address the treatment policy estimand. The primary radiograph evaluation week for studies M13-545 and M14-465 is at Week 24 and Week 26 respectively.

In study M14-465, to address the primary estimand of interest, when patients prematurely discontinued originally randomized study drug at or after Week 14 and at or before Week 18, their radiograph data collected at Week 26 were excluded from the primary analysis. For these patients, if a Week 14 x-ray, prior to Day 141, were taken, then the applicant used the baseline

and Week 14 x-ray and fit a straight line to predict the Week 26 mTSS outcome for these patients. This approach is called linear extrapolation. If the Week 26 x-ray data was missing, and if a Week 14 x-ray was collected, then linear extrapolation would be used to extrapolate the x-ray outcome at Week 26.

In study M13-545, the linear extrapolation procedure imputed patients who discontinued randomized treatment at or after Week 12 and prior to Week 16. X-rays collected from patients who discontinued randomized treatment after Week 16 will be treated as missing. For these patients, if an x-ray was collected between Week 12 and 16, then the x-ray will be used to extrapolate Week 24 results. If patients were missing Week 24 x-ray data but had Week 12 x-ray evaluated, then the Week 12 x-ray will be used to extrapolate Week 24 results.

The primary statistical model used is a linear regression fit to the change from baseline in mTSS at week of interest adjusting for baseline mTSS, treatment group, stratification factor listed in Table 4. In M13-545, data from upadacitinib 15 mg and 30 mg arm were included in the regression model. Likewise, for study M14-465, radiograph data from the adalimumab arm was included in the same statistical model.

To analyze the proportion of patients with no radiographic progression, linear extrapolation was used to impute x-ray scores at Week 24/26 (M13-545/M14-465) for patients who discontinue at or after Week 12/14 and at or before Week 16/18. Patients with missing baseline or Week 24/26 data were excluded from the statistical analysis. Point estimates based on difference in the proportion of patients with no radiographic progression (change ≤ 0), respective 95% CI based on normal approximation to difference in binomial proportions, and p-values from the CMH test stratified by geographic region were reported.

3.2.4 Additional Reviewer Analyses

During the review of the application, I identified the following limitations with the SAP. Chiefly, they are related to the level of details specified.

There were limited details described in the SAP for the multiple imputation procedure. First, the covariates that were used to impute missing observations were not clearly specified in the SAP. Because the multiple imputation inference assumed that the model used to analyze the multiply imputed data was like the model used to impute missing values in the multiple imputation, the MI procedure in the SAP did not target similar estimand as the primary statistical model. Second, the applicant used last-observation carried forward, a single imputation procedure, for data that were not monotone missing to create monotone missing pattern. Even though the proportion of non-monotone observations is low, the single imputation step would not handle the uncertainty of the imputed data. Third, the applicant had only conducted five multiple imputations, which were insufficient to minimize the between imputation variability.

The tipping point analysis in the final version of the SAP for each study were not prospectively defined. Further, the tipping point analyses were dependent on the multiple imputation model. Because of these reasons, it was unclear whether the applicant had chosen covariates or models such that the tipping point analysis was consistent with the primary findings.

To explore the sensitivity of the results to violations of these missing data assumptions, I included simple tipping point analyses for ACR20 (ACR50), HAQ-DI, and mTSS. We performed separate tipping point analyses for each dose versus placebo. This tipping point analysis varied the assumed mean differences between outcomes in dropouts and outcomes in completers on each of the two treatment arms. Therefore, we can determine how much worse outcomes in patients who dropped out from the study from the treatment arm would have to be than outcomes in patients who dropped out from the study from the placebo arm such that conclusions of efficacy would change. The details of this tipping point approach are provided in the 6.2.

As previously stated in 3.2.3.2, the MMRM model used for other continuous key secondary endpoints made unverifiable assumptions about the unobserved missing values after discontinuation of randomized treatment, in addition to constant variance and normality. Because these endpoints were generally assessed prior to time of crossing-over, in this review, I conducted and reported analyses that targeted the intent-to-treat or de facto estimand, i.e., the difference in outcomes between treatment arms in all randomized patients regardless of adherence to randomized treatment at the primary efficacy timepoint. I reported the observed means, standard deviations, and applied a linear regression model, similar to the statistical model for HAQ-DI, fit to all data collected at the week of analysis regardless of adherence to randomized treatment and reported the results.

For these endpoints, I regarded the statistical analyses based on all randomized patients regardless of adherence to randomized treatment, or use of ancillary treatments (i.e., targeting the intent-to-treat or de facto estimand) to be the important analyses to evaluate the effects of upadacitinib relative to the reference arm on these direct measures of how patients function and feel. However, follow-up efficacy data collected from patients after discontinuation of randomized treatment can continue to have an effect on signs and symptoms. Therefore, the estimated treatment effects from such analyses for symptomatic endpoints, such as ACR20 and HAQ-DI, may be smaller in magnitude for these patients than those who continue to receive active treatment when included in the analysis.

In study M14-465, the presence of crossovers between treatments, i.e., from upadacitinib to adalimumab, or placebo to upadacitinib, creates challenges in analyses of endpoints after Week 14. Although NRI was used for responder type endpoint, the estimand of interest is unclear and inconsistent with the interpretation of results at the primary efficacy timepoint, assessed prior to escape. Therefore, in this review, I focused evaluation of efficacy data prior to Week 14.

In study M14-465, the applicant had pre-specified a non-inferiority margin of 10% at Week 12 to compare upadacitinib with adalimumab. As noted in 2.1.2, concerns were raised on the use of an NI margin. I did not agree with the proposed margin for the following reasons. First, it is reasonable to conduct comparative effectiveness studies between products in an active controlled study. However, with a placebo arm in place, the objective is less clear. Second, it is more reasonable to conduct the NI comparison at a later timepoint due to the chronic nature of the disease. Thirdly, although later efficacy timepoints were evaluated in M14-465, escape was potentially possible at multiple timepoints. Patients randomized to active comparator can potentially escape to investigational product, breaking the integrity of the randomization, further limits efficacy and safety comparisons between active control and upadacitinib.

In several communications as noted in 2.1.2, concerns were raised with the applicant on the appropriateness of the method of linear extrapolation, typically used in RA studies. The linear extrapolation approach assumed that placebo patients' scores would, in the absence of crossover to another treatment, continue to change at the same linear rate as was observed through the time of escape. This assumption is strong and unverifiable, and may in some situations tend to overstate true progression on placebo. Further, the linear extrapolation approach is a single-imputation method that, even if the overall progression is truly linear on average, does not appropriately consider the statistical uncertainty of the imputation process. This can potentially lead to underestimates of the variability and potentially overestimates of the degree of evidence of a treatment effect.

Therefore, I included analyses to more reliably evaluate the hypothetical estimand in the absence of crossover. I considered the following estimand, "the difference in mean rate of change through 26 weeks between treatment conditions (upadacitinib dose relative to placebo) in mTSS in the hypothetical scenario where escape therapy was not available to any patients on any arm and that patients continue to receive randomized treatment through Week 26, and that these patients were likely to progress in the same linear rate like those remaining on randomized treatment, in the targeted patient population of inadequate responders to MTX who are on a stable background dose of MTX. I included a random coefficient model, allowing for random slope and random intercept, allowing for unequal variance between treatment groups, to estimate the difference in the mean rate of mTSS progression up to Week 26 comparing upadacitinib and placebo based on all radiograph data collected prior to escape. I note that similar model can be applied to attempt to address the treatment policy estimand with the additional assumption of linear progression within 6 months.

We find merit in the random coefficient model since it more appropriately accounts for the statistical uncertainty. However, this approach still relies on a missing at random assumption. Therefore, sensitivity analyses with two-dimensional tipping point analysis would be important to address the robustness of the results to deviations of missing data assumptions.

An integrated summary of efficacy (ISE) for cDMARD add-on studies is presented for reporting purposes in drug trial snapshot. Therefore, the SAP for the ISE submitted by the sponsor was reviewed. In the ISE, the subgroups were refined for age (<65, 65 – 75, ≥ 75), weight (<60, 60 to <100, ≥ 100), and BMI (< 25, 25 to < 30, ≥ 30). Although the categories for geographical region was kept similar for studies M13-542, M13-549, and M14-465, study M13-542 did not randomize any patients from South/Central America. Therefore, I refined the categories to North America, Western Europe, Eastern Europe, Asia/Other to avoid Simpson paradox. I also reported age subgroups based on categories used in the individual studies, namely, < 40, 40 to 64, and ≥ 65.

3.2.5 Patient Disposition, Demographic and Baseline Characteristics

3.2.5.1 Patient Disposition

In study M13-542, 92% of the randomized patients completed Week 12 visit regardless of adherence to randomized treatment (Table 5). Patients on upadacitinib 15 mg QD arm had the lowest rate of discontinuation from study. A total of 91% of the randomized patients completed randomized treatment at Week 12 with patients on upadacitinib 15 mg QD having the lowest rate of discontinuation from randomized treatment. The most common reason for study discontinuation or treatment discontinuation was adverse events and a markedly higher proportion of patients on upadacitinib 30 mg QD arm discontinued the study for reasons of adverse events.

Table 5 Patient Disposition at Week 12, Study M13-542

Study M13-542	Placebo + cDMARD (N=169)	UPA 15 mg QD + cDMARD (N=164)	UPA 30 mg QD + cDMARD (N=165)	All (N=498)
Completed Week 12 Study	151 (89%)	157 (96%)	149 (90%)	457 (92%)
Discontinued Week 12 Study	18 (11%)	7 (4%)	16 (10%)	41 (8%)
Adverse event	4 (2%)	1 (1%)	12 (7%)	17 (3%)
Lack of efficacy	-	-	-	-
Lost to follow-up	3 (2%)	-	-	3 (1%)
Other	8 (5%)	2 (1%)	2 (1%)	12 (2%)
Withdrew consent	3 (2%)	4 (2%)	2 (1%)	9 (2%)
Completed Week 12 Treatment	147 (87%)	156 (95%)	148 (90%)	451 (91%)
Discontinued Week 12 Treatment	22 (13%)	8 (5%)	17 (10%)	47 (9%)
Adverse event	7 (4%)	3 (2%)	13 (8%)	23 (5%)
Lack of efficacy	10 (6%)	1 (1%)	-	11 (2%)
Lost to follow-up	3 (2%)	-	-	3 (1%)
Other	1 (1%)	1 (1%)	2 (1%)	4 (1%)
Withdrew consent	1 (1%)	3 (2%)	2 (1%)	6 (1%)

Counts and percentages relative to N in parenthesis.
Abbreviations: UPA=upadacitinib; QD=once daily; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In study M13-549, 94% of the randomized patients completed Week 12 visit regardless of adherence to randomized treatment (Table 6). A total of 93% of the randomized patients completed randomized treatment at Week 12. The most common reason for study discontinuation or discontinuation of randomized treatment was adverse events and the rates were similar across arms.

Table 6 Patient Disposition at Week 12, Study M13-549

Study M13-549	Placebo + cDMARD (N=221)	UPA 15 mg QD + cDMARD (N=221)	UPA 30 mg QD + cDMARD (N=219)	All (N=661)
Completed Week 12 Study	208 (94%)	213 (96%)	201 (92%)	622 (94%)
Discontinued Week 12 Study	13 (6%)	8 (4%)	18 (8%)	39 (6%)
Adverse event	6 (3%)	2 (1%)	9 (4%)	17 (3%)
Lack of efficacy	-	-	-	-
Lost to follow-up	1 (0%)	-	2 (1%)	3 (0%)
Other	3 (1%)	1 (0%)	2 (1%)	6 (1%)
Withdrew consent	3 (1%)	5 (2%)	5 (2%)	13 (2%)
Completed Week 12 Treatment	207 (94%)	210 (95%)	201 (92%)	618 (93%)
Discontinued Week 12 Treatment	14 (6%)	11 (5%)	18 (8%)	43 (7%)
Adverse event	5 (2%)	5 (2%)	8 (4%)	18 (3%)
Lack of efficacy	4 (2%)	-	-	4 (1%)
Lost to follow-up	1 (0%)	-	2 (1%)	3 (0%)
Other	2 (1%)	1 (0%)	2 (1%)	5 (1%)
Withdrew consent	2 (1%)	5 (2%)	6 (3%)	13 (2%)

Counts and percentages relative to N in parenthesis.
Abbreviations: UPA=upadacitinib; QD=once daily; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In study M14-465, 95% of the randomized patients completed Week 12 visit regardless of adherence to randomized treatment (Table 7). At Week 26, 92% of the patients completed Week 26 visit regardless of adherence to randomized treatment, crossing over between study drugs. The rate of discontinuation from study was similar between arms and withdrawal of informed consent was the more common reasons for study discontinuation.

Table 7 Patient Disposition at Visit Weeks of Interest, Study M14-465

Study M14-465	Placebo + cDMARD (N=651)	UPA 15 mg QD + cDMARD (N=651)	ADA 40 mg EOW + cDMARD (N=327)	All (N=1629)
Completed Week 12 Study	625 (96%)	625 (96%)	305 (93%)	1555 (95%)
Discontinued Study prior to Week 12	26 (4%)	26 (4%)	22 (7%)	74 (5%)
Adverse event	7 (1%)	12 (2%)	10 (3%)	29 (2%)
Lack of efficacy	-	-	-	-
Lost to follow-up	4 (1%)	2 (0%)	1 (0%)	7 (0%)
Other	3 (0%)	4 (1%)	1 (0%)	8 (0%)
Withdrew consent	12 (2%)	8 (1%)	10 (3%)	30 (2%)
Completed Week 18 Study	618 (95%)	619 (95%)	299 (91%)	1536 (94%)
Discontinued Study prior to Week 18	33 (5%)	32 (5%)	28 (9%)	93 (6%)
Adverse event	9 (1%)	13 (2%)	12 (4%)	34 (2%)
Lack of efficacy	-	-	-	-
Lost to follow-up	4 (1%)	3 (0%)	2 (1%)	9 (1%)
Other	5 (1%)	5 (1%)	2 (1%)	12 (1%)
Withdrew consent	15 (2%)	11 (2%)	12 (4%)	38 (2%)
Completed Week 22 Study	610 (94%)	615 (94%)	297 (91%)	1522 (93%)
Discontinued Study prior to Week 22	41 (6%)	36 (6%)	30 (9%)	107 (7%)
Adverse event	10 (2%)	14 (2%)	14 (4%)	38 (2%)
Lack of efficacy	2 (0%)	-	-	2 (0%)
Lost to follow-up	5 (1%)	3 (0%)	2 (1%)	10 (1%)
Other	7 (1%)	6 (1%)	2 (1%)	15 (1%)
Withdrew consent	17 (3%)	13 (2%)	12 (4%)	42 (3%)
Completed Week 26 Study	605 (93%)	606 (93%)	294 (90%)	1505 (92%)
Discontinued Study prior to Week 26	46 (7%)	45 (7%)	33 (10%)	124 (8%)
Adverse event	10 (2%)	16 (2%)	14 (4%)	40 (2%)
Lack of efficacy	2 (0%)	-	-	2 (0%)
Lost to follow-up	7 (1%)	4 (1%)	4 (1%)	15 (1%)
Other	8 (1%)	7 (1%)	2 (1%)	17 (1%)
Withdrew consent	19 (3%)	15 (2%)	13 (4%)	47 (3%)

Counts and percentages relative to N in parenthesis.
Abbreviations: UPA=upadacitinib; QD=once daily; ADA=adalimumab; EOW=every other week; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In study M14-465, 91% of the randomized patients completed treatment at Week 26 (Table 8). Patients on adalimumab 40 mg EOW had the highest discontinuation rate across study visit, with adverse events and withdrawal of informed consent being the most common reason for study

discontinuation. However, 31% of the randomized patients were rescued prior to Week 26 to other treatment arms, with majority of them rescued at Week 14 (Table 9). At Week 14, 35% of the placebo patients were crossover to upadacitinib, 12% of the upadacitinib patients were crossover to adalimumab, and 17% of the adalimumab patients were crossover to upadacitinib.

Table 8 Patient Disposition for Completing Treatment at Various Weeks, Study M14-465

Study M14-465	Placebo + cDMARD (N=651)	UPA 15 mg QD + cDMARD (N=651)	ADA 40 mg EOW + cDMARD (N=327)	All (N=1629)
Completed Week 12 Treatment	620 (95%)	620 (95%)	300 (92%)	1540 (95%)
Discontinued Treatment prior to Week 12	31 (5%)	31 (5%)	27 (8%)	89 (5%)
Adverse event	10 (2%)	15 (2%)	15 (5%)	40 (2%)
Lack of efficacy	1 (0%)	-	-	1 (0%)
Lost to follow-up	4 (1%)	3 (0%)	1 (0%)	8 (0%)
Other	2 (0%)	4 (1%)	2 (1%)	8 (0%)
Withdrew consent	14 (2%)	9 (1%)	9 (3%)	32 (2%)
Completed Week 18 Treatment	611 (94%)	616 (95%)	297 (91%)	1524 (94%)
Discontinued Treatment prior to Week 18	40 (6%)	35 (5%)	30 (9%)	105 (6%)
Adverse event	13 (2%)	17 (3%)	15 (5%)	45 (3%)
Lack of efficacy	1 (0%)	-	-	1 (0%)
Lost to follow-up	4 (1%)	3 (0%)	2 (1%)	9 (1%)
Other	4 (1%)	5 (1%)	3 (1%)	12 (1%)
Withdrew consent	18 (3%)	10 (2%)	10 (3%)	38 (2%)
Completed Week 22 Treatment	602 (92%)	610 (94%)	293 (90%)	1505 (92%)
Discontinued Treatment prior to Week 22	49 (8%)	41 (6%)	34 (10%)	124 (8%)
Adverse event	15 (2%)	20 (3%)	19 (6%)	54 (3%)
Lack of efficacy	3 (0%)	1 (0%)	-	4 (0%)
Lost to follow-up	5 (1%)	3 (0%)	2 (1%)	10 (1%)
Other	5 (1%)	5 (1%)	3 (1%)	13 (1%)
Withdrew consent	21 (3%)	12 (2%)	10 (3%)	43 (3%)
Completed Week 26 Treatment	595 (91%)	600 (92%)	288 (88%)	1483 (91%)
Discontinued Treatment prior to Week 26	56 (9%)	51 (8%)	39 (12%)	146 (9%)
Adverse event	17 (3%)	22 (3%)	20 (6%)	59 (4%)
Lack of efficacy	4 (1%)	1 (0%)	-	5 (0%)
Lost to follow-up	7 (1%)	5 (1%)	4 (1%)	16 (1%)
Other	6 (1%)	8 (1%)	4 (1%)	18 (1%)
Withdrew consent	22 (3%)	15 (2%)	11 (3%)	48 (3%)

Counts and percentages relative to N in parenthesis.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; ADA=adalimumab; EOW=every other week;
[Source: Statistical Reviewer]

Table 9 Proportion of Randomized Patients Rescued, Study M14-465

	Placebo + cDMARD (N=651)	UPA 15 mg QD + cDMARD (N=651)	ADA 40 mg EOW + cDMARD (N=327)	All (N=1629)
Rescued at Week 14	231 (35%)	78 (12%)	56 (17%)	365 (22%)
Rescued at Week 18	48 (7%)	29 (4%)	14 (4%)	91 (6%)
Rescued at Week 22	26 (4%)	18 (3%)	7 (2%)	51 (3%)
Rescued prior to Week 26	305 (47%)	125 (19%)	77 (24%)	507 (31%)

Patients who did not have >20% improvement in both TJC and SJC were rescued to other treatment arms.
Counts and percentages relative to N in parenthesis.
Abbreviations: UPA=upadacitinib; QD=once daily; ADA=adalimumab; EOW=every other week; TJC=tender joint counts; cDMARD=conventional disease modifying anti-rheumatic drugs; SJC=swollen joint counts
[Source: Statistical Reviewer]

In study M15-555, 94% of the randomized patients completed Week 14 study visit regardless of adherence to randomized treatment (Table 10). The most common reason for discontinuation of study was withdrawal of informed consent. A total of 92% of the randomized patients completed randomized treatment at Week 14. The rates of treatment discontinuation were similar across arms with adverse events and patients initiating withdrawal from study drug being the most common reason for treatment discontinuation.

Table 10 Patient Disposition at Week 14, Study M15-555

	MTX (N=216)	UPA 15 mg QD (N=217)	UPA 30 mg QD (N=215)	All (N=648)
Completed Week 14 Study	202 (94%)	201 (93%)	205 (95%)	608 (94%)
Discontinued Week 14 Study	14 (6%)	16 (7%)	10 (5%)	40 (6%)
Adverse event	1 (0%)	5 (2%)	3 (1%)	9 (1%)
Lack of efficacy	-	-	-	-
Lost to follow-up	-	4 (2%)	1 (0%)	5 (1%)
Other	2 (1%)	1 (0%)	-	3 (0%)
Withdrew consent	11 (5%)	6 (3%)	6 (3%)	23 (4%)
Completed Week 14 Treatment	197 (91%)	199 (92%)	202 (94%)	598 (92%)
Discontinued Week 14 Treatment	19 (9%)	18 (8%)	13 (6%)	50 (8%)
Adverse event	5 (2%)	6 (3%)	6 (3%)	17 (3%)
Lack of efficacy	4 (2%)	1 (0%)	-	5 (1%)
Lost to follow-up	-	4 (2%)	2 (1%)	6 (1%)
Other	3 (1%)	-	-	3 (0%)
Withdrew consent	7 (3%)	7 (3%)	5 (2%)	19 (3%)

Counts and percentages relative to N in parenthesis.
Screened patients were randomized to MTX monotherapy, or upadacitinib 15 mg QD, or upadacitinib 30 mg QD.
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate
[Source: Statistical Reviewer]

In study M13-545, 94% and 91% of the randomized patients completed Week 12 and Week 24 study visits respectively regardless of adherence to randomized treatment (Table 11). The study discontinuation rates were similar across arms at Week 12 and Week 24. The most common reason for study discontinuation across arms was withdrawal of informed consent. A total of 89% of the randomized patients completed randomized treatment at Week 24. The treatment discontinuation rates were similar across arms with adverse events and patients initiating withdrawal from study drug being the most common reason for treatment discontinuation.

In study M13-545, 5% of the randomized patients were rescued prior to Week 24, with majority of them rescued at Week 16 (Table 52). At Week 16, 6% of the patients on MTX monotherapy were rescued with upadacitinib, 3% of the patients on upadacitinib 15 mg QD monotherapy were rescued with MTX, and 1% of the patients on upadacitinib 30 mg QD monotherapy were rescued with MTX.

Table 11 Patient Disposition at Various Weeks, Study M13-545

	MTX (N=314)	UPA 15 mg QD (N=317)	UPA 30 mg QD (N=314)	All (N=945)
Completed Study at Week 12	288 (91%)	302 (95%)	299 (95%)	889 (94%)
Discontinued Study prior to Week 12	26 (8%)	15 (5%)	15 (5%)	56 (6%)
Adverse event	4 (1%)	7 (2%)	3 (1%)	14 (1%)
Lack of efficacy	5 (2%)	-	2 (1%)	7 (1%)
Lost to follow-up	2 (1%)	2 (1%)	1 (0%)	5 (1%)
Other	4 (1%)	1 (0%)	2 (1%)	7 (1%)
Withdrew consent	11 (3%)	5 (2%)	7 (2%)	23 (2%)
Completed Study at Week 24	277 (88%)	293 (92%)	290 (92%)	860 (91%)
Discontinued Study prior to Week 24	37 (12%)	24 (8%)	23 (7%)	84 (9%)
Ongoing	-	-	1 (0%)	1 (0%)
Adverse event	8 (3%)	10 (3%)	5 (2%)	23 (2%)
Lack of efficacy	7 (2%)	1 (0%)	2 (1%)	10 (1%)
Lost to follow-up	3 (1%)	4 (1%)	2 (1%)	9 (1%)
Other	4 (1%)	1 (0%)	2 (1%)	7 (1%)
Withdrew consent	14 (4%)	8 (3%)	12 (4%)	34 (4%)
Missing Status	1 (0%)	-	-	1 (0%)
Completed Treatment at Week 12	284 (90%)	301 (95%)	297 (94%)	882 (93%)
Discontinued Treatment prior to Week 12	30 (10%)	16 (5%)	17 (5%)	63 (7%)
Adverse event	5 (2%)	7 (2%)	6 (2%)	18 (2%)
Lack of efficacy	7 (2%)	-	3 (1%)	10 (1%)
Lost to follow-up	2 (1%)	2 (1%)	1 (0%)	5 (1%)
Other[2]	5 (2%)	2 (1%)	2 (1%)	9 (1%)
Withdrew consent	11 (3%)	5 (2%)	5 (2%)	21 (2%)
Completed Treatment at Week 24	268 (85%)	290 (91%)	282 (90%)	840 (89%)
Discontinued Treatment prior to Week 24	46 (15%)	27 (9%)	31 (10%)	105 (11%)
Ongoing	-	-	1 (0%)	1 (0%)
Adverse event	13 (4%)	13 (4%)	12 (4%)	38 (4%)
Lack of efficacy	10 (3%)	-	3 (1%)	13 (1%)
Lost to follow-up	3 (1%)	4 (1%)	2 (1%)	9 (1%)
Other[2]	6 (2%)	2 (1%)	3 (1%)	11 (1%)
Withdrew consent	15 (5%)	8 (3%)	11 (3%)	34 (4%)

Counts and percentages relative to N in parenthesis.
Screened patients were randomized to MTX monotherapy, or upadacitinib 15 mg QD monotherapy, or upadacitinib 30 mg QD monotherapy.
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate
[Source: Statistical Reviewer]

3.2.5.2 Demographics and Baseline Characteristics

Majority of the patients were female, white (82% - 88%), were on average at least 54 years old (Table 12). Approximately 20% of the patients were at least 65 years of age. Studies M13-545 and M14-465 randomized the highest percentage of Hispanics (at least 32%). The distribution of baseline anthropometric variables was similar across studies and balanced within studies.

Table 12 Baseline Demographics for all Studies

		M13-542 (N=498)	M13-549 (N=661)	M14-465 (N=1629)	M15-555 (N=648)	M13-545 (N=945)
Age	*Mean (SD)*	57 (11.4)	56 (11.7)	54 (12.1)	54 (12.1)	53 (12.7)
	Min – Max	23 – 87	21 – 86	19 – 86	20 – 80	18 – 85
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
Age > 65	*Counts (%)*	135 (27%)	152 (23%)	310 (19%)	130 (20%)	179 (19%)
Sex: Female	*Counts (%)*	418 (84%)	520 (79%)	1292 (79%)	523 (81%)	721 (76%)
Sex: Male	*Counts (%)*	80 (16%)	141 (21%)	337 (21%)	125 (19%)	224 (24%)
Race: *White*	*Counts (%)*	433 (87%)	561 (85%)	1429 (88%)	529 (82%)	766 (81%)
Race: *Black*	*Counts (%)*	48 (10%)	31 (5%)	88 (5%)	35 (5%)	33 (3%)
Race: *Asian*	*Counts (%)*	9 (2%)	59 (9%)	85 (5%)	69 (11%)	106 (11%)
Race: *Other*	*Counts (%)*	0 (0%)	8 (1%)	22 (1%)	5 (1%)	17 (2%)
Ethnic: *Hispanic*		86 (17%)	80 (12%)	527 (32%)	156 (24%)	316 (33%)
Ethnic: *Not Hispanic*		412 (83%)	581 (88%)	1102 (68%)	492 (76%)	629 (67%)
Region: *North America*	*Counts (%)*	328 (66%)	267 (40%)	303 (19%)	192 (30%)	140 (15%)
US only	*Counts (%)*	320 (64%)	255 (39%)	288 (18%)	189 (29%)	124 (13%)
Region: *South America*	*Counts (%)*	-	29 (4%)	432 (27%)	91 (14%)	272 (29%)
Region: *Eastern Europe*	*Counts (%)*	67 (13%)	223 (34%)	656 (40%)	239 (37%)	259 (27%)
Region: *Western Europe*	*Counts (%)*	97 (19%)	69 (10%)	89 (5%)	24 (4%)	109 (12%)
Region: *Asia*	*Counts (%)*	1 (0%)	48 (7%)	52 (3%)	65 (10%)	95 (10%)
Region: *Other*	*Counts (%)*	5 (1%)	25 (4%)	97 (6%)	37 (6%)	70 (7%)
Height (cm)	*Mean (SD)*	164 (9.0)	165 (9.2)	163 (9.4)	163 (9.2)	163 (9.7)
	Min – Max	142 – 193	140 – 196	139 – 197	136 – 198	135 – 197
	n	(n=495)	(n=661)	(n=1625)	(n=640)	(n=945)
Weight (kg)	*Mean (SD)*	81 (19.8)	81.4 (21.8)	76.7 (19.1)	76.1 (19.7)	74.8 (19.6)
	Min – Max	38 – 148	36 – 172	39 – 169	39 – 167	37 – 196
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
BMI (kg/m^2)	*Mean (SD)*	30.2 (7.0)	29.7 (7.2)	28.9 (6.6)	28.6 (6.7)	28.2 (6.6)
	Min – Max	15 – 64	17 – 58	15 – 68	16 – 58	13 – 72
	n	(n=495)	(n=661)	(n=1625)	(n=640)	(n=945)

Counts (%): Counts and percentages relative to N in parenthesis.
Abbreviations: N=total patients randomized who taken at least one dose of drug; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; BMI=body mass index

[Source: Statistical Reviewer]

In summary, the patient's RA symptoms were consistent with the patient population each study enrolled (Table 13). Patients in study M13-542 averaged longer duration of RA symptoms than patients in other studies. Majority of the patients in study M13-542, M13-549 and M14-465 received background MTX.

Table 13 Disease-Related Baseline Characteristics for all Studies

		M13-542 (N=498)	M13-549 (N=661)	M14-465 (N=1629)	M15-555 (N=648)	M13-545 (N=945)
Duration of RA symptoms	*Mean (SD)*	14.6 (9.5)	8.9 (8.5)	9.9 (8.4)	8.6 (8.4)	4.4 (6.3)
	Min – Max	1 – 52	0 – 60	0 – 52	0 – 57	0 – 46
	n	(N=498)	(N=661)	(N=1620)	(N=643)	(N=945)
RF+	*Counts (%)*	345 (69%)	473 (72%)	1303 (80%)	457 (71%)	717 (76%)
anti-CCP+	*Counts (%)*	356 (71%)	496 (75%)	1318 (81%)	463 (71%)	724 (77%)
At least 1 RF+ or anti-CCP+	*Counts (%)*	391 (79%)	529 (80%)	1425 (87%)	512 (79%)	833 (83%)

Counts (%): Counts and percentages relative to N in parenthesis.
Abbreviations: N= total patients randomized who taken at least one dose of drug; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; RF=rheumatoid factor; anti-CCP = anti-cyclic citrullinated peptide; RA=rheumatoid arthritis
[Source: Statistical Reviewer]

Baseline ACR-related characteristics were on average similar across all studies except for CRP (Table 14). The average baseline CRP was higher in studies M13-545 and M14-465, consistent with inclusion criteria of the study protocols, compared to the other studies.

Table 14 Baseline Efficacy Related Components for all Studies

		M13-542 (N=498)	M13-549 (N=661)	M14-465 (N=1629)	M15-555 (N=648)	M13-545 (N=945)
SJC (Out of 28)	*Mean (SD)*	11.6 (5.7)	11.0 (5.3)	11.6 (5.5)	11.3 (5.8)	11.7 (5.8)
	Min -Max	0 - 28	0 - 28	0 - 28	0 - 28	0 - 28
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
SJC (Out of 66)	*Mean (SD)*	16.8 (10.6)	15.8 (9.9)	16.4 (9.6)	16.7 (10.9)	16.5 (10.2)
	Min -Max	4 - 64	1 - 62	1 - 64	3 - 66	0 - 60
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
TJC (Out of 28)	*Mean (SD)*	15.8 (7.0)	14.5 (6.7)	15.2 (6.9)	14.3 (7.2)	15.0 (7.1)
	Min -Max	1 - 28	0 - 28	0 - 28	0 - 28	1 - 28
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
TJC (Out of 68)	*Mean (SD)*	27.9 (15.6)	25.4 (14.3)	26.2 (14.8)	24.8 (15.4)	25.7 (15.2)
	Min -Max	6 - 68	2 - 68	5 - 68	3 - 68	2 - 68
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)
HAQ-DI	*Mean (SD)*	1.6 (0.6)	1.5 (0.6)	1.6 (0.6)	1.5 (0.7)	1.6 (0.7)
	Min -Max	0 - 3	0 - 3	0 - 3	0 - 3	0 - 3
	n	(n=490)	(n=656)	(n=1620)	(n=647)	(n=942)
Physician Global Assessment	*Mean (SD)*	67.3 (16.4)	63.9 (17.3)	65.7 (17.6)	63.5 (18.0)	67.0 (16.7)
	Min -Max	13 - 100	8 - 100	4 - 100	17 - 100	8 - 100
	n	(n=475)	(n=633)	(n=1541)	(n=611)	(n=904)
Patient Global Assessment	*Mean (SD)*	66.1 (21.2)	62.1 (20.9)	64.4 (21.5)	60.4 (22.3)	65.8 (21.7)
	Min -Max	9 - 100	0 - 100	0 - 100	0 - 100	0 - 100
	n	(n=492)	(n=657)	(n=1620)	(n=647)	(n=942)
Patient Pain VAS	*Mean (SD)*	67.5 (20.5)	63.2 (20.0)	65.5 (20.8)	62.3 (21.9)	66.5 (21.2)
	Min -Max	4 - 100	0 - 100	0 - 100	4 - 100	0 - 100
	n	(n=490)	(n=657)	(n=1621)	(n=647)	(n=942)
CRP (mg/L)	*Mean (SD)*	16.2 (20.3)	14.7 (16.9)	18.3 (21.9)	14.9 (18.3)	21.2 (24.2)
	Min -Max	0.2 - 150.0	0.2 - 122.0	0.2 - 198.0	0.2 - 136.0	0.2 - 207.0
	n	(n=498)	(n=661)	(n=1629)	(n=648)	(n=945)

Abbreviations: N= total patients randomized who taken at least one dose of drug; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP=C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI=Health Assessment Questionnaire-Disability Index;
[Source: Statistical Reviewer]

Baseline SF-36 PCS and MCS, duration of morning stiffness, and severity of morning stiffness were on average similar across studies (Table 15). In studies M13-549, M14-465, and M13-545, the average FACIT-F scores were similar.

Table 15 Baseline PRO Endpoints for all Studies

		M13-542 (N=498)	M13-549 (N=661)	M14-465 (N=1629)	M15-555 (N=648)	M13-545 (N=945)
SF-36 PCS	*Mean (SD)*	31.3 (7.4)	33.0 (7.6)	32.5 (7.0)	33.5 (7.7)	33.2 (7.5)
	Min -Max	9.9 - 56.7	11.6 - 58.2	9.2 - 58.6	13.2 - 60.4	12.3 - 60.9
	n	(n=491)	(n=657)	(n=1622)	(n=647)	(n=940)
SF-36 MCS	*Mean (SD)*	45.2 (12.2)	46.2 (11.5)	42.9 (10.8)	44.6 (11.3)	43.0 (11.0)
	Min -Max	10.5 - 68.8	11.7 - 78.4	13.9 - 73.7	14.6 - 72.4	13.5 - 70.9
	n	(n=491)	(n=657)	(n=1622)	(n=647)	(n=940)
Duration of Morning Stiffness	*Mean (SD)*	154.4 (223.2)	140.0 (207.1)	142.8 (180.0)	143.7 (198.9)	144.7 (181.1)
	Min -Max	0 - 1440	0 - 1440	0 - 1440	0 - 1830	0 - 1440
	n	(n=498)	(n=648)	(n=1624)	(n=646)	(n=942)
Severity of Morning Stiffness	*Mean (SD)*	6.7 (2.2)	6.1 (2.2)	6.3 (2.3)	6.0 (2.3)	6.5 (2.3)
	Min -Max	0 - 10	0 - 10	0 - 10	0 - 10	0 - 10
	n	(n=498)	(n=647)	(n=1623)	(n=646)	(n=942)
FACIT-F	*Mean (SD)*	-	28.0 (11.7)	26.8 (11.1)	-	26.9 (11.6)
	Min -Max	-	0 - 51	0 - 52	-	0 - 52
	n	-	(n=654)	(n=1615)	-	(n=940)

FACIT-F was not collected in study M13-542 and M15-555.
Abbreviations: N= total patients randomized who taken at least one dose of drug; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; SF-36=short form 36-item survey; PCS=physical component score; MCS=mental component score; FACIT-F=functional assessment of chronic illness therapy-fatigue; PRO=patient reported outcomes
[Source: Statistical Reviewer]

In general, the distribution of baseline radiograph scores was similar within each study but were different across studies (Table 16). Patients in the cDMARD add-on study averaged higher mTSS, scores than patients in the monotherapy study. In addition, there was a lower proportion of patients who had a mTSS score of 0 in M14-465 than those in M13-545.

Table 16 Baseline Radiograph Scores and Components, Studies M14-465 and M13-545

		M14-465 (N=1629)	M13-545 (N=945)
mTSS	*Mean (SD)*	34.9 (50.1)	16.2 (35.9)
	Min -Max	0.0 - 324.5	0.0 - 260.5
	n	(n=1619)	(n=927)
Erosion Score	*Mean (SD)*	16.5 (26.2)	7.6 (18.0)
	Min -Max	0.0 - 192.0	0.0 - 147.0
	n	(n=1619)	(n=927)
Joint Space Narrowing	*Mean (SD)*	18.4 (25.7)	8.7 (19.0)
	Min -Max	0.0 - 136.5	0.0 - 118.0
	n	(n=1619)	(n=927)
Proportion with mTSS = 0	*Count (%)*	213 (13%)	311 (33%)

Counts (%): Counts and percentages relative to N in parenthesis.
Abbreviations: N= total patients randomized who taken at least one dose of drug; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; mTSS=modified total sharp score
[Source: Statistical Reviewer]

3.2.6 Results and Conclusions

3.2.6.1 Primary Endpoint

In studies M13-542 and M13-549, there were statistically significantly higher probabilities of ACR20 response comparing each dose of upadacitinib with placebo (Table 17). There were statistically significantly higher probabilities of ACR20 response comparing upadacitinib 15 mg QD with placebo in study M14-465. There were statistically significantly higher probabilities of ACR20 and ACR50 response respectively for monotherapy studies M15-555 and M13-545 comparing each dose of upadacitinib monotherapy with MTX monotherapy.

The estimated probabilities of ACR20 response at the primary timepoint were between 28 - 36% higher on the absolute scale comparing upadacitinib 15 mg QD to placebo in MTX-add on studies.

The estimated probabilities of ACR20 response at the primary timepoint was 26.5% (95% CI: 17.5%, 35.6%) significantly higher on the absolute scale comparing upadacitinib 15 mg QD monotherapy to MTX monotherapy in study M15-555. In study M13-545, the estimated probabilities of ACR50 response at the primary timepoint was 23.7% (95% CI: 16.0%, 35.6%) significantly higher on the absolute scale comparing upadacitinib 15 mg QD monotherapy to MTX monotherapy.

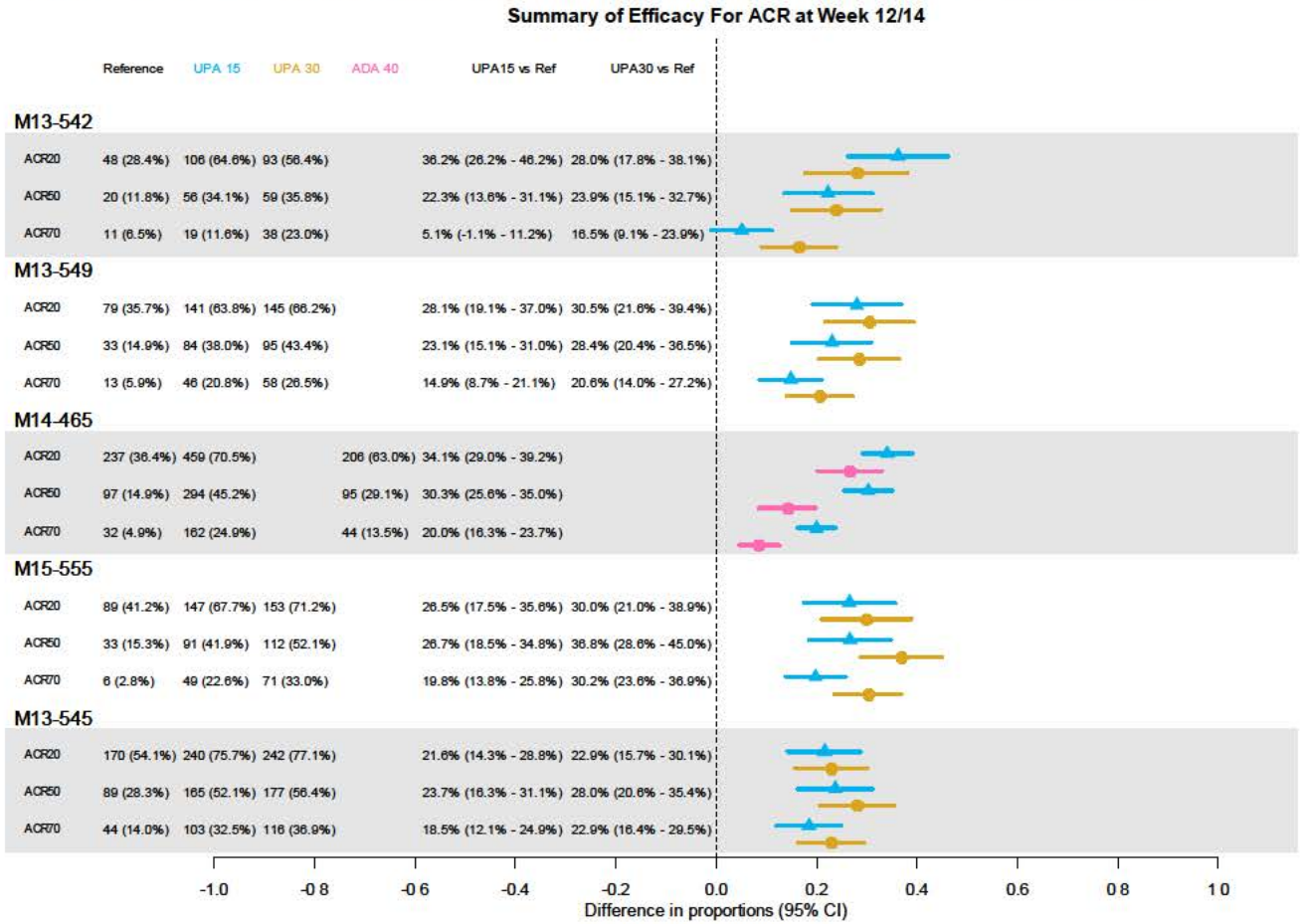
Table 17 Summary of Primary Efficacy Endpoint Results at Primary Timepoint in all Studies

Treatment Arm	N	Count (%)[1]	Diff (%) [95 % CI]; P-value[2]
cDMARD Add-on Studies			
M13-542			
Placebo	169	48 (28%)	
UPA 15 mg QD	164	106 (65%)	36.2% (26.2% - 46.2%); <0.001
UPA 30 mg QD	165	93 (56%)	28.0% (17.8% - 38.1%); <0.001
M13-549			
Placebo	221	79 (36%)	
UPA 15 mg QD	221	141 (64%)	28.1% (19.1% - 37.0%); <0.001
UPA 30 mg QD	219	145 (66%)	30.5% (21.6% - 39.4%); <0.001
M14-465			
Placebo	651	237 (36%)	
UPA 15 mg QD	651	459 (71%)	34.1% (29.0% - 39.2%); <0.001
ADA 40 mg EOW	327	206 (63%)	26.6% (20.2% - 33.0%); <0.001
MTX Monotherapy Studies			
M15-555[3]			
MTX	216	89 (41%)	
UPA 15 mg QD	217	147 (68%)	26.5% (17.5% - 35.6%); <0.001
UPA 30 mg QD	215	153 (71%)	30.0% (21.0% - 38.9%); <0.001
M13-545 (ACR50)			
MTX	314	89 (28%)	
UPA 15 mg QD	317	165 (52%)	23.7% (16.3% - 31.1%); <0.001
UPA 30 mg QD	314	177 (56%)	28.0% (20.6% - 35.4%); <0.001

Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.
1: Counts and percentages relative to N in parenthesis are reported for the probability of ACR20 response.
2: Difference in the probability of ACR response, respective 95% CI using normal approximation to difference in binomial proportions (vs placebo or MTX), p-values from the Cochran Mantel Haenszel test were reported.
3: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
Abbreviations: ACR=American College of Rheumatology; MTX=methotrexate; UPA=upadacitinib; QD=once daily; CI=confidence intervals; EOW=every other week; ADA=adalimumab; Diff=difference; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In studies M13-549, M15-555, and M13-545, there was a numerical trend towards greater ACR20/50/70 response probability for upadacitinib 30 mg arm relative to upadacitinib 15 mg arm (Figure 1). In study M13-542, this numerical dose-response trend was not observed for ACR20 response probability but was observed for ACR50/70 response probability.

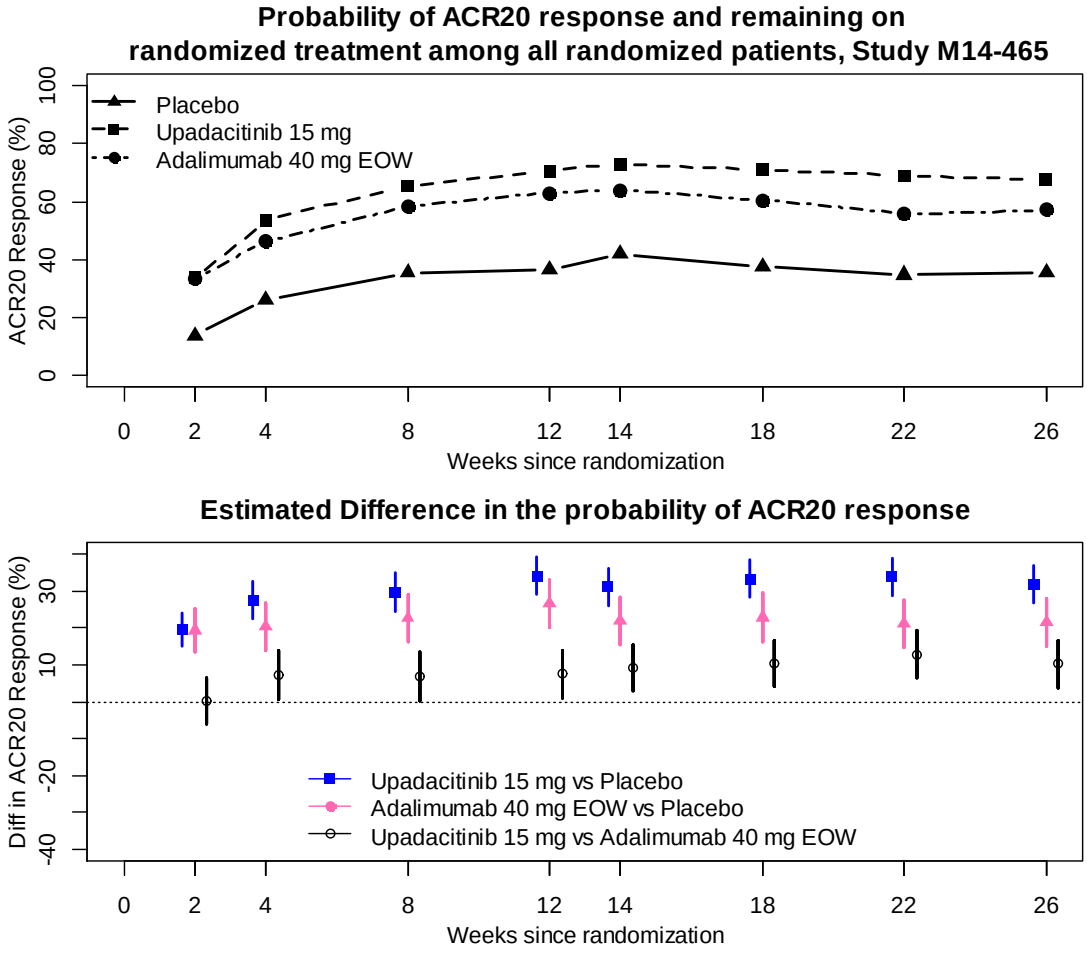
Figure 1 ACR20/50/70 Efficacy for the Doses of Upadacitinib with Reference for all Studies



Counts (%): Counts and percentages relative to N in parenthesis are reported for the probability of ACR20 response.
Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.
Point estimate for the difference in probability of ACR response, respective 95% CI using normal approximation to difference in binomial proportions are presented. The reference arm for studies M13-542, M13-549, M14-465 is placebo. The reference arm for studies M15-555 and M13-545 is MTX monotherapy.
ACR results for all studies except M15-555 were evaluated at Week 12. In study M15-555, ACR was evaluated at Week 14.
Abbreviations: MTX=methotrexate; ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; ADA=adalimumab
[Source: Statistical Reviewer]

Trends of efficacy over time were summarized for studies M14-465 and M13-545. In Study M14-465, there was greater probability in ACR20 response and remaining on randomized treatment comparing randomized upadacitinib 15 mg QD patients relative to randomized placebo patients across all study visits up to Week 14 (Figure 2). In addition, there was greater probability in ACR20 response and remaining on randomized treatment comparing randomized upadacitinib 15 mg QD patients over randomized adalimumab patients after Week 4 study visit through Week 14 (Table 53). Interpretation of the results in the figure after Week 14 is affected by differential discontinuation of randomized treatment due to escape across all treatment arms after Week 14 (Table 8).

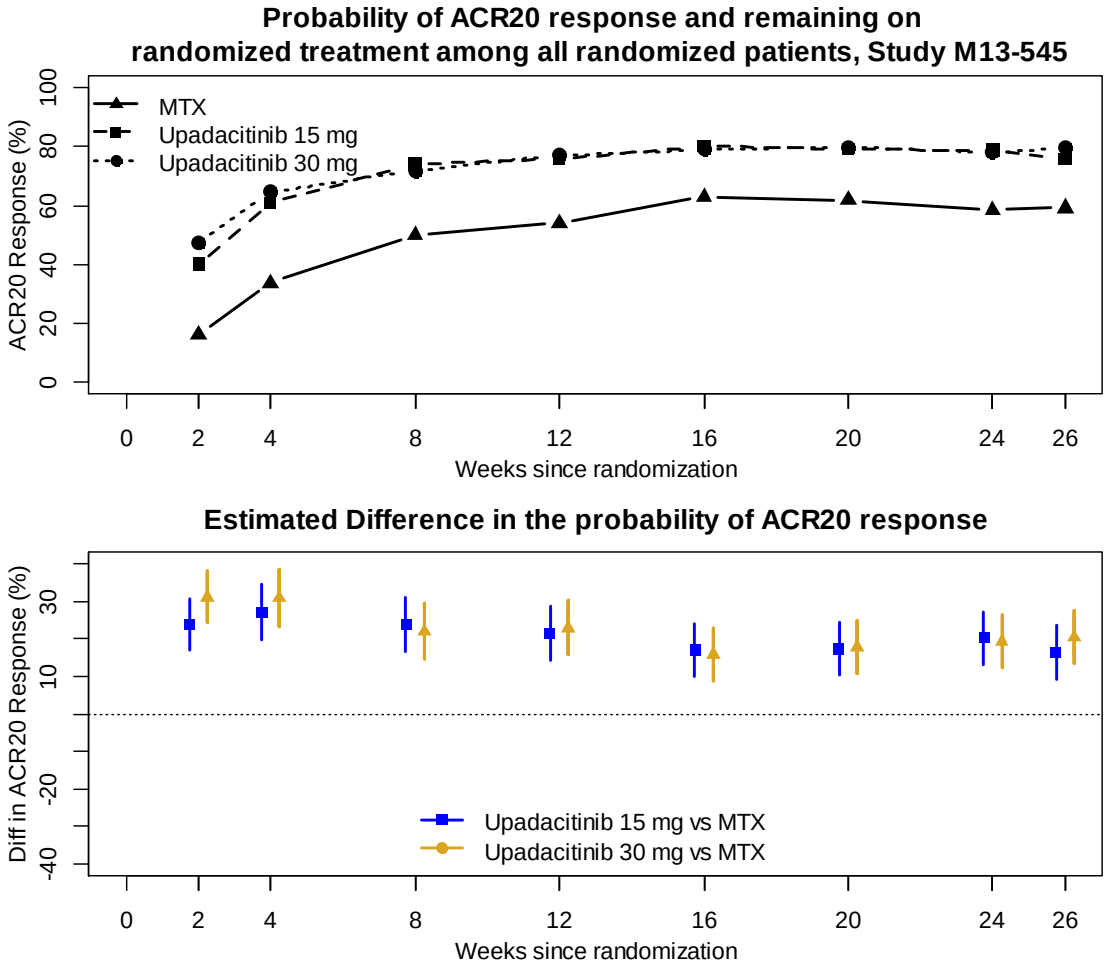
Figure 2 Probability of ACR20 Response over Time, Study M14-465



At or prior to Week 14, patients who discontinued randomized treatment, or were lost-to-follow u, or had missing ACR components to allow derivation of ACR20 response were imputed as non-response.
After Week 14 or 18 or 22, patients who remained on randomized treatment, did not have 20% improvement in both SJC and TJC at Week 14 or 18 or 22 respectively, had not been previously rescued, and were rescued to another treatment arm from Week 14 or 18 or 22 respectively were further imputed as non-response.
Top: Trends in the probability of ACR20 response and remaining in the study and on randomized treatment among all randomized patients across visit weeks.
Bottom: Estimated difference in the probability of ACR20 (represented by the solid squares, solid triangles, and open circles) with respective 95% CI (represented by the vertical lines through the points) comparing the three arms.
Abbreviations: ACR=American College of Rheumatology; EOW=every other week
[Source: Statistical Reviewer]

In Study M13-545, there was a greater ACR20 response for both upadacitinib monotherapy dosing regimens relative to MTX monotherapy across all visit weeks based on the proportion of randomized patients who were ACR20 responders and remained on randomized treatment and in the study during the MTX-controlled period (Figure 3,Table 54). There were no observed numerical differences between the two doses across the study visits. The interpretation of the results in the figure after Week 12 is difficult because both MTX and upadacitinib patients could be rescued by modifying background medications (Table 52).

Figure 3 Probability of ACR20 Response over Time, Study M13-545



Prior to Week 12, patients who discontinued randomized treatment, or were lost-to-follow up, or had missing ACR components to allow derivation of ACR20 response were imputed as non-response.
At and after Week 12, patients who did not have 20% improvement in both SJC and TJC at two consecutive visits and were rescued were imputed as non-response.
Top: Trends in the probability of ACR20 response and remaining in the study and on randomized treatment among all randomized patients across visit weeks.
Bottom: Estimated difference in the probability of ACR20 comparing the doses of upadacitinib vs MTX
Abbreviations: ACR=American College of Rheumatology; MTX=methotrexate
[Source: Statistical Reviewer]

There were consistent significant trends of improvements observed at the primary timepoint for the individual components of ACR regardless of adherence to study drug comparing upadacitinib

15 mg with placebo in cDMARD add-on studies (Table 47, Table 48, and Table 49). Consistent significant trends of improvements were observed for the individual components of ACR regardless of adherence to study drug comparing upadacitinib 15 mg monotherapy with MTX monotherapy in monotherapy studies (Table 50, Table 51). There were consistent trends of improvements in the components of ACR comparing upadacitinib 30 mg QD with placebo in cDMARD add-on studies M13-542 and M13-549. Such significant trends were observed in the monotherapy studies. There was a lack of consistent numerical dose-response trends of improvements in all ACR components in all studies evaluating both the 15 mg and 30 mg doses.

The tipping point analysis for the primary efficacy endpoint in each study were robust to missing data assumptions (Figure 13, Figure 14, Figure 15, Figure 16, Figure 17). Tipping point analysis conducted for other observed ACR thresholds were also robust to missing data assumptions.

In summary, the various sensitivity and tipping point analyses conducted were generally supportive of the applicant's primary analysis of ACR20 at the primary efficacy time points for the proposed upadacitinib 15 mg QD in the placebo-controlled studies. In the monotherapy studies, similar conclusions were observed. In studies where both doses of upadacitinib were investigated, there were numerical trends towards higher ACR response except for study M13-542 where such trend was not consistent for ACR20. There were significant trends of improvement across study visits for ACR20 in studies M13-545 and M14-465 up to Week 24/26. Significant improvements in all components of ACR were observed for all studies comparing upadacitinib 15 mg QD with the reference.

3.2.6.2 HAQ-DI

HAQ-DI is one of the key patients reported outcomes commonly used in RA studies. In the cDMARD add-on studies, there were statistically significant reduction in the mean changes from baseline in HAQ-DI, indicating improvements in signs and symptoms, at primary timepoint comparing upadacitinib 15mg QD with placebo (Table 18). In the monotherapy studies, there were also statistically significant reduction in the mean changes from baseline in HAQ-DI at the primary timepoint for monotherapy studies M15-555 and M13-545 comparing upadacitinib 15 mg QD monotherapy with MTX monotherapy at the primary time-point of assessment.

Table 18 Change from Baseline in HAQ-DI at Primary Timepoint for all Studies

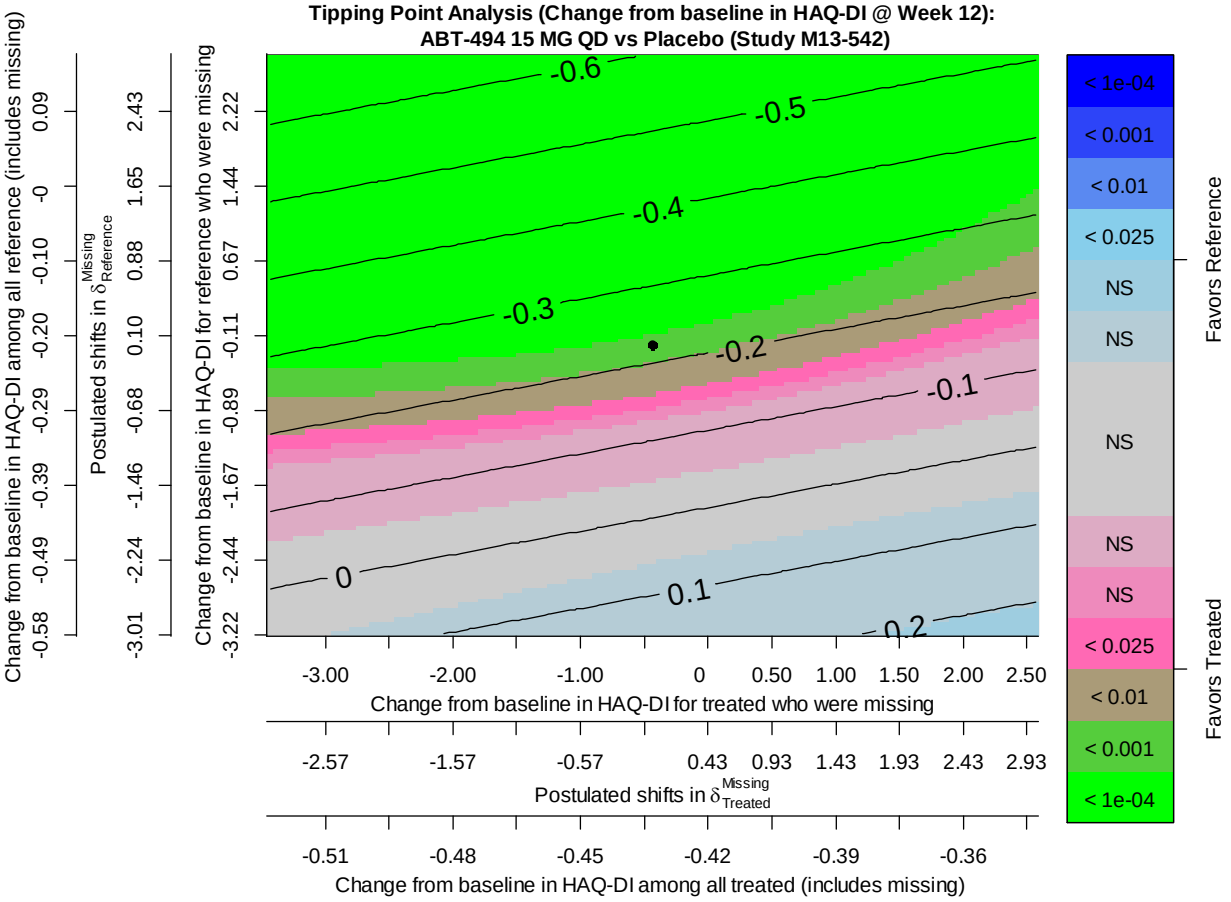
Treatment Arm	N	Baseline Mean (SD)	n	Visit [1] Mean (SD)	Adj Mean Diff [2] (95% CI)	P-value [2]
cDMARD Add-on studies						
M13-542						
Placebo	166	1.6 (0.6)	150	1.3 (0.7)		
UPA 15 mg QD	163	1.7 (0.6)	160	1.2 (0.8)	-0.2 (-0.3, -0.1)	0.002
UPA 30 mg QD	161	1.6 (0.6)	154	1.2 (0.7)	-0.2 (-0.3, -0.1)	<0.001
M13-549						
Placebo	221	1.4 (0.6)	206	1.1 (0.7)		
UPA 15 mg QD	216	1.5 (0.6)	210	0.9 (0.7)	-0.3 (-0.4, -0.2)	<0.001
UPA 30 mg QD	219	1.5 (0.6)	200	0.9 (0.7)	-0.3 (-0.4, -0.2)	<0.001
M14-465						
Placebo	650	1.6 (0.6)	617	1.3 (0.7)		
UPA 15 mg QD	645	1.6 (0.6)	617	1.0 (0.7)	-0.3 (-0.4, -0.3)	<0.001
ADA 40 mg EOW	325	1.6 (0.6)	309	1.1 (0.7)	-0.2 (-0.3, -0.1)	<0.001
Monotherapy studies						
M15-555						
MTX	216	1.5 (0.7)	195	1.2 (0.7)		
UPA 15 mg QD	216	1.5 (0.7)	199	0.8 (0.7)	-0.3 (-0.5, -0.2)	<0.001
UPA 30 mg QD	215	1.5 (0.7)	201	0.8 (0.7)	-0.4 (-0.5, -0.3)	<0.001
M13-545						
MTX	314	1.6 (0.7)	278	1.1 (0.7)		
UPA 15 mg QD	317	1.6 (0.7)	302	0.8 (0.7)	-0.3 (-0.4, -0.2)	<0.001
UPA 30 mg QD	311	1.5 (0.7)	298	0.7 (0.7)	-0.3 (-0.4, -0.2)	<0.001

1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in HAQ-DI adjusting for treatment groups, baseline HAQ-DI, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: N=total randomized with evaluable baseline; n=total observed at visit; SD=standard deviation; CI=confidence interval; UPA=upadacitinib; QD=once daily; MTX=methotrexate; EOW=every other week; ADA=adalimumab; Adj=adjusted; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

Tipping point analyses were conducted to evaluate the robustness of the results for the change from baseline in HAQ-DI comparing upadacitinib 15 mg QD with reference for the placebo-controlled studies and the MTX-controlled studies. In study M13-542, placebo patients with missing HAQ-DI response at Week 12 would have to average at least 1 point improvement (i.e., negative change from baseline) from baseline, together with missing upadacitinib patients to average 0.5 point worsening (i.e., positive change from baseline) from baseline in order to tip the conclusions, such that there would be no longer evidence of an effect (Figure 4). The tipping

point analysis in other studies were more extreme and less likely plausible (Figure 18, Figure 19, Figure 20, Figure 21).

Figure 4 Tipping Point Sensitivity Analysis for the Change from Baseline in HAQ-DI at Week 12 Comparing Upadacitinib (ABT-494) 15 mg QD vs Placebo, M13-542



Abbreviations: HAQ-DI=health assessment questionnaire-disability index; QD=once daily; NS=non-significant at a conservative 2-sided level alpha of 0.025 based on the first alpha split for the graphical testing procedure.
[Source: Statistical Reviewer]

In summary, the various sensitivity and tipping point analyses conducted were generally supportive of the applicant's finding for HAQ-DI that there was evidence of improvement in physical function at the primary efficacy time points for the proposed upadacitinib 15 mg QD in all the studies.

3.2.6.3 DAS28(CRP)

In the cDMARD add-on studies, there were statistically significant reduction in the mean change from baseline in DAS28(CRP), indicating improvements in signs and symptoms, at the primary timepoint of assessment comparing upadacitinib 15mg QD with placebo (Table 19). In the monotherapy studies, there were statistically significant reduction in the mean change from

baseline in DAS28(CRP) at the primary timepoint of assessment comparing upadacitinib 15 mg QD monotherapy with MTX monotherapy.

Table 19 Change from Baseline in DAS28(CRP) for all Studies

Treatment Arm	N	Baseline Mean (SD)	n	Visit[1] Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
cDMARD Add-on Studies						
M13-542						
Placebo	166	5.8 (1.0)	147	4.7 (1.4)		
UPA 15 mg QD	163	5.9 (0.9)	157	3.5 (1.3)	-1.2 (-1.5, -0.9)	<0.001
UPA 30 mg QD	163	5.8 (0.9)	149	3.5 (1.5)	-1.1 (-1.4, -0.8)	<0.001
M13-549						
Placebo	221	5.6 (0.8)	206	4.5 (1.5)		
UPA 15 mg QD	217	5.7 (1.0)	206	3.4 (1.4)	-1.2 (-1.4, -0.9)	<0.001
UPA 30 mg QD	219	5.7 (0.9)	200	3.3 (1.2)	-1.3 (-1.6, -1.1)	<0.001
M14-465						
Placebo	649	5.8 (0.9)	595	4.7 (1.4)		
UPA 15 mg QD	647	5.8 (1.0)	586	3.3 (1.3)	-1.3 (-1.5, -1.2)	<0.001
ADA 40 mg EOW	324	5.9 (1.0)	295	3.8 (1.4)	-0.9 (-1.0, -0.7)	<0.001
Monotherapy Studies						
M15-555						
MTX	216	5.6 (1.0)	194	4.4 (1.4)		
UPA 15 mg QD	216	5.6 (0.9)	195	3.3 (1.4)	-1.1 (-1.3, -0.9)	<0.001
UPA 30 mg QD	215	5.6 (1.1)	198	3.0 (1.3)	-1.4 (-1.7, -1.2)	<0.001
M13-545						
MTX	314	5.9 (1.0)	290	4.0 (1.4)		
UPA 15 mg QD	317	5.9 (1.0)	303	3.2 (1.4)	-0.9 (-1.1, -0.7)	<0.001
UPA 30 mg QD	311	5.8 (1.0)	297	3.0 (1.3)	-1.0 (-1.2, -0.7)	<0.001

The mean and standard deviation in parenthesis of the observed data were reported.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in DAS28(CRP) adjusting for treatment groups, baseline DAS28(CRP), and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: N=total randomized with evaluable baseline; n=total observed at visit; SD=standard deviation; CI=confidence interval; UPA=upadacitinib; QD=once daily; MTX=methotrexate; EOW=every other week; ADA=adalimumab; Adj=adjusted; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

The results based on the analyses of DAS28(CRP)<2.6 or DAS28(CRP) < 3.3 supported the efficacy of upadacitinib as a potential treatment for RA (Table 20).

Table 20 Proportion of Randomized Patients with DAS28(CRP) < 2.6 or DAS28(CRP) < 3.3 and Remained on Randomized Treatment through Primary Time Point for all Studies

Treatment Arm	Total	DAS28(CRP) <2.6		DAS28(CRP)<3.3	
		Counts (%)	Diff (%) [95 % CI]	Counts (%)	Diff (%) [95 % CI]
cDMARD Add-on Studies					
M13-542					
Placebo	169	16 (9%)		24 (14%)	
UPA 15 mg QD	164	47 (29%)	19.2% (11.0%, 27.4%)	71 (43%)	29.1% (19.9%, 38.3%)
UPA 30 mg QD	165	39 (24%)	14.2% (6.3%, 22.0%)	70 (42%)	28.2% (19.0%, 37.4%)
M13-549					
Placebo	221	22 (10%)		38 (17%)	
UPA 15 mg QD	221	68 (31%)	20.8% (13.6%, 28.1%)	107 (48%)	31.2% (23.0%, 39.5%)
UPA 30 mg QD	219	62 (28%)	18.4% (11.2%, 25.5%)	105 (48%)	30.8% (22.5%, 39.0%)
M14-465					
Placebo	651	40 (6%)		90 (14%)	
UPA 15 mg QD	651	187 (29%)	11.9% (7.3%, 16.5%)	293 (45%)	14.9% (9.3%, 20.5%)
ADA 40 mg EOW	327	59 (18%)	22.6% (18.6%, 26.5%)	94 (29%)	31.2% (26.5%, 35.8%)
Monotherapy Studies					
M15-555					
MTX	216	18 (8%)		42 (19%)	
UPA 15 mg QD	217	61 (28%)	19.8% (12.8%, 26.8%)	97 (45%)	25.3% (16.8%, 33.7%)
UPA 30 mg QD	215	87 (40%)	32.1% (24.6%, 39.7%)	114 (53%)	33.6% (25.1%, 42.1%)
M13-545					
MTX	314	42 (13%)		88 (28%)	
UPA 15 mg QD	317	113 (36%)	22.3% (15.8%, 28.7%)	169 (53%)	25.3% (17.9%, 32.7%)
UPA 30 mg QD	314	128 (41%)	27.4% (20.8%, 34.0%)	172 (55%)	26.8% (19.3%, 34.2%)

NRI was used to impute patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study.
Counts (%): Counts and percentages relative to N in parenthesis are reported for the probability of response.
Diff (95% CI); p-value: Difference in the probability of response, respective 95% CI using normal approximation to difference in binomial proportions.
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; Adj=adjusted; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In study M13-542, of the patients who achieved a DAS28(CRP)<2.6 response at Week 12 on placebo, upadacitinib 15 mg QD, and 30 mg QD, 19% (3/16), 28% (13/47), and 23% (9/39) of these DAS28(CRP)<2.6 responders had at least three active joint counts respectively.

In study M13-549, of the patients who achieved a DAS28(CRP)<2.6 response at Week 12 on placebo, upadacitinib 15 mg QD, and 30 mg QD, 18% (4/22), 25% (17/68), and 19% (12/62) of these DAS28(CRP)<2.6 responders had at least three active joint counts respectively.

In study M14-465, of the patients who achieved a DAS28(CRP)<2.6 on placebo, upadacitinib 15 mg QD, and adalimumab 40 mg EOW, 5% (2/40), 17% (31/187), and 10% (6/59) of these DAS28(CRP)<2.6 responders had at least three active joint counts respectively.

In study M15-555, of the patients who achieving a DAS28(CRP)<2.6 response at Week 14 on MTX monotherapy, upadacitinib 15 mg QD monotherapy, and 30 mg QD monotherapy, 6% (1/18), 15% (9/61), and 11% (10/87) of these DAS28(CRP)<2.6 responders had at least three active joint counts respectively.

In study M13-545, of the patients who achieving a DAS28(CRP)<2.6 response at Week 12 on MTX monotherapy, upadacitinib 15 mg QD monotherapy, and 30 mg QD monotherapy, 5% (2/42), 15% (17/113), and 17% (22/128) of these DAS28(CRP)<2.6 responders had at least three active joint counts respectively.

3.2.6.4 Other Endpoints

The applicant proposed to include both CDAI and SDAI (not controlled for multiplicity) in the product label. However, these endpoints are functions of the components of ACR and DAS28(CRP). Therefore, the results were not included in this section of the review. In summary, the results for these endpoints based on my evaluation were consistent with the primary findings of ACR and the components (Table 55, Table 56, Table 57, and Table 58).

3.2.6.5 Patient Reported Outcomes

The applicant proposed to include the following PROs, namely SF-36, FACIT-F, and morning stiffness in the label. Only the change from baseline in SF-36 PCS component was included in the multiplicity hierarchy. The applicant submitted the white papers for these endpoints. We refer you to the Clinical Outcomes Assessment consult for further information.

3.2.6.5.1 SF-36

In cDMARD add-on and monotherapy studies, the mean adjusted change from baseline in SF-36 PCS score at the primary timepoint was statistically significantly greater comparing patients randomized to upadacitinib 15 mg QD relative to the reference arm (Table 21).

Table 21 Change from Baseline in SF-36 PCS Summary Scores at Week 12/14

Treatment Arm	N	Baseline Mean (SD)	n	Visit[1] Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
cDMARD Add-on studies						
M13-542						
Placebo	166	31.6 (7.2)	147	34.6 (9.3)		
UPA 15 MG QD	163	30.6 (7.8)	157	37.1 (10.1)	3.1 (1.4, 4.8)	<0.001
UPA 30 MG QD	162	31.5 (7.3)	149	38.7 (9.4)	4.5 (2.8, 6.2)	<0.001
M13-549						
Placebo	221	33.1 (7.7)	207	36.8 (9.2)		
UPA 15 MG QD	219	33.4 (7.4)	211	41.3 (8.9)	4.4 (3.0, 5.8)	<0.001
UPA 30 MG QD	217	32.6 (7.9)	199	41.8 (8.9)	5.1 (3.6, 6.5)	<0.001
M14-465						
Placebo	647	32.5 (6.8)	632	36.6 (8.0)		
UPA 15 mg QD	648	32.5 (7.3)	634	40.8 (9.1)	4.3 (3.5, 5.1)	<0.001
ADA 40 mg EOW	327	32.2 (7.0)	322	38.9 (8.7)	2.5 (1.6, 3.5)	<0.001
Monotherapy Studies						
M15-555						
MTX	216	33.3 (7.3)	195	37.1 (8.1)		
UPA 15 MG QD	217	33.3 (7.9)	200	41.3 (9.1)	4.2 (2.8, 5.7)	<0.001
UPA 30 MG QD	214	33.9 (7.8)	202	43.6 (9.1)	6.2 (4.7, 7.6)	<0.001
M13-545						
MTX	313	33.1 (7.5)	299	38.8 (8.9)		
UPA 15 MG QD	315	32.7 (7.7)	306	43.0 (9.7)	4.5 (3.3, 5.8)	<0.001
UPA 30 MG QD	312	33.7 (7.2)	306	43.7 (8.5)	4.6 (3.3, 5.8)	<0.001

Mean(SD): Mean and standard deviation in parenthesis of the observed data were reported.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in PCS component adjusting for treatment groups, baseline PCS component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: SF-36=short form-36; PCS=physical component score; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; Adj=adjusted; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

The mean adjusted change from baseline in SF-36 MCS score at Week 12 in patients treated with upadacitinib 15 mg QD was nominally statistically significantly greater compared to patients treated with placebo in studies M13-549 and M14-465 (Table 22). In study M13-542, the mean adjusted change from baseline in SF-36 MCS score at Week 12 in patients treated with upadacitinib 15 mg QD was numerically higher compared to placebo treated patients. In the monotherapy studies, the mean adjusted changes from baseline in SF-36 MCS score at the primary timepoint were statistically significantly greater comparing patients treated upadacitinib 15 mg QD monotherapy to patients treated with MTX monotherapy.

Table 22 Change from Baseline in SF-36 MCS Summary Scores at Week 12/14

Treatment Arm	N	Baseline Mean (SD)	n	Visit[1] Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
cDMARD Add-on studies						
M13-542						
Placebo	166	45.9 (12.6)	147	49.2 (11.2)		
UPA 15 MG QD	163	44.0 (11.7)	157	49.0 (12.1)	1.0 (-0.8, 2.9)	0.28
UPA 30 MG QD	162	45.9 (12.3)	149	49.3 (11.4)	-0.2 (-2.1, 1.7)	0.82
M13-549						
Placebo	221	46.5 (11.7)	207	48.5 (11.4)		
UPA 15 MG QD	219	45.9 (10.9)	211	50.1 (9.9)	2.0 (0.5, 3.6)	0.01
UPA 30 MG QD	217	46.1 (12.0)	199	49.8 (10.4)	1.1 (-0.5, 2.7)	0.18
M14-465						
Placebo	647	43.0 (11.0)	632	46.2 (11.1)		
UPA 15 mg QD	648	43.0 (10.6)	634	48.8 (10.1)	2.7 (1.8, 3.6)	<0.001
ADA 40 mg EOW	327	42.7 (10.6)	322	47.7 (10.1)	1.7 (0.5, 2.8)	0.004
Monotherapy Studies						
M15-555						
MTX	216	45.1 (11.0)	195	47.5 (10.6)		
UPA 15 MG QD	217	44.1 (11.3)	200	49.4 (10.2)	2.5 (0.9, 4.1)	0.002
UPA 30 MG QD	214	44.5 (11.5)	202	50.0 (10.7)	2.7 (1.1, 4.3)	<0.001
M13-545						
MTX	313	43.2 (10.7)	299	46.9 (10.5)		
UPA 15 MG QD	315	42.5 (10.6)	306	48.6 (10.3)	2.1 (0.7, 3.5)	0.003
UPA 30 MG QD	312	43.3 (11.6)	306	48.9 (10.8)	2.1 (0.7, 3.5)	0.003

Mean(SD): Mean and standard deviation in parenthesis of the observed data were reported.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX),, respective 95% CI were reported based on a linear regression fit to the change from baseline in MCS component adjusting for treatment groups, baseline MCS component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: SF-36=short form-36; MCS=mental component score; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

In all studies, at the primary analysis timepoint, there was consistent evidence of improvement from baseline in the components of physical functioning, namely, role-physical, bodily-pain, and general health domains. In all studies except M13-542, there was consistent evidence of improvement from baseline in the SF36 MCS as well as the MCS components (See Appendix Table 59, Table 63, Table 60, Table 61, and Table 62).

3.2.6.5.2 Morning Stiffness

In study M13-549 and M15-555, there was statistically significantly reduction in the estimated mean adjusted change from baseline in duration of morning stiffness at the primary timepoint comparing randomized patients on upadacitinib 15 mg QD arm relative to the reference arm (Table 23). These results were also consistent in the remaining studies where the endpoint was not included in the multiplicity hierarchy.

Table 23 Duration of Morning Stiffness at Week of Analysis by Study by Arm

Treatment Arm	N	Baseline Mean (SD)	n	Visit1 Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
cDMARD Add-on Studies						
M13-542						
Placebo	169	138.4 (179)	151	133.1 (249)		
UPA 15 mg QD	164	140.4 (190)	160	68.3 (133)	-64.4 (-104.1, -24.7)	0.001
UPA 30 mg QD	165	184.5 (285)	153	89.8 (203)	-59.5 (-99.7, -19.3)	0.004
M13-549						
Placebo	216	138.9 (214)	202	95.7 (166)		
UPA 15 mg QD	217	152.4 (242)	207	54.3 (114)	-43.0 (-64.2, -21.7)	<0.001
UPA 30 mg QD	215	128.6 (156)	197	43.2 (67)	-49.4 (-70.9, -27.9)	<0.001
M14-465						
Placebo	651	142.4 (170)	625	91.8 (135)		
UPA 15 mg QD	324	146.1 (185)	314	61.0 (105)	-32.1 (-45.8, -18.5)	<0.001
ADA 40 mg EOW	649	141.5 (188)	625	47.9 (95)	-42.9 (-54.0, -31.7)	<0.001
Monotherapy Studies						
M15-555						
MTX	215	153.0 (222)	196	102.3 (190)		
UPA 15 mg QD	217	144.2 (215)	199	55.8 (111)	-44.5 (-69.2, -19.8)	<0.001
UPA 30 mg QD	214	133.9 (153)	202	43.2 (81)	-54.4 (-79.0, -29.8)	<0.001
M13-545						
MTX	314	128.1 (134)	291	71.9 (144)		
UPA 15 mg QD	316	168.9 (228)	303	43.9 (69)	-33.2 (-48.3, -18.1)	<0.001
UPA 30 mg QD	313	136.4 (167)	301	33.7 (49)	-39.0 (-54.0, -23.9)	<0.001

Mean and standard deviation in parenthesis were based on all observed data.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in duration of morning stiffness adjusting for treatment groups, baseline duration of morning stiffness, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate; ADA=adalimumab; EOW=every other week; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

There was nominally statistically significant reduction in mean adjusted change from baseline in the severity of morning stiffness comparing patients treated with upadacitinib 15 mg QD relative to the reference arm (Table 24).

Table 24 Severity of Morning Stiffness by Study by Arm

Treatment Arm	N	Baseline Mean (SD)	n	Visit[1] Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
cDMARD Add-on Studies						
M13-542						
Placebo	169	6.8 (2.3)	150	5.0 (2.8)		
UPA 15 mg QD	164	6.8 (2.1)	160	3.9 (2.7)	-1.1 (-1.7, -0.6)	<0.001
UPA 30 mg QD	165	6.5 (2.2)	152	3.4 (2.6)	-1.5 (-2.1, -0.9)	<0.001
M13-549						
Placebo	216	6.1 (2.2)	202	4.6 (2.7)		
UPA 15 mg QD	217	6.1 (2.3)	207	3.1 (2.5)	-1.5 (-1.9, -1.0)	<0.001
UPA 30 mg QD	214	6.2 (2.2)	196	2.8 (2.2)	-1.8 (-2.3, -1.4)	<0.001
M14-465						
Placebo	651	6.3 (2.3)	625	4.3 (2.6)		
UPA 15 mg QD	648	6.3 (2.3)	624	2.8 (2.4)	-1.5 (-1.8, -1.3)	<0.001
ADA 40 mg EOW	324	6.3 (2.1)	314	3.3 (2.4)	-1.0 (-1.3, -0.7)	<0.001
Monotherapy Studies						
M15-555						
MTX	215	6.0 (2.2)	196	4.4 (2.6)		
UPA 15 mg QD	217	5.9 (2.4)	199	2.8 (2.6)	-1.6 (-2.0, -1.1)	<0.001
UPA 30 mg QD	214	5.9 (2.4)	202	2.4 (2.3)	-2.0 (-2.5, -1.6)	<0.001
M13-545						
MTX	314	6.3 (2.4)	291	3.4 (2.6)		
UPA 15 mg QD	316	6.6 (2.3)	302	2.6 (2.6)	-1.0 (-1.4, -0.6)	<0.001
UPA 30 mg QD	313	6.4 (2.2)	300	2.4 (2.5)	-1.1 (-1.4, -0.7)	<0.001

Mean(SD): Mean and standard deviation in parenthesis of the observed data were reported.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX),, respective 95% CI were reported based on a linear regression fit to the change from baseline in severity of morning stiffness adjusting for treatment groups, baseline severity of morning stiffness, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate; ADA=adalimumab; EOW=every other week; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

3.2.6.5.3 FACIT-F

In studies M13-542 and M14-465, the adjusted mean changes from baseline in FACIT-F at Week 12 in patients treated with upadacitinib 15 mg QD were nominally statistically

significantly greater compared to patients treated with placebo (Table 25). The adjusted mean change from baseline in FACIT-F at Week 12 was nominally statistically significantly greater in patients on upadacitinib 15 mg QD monotherapy compared to patients on MTX monotherapy in study M13-545. There was lack of numerical trends towards greater mean change from baseline in FACIT-F at Week 12 for patients treated with the higher dose of upadacitinib for studies M13-542 and M13-545.

Table 25 Change from Baseline in FACIT-F at Week 12 in Studies M13-542, M14-465, and M13-545

Treatment Arm	N	Baseline Mean (SD)	n	Visit[1] Mean (SD)	Adj Mean Diff[2] (95% CI)	P-value[2]
M13-542						
Placebo	221	28.3 (11.5)	207	31.6 (11.8)		
UPA 15 mg QD	216	28.1 (11.1)	211	36.0 (10.4)	4.8 (3.2, 6.5)	<0.001
UPA 30 mg QD	217	27.5 (12.6)	199	36.4 (11.4)	4.9 (3.2, 6.5)	<0.001
M14-465						
Placebo	644	27.0 (11.1)	632	31.5 (11.6)		
UPA 15 mg QD	646	26.9 (11.1)	632	35.3 (10.5)	4.0 (3.0, 5.0)	<0.001
ADA 40 mg EOW	325	26.2 (11.4)	322	33.8 (11.3)	2.7 (1.5, 3.9)	<0.001
M13-545						
MTX	314	26.6 (11.7)	300	33.6 (11.3)		
UPA 15 mg QD	316	26.4 (11.9)	308	37.0 (10.8)	3.6 (2.2, 5.1)	<0.001
UPA 30 mg QD	310	27.8 (11.1)	306	37.4 (10.7)	3.2 (1.7, 4.7)	<0.001

Mean(SD): Mean and standard deviation in parenthesis of the observed data were reported.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Estimated difference in the adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in FACIT-F adjusting for treatment groups, baseline FACIT-F, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate; ADA=adalimumab; EOW=every other week; FACIT-F=functional assessment of chronic illness-fatigue
[Source: Statistical Reviewer]

3.2.6.6 Radiographic Endpoint

3.2.6.6.1 Study M14-465

At Week 14, 26% of the randomized patients underwent radiograph evaluation at Week 14, with a majority of these patients coming from placebo arm (Table 26). Of these patients, fewer than 1% of these patients had discontinued randomized treatment prior to Week 14.

At Week 26, 90% of the patients, remaining in the study regardless of discontinuation of randomized or rescue treatment, had an evaluable radiograph collected within the visit window. However, only 45%, 71%, and 65% of the patients randomized to placebo, upadacitinib 15 mg QD arm, and adalimumab 40 mg EOW, remained on randomized treatment.

Table 26 Disposition of Patients with mTSS Evaluated during the First 26 Weeks, Study M14-465

	Placebo (N=651)	UPA 15 MG QD (N=651)	ADA 40 MG EOW (N=327)	All (N=1629)
Baseline				
Patients with evaluable mTSS	649 (100%)	644 (99%)	326 (100%)	1619 (99%)
Patients with ≥ 1 post-baseline mTSS value	608 (93%)	599 (92%)	298 (91%)	1505 (92%)
Patients without post-baseline mTSS value	41 (6%)	45 (7%)	28 (9%)	114 (7%)
Patients missing mTSS at baseline	2 (0%)	7 (1%)	1 (0%)	10 (1%)
Week 14				
Patients with evaluable mTSS	243 (37%)	110 (17%)	67 (20%)	420 (26%)
Patients without baseline mTSS	2	5	1	8
On randomized treatment (prior to rescue)	238 (37%)	105 (16%)	63 (19%)	406 (25%)
Discontinued randomized treatment [1]	5 (1%)	5 (1%)	4 (1%)	14 (1%)
Week 26				
Patients with evaluable mTSS	588 (90%)	588 (90%)	291 (89%)	1467 (90%)
Patients without baseline mTSS	1	6	1	8
On randomized treatment (never rescued)	295 (45%)	465 (71%)	211 (65%)	971 (60%)
Discontinued randomized treatment	7 (1%)	6 (1%)	7 (2%)	20 (1%)
Remained on rescue treatment [2]	282 (43%)	113 (17%)	73 (22%)	468 (29%)
Discontinued rescue treatment [2]	4 (1%)	4 (1%)	-	8 (0%)
Patients without mTSS	63 (10%)	63 (10%)	36 (11%)	162 (10%)
Discontinued rescue/randomized treatment [3]	18 (3%)	22 (3%)	4 (1%)	44 (3%)
Discontinued study participation	43 (7%)	39 (6%)	32 (10%)	114 (7%)
On either rescue/randomized treatment	2 (<1%)	2 (<1%)	-	4 (<1%)

1: Prior to Week 14
2: Patients were rescued to different study treatment prior to Week 26 but after Week 14
3: Patients remained in the study but had discontinued either randomized treatment or rescue treatment.
Abbreviations: mTSS=modified total sharp score; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation
[Source: Statistical Reviewer]

Based on all radiographs collected regardless of escape or treatment discontinuation, there were statistically significant evidence that upadacitinib delays radiograph progression at Week 26 compared to placebo arm (Table 27) as evaluated using mTSS as well as its components.

Table 27 Analysis for the Mean Change from Baseline in mTSS and Components at Week 26 including Data Collected after Rescue, Study M14-465

Treatment Arm	N	Baseline Mean (SD)	n	Change from Baseline Mean (SD)	Est Diff (95% CI)[1,2]	P-value	Est Diff (95% CI)[1,3]	P-value
mTSS								
Placebo	649	35.9 (51.7)	588	0.7 (3.0)				
UPA 15 mg QD	644	34.0 (50.1)	588	0.2 (1.4)	-0.56 (-0.83, -0.29)	<0.001	0.16 (-0.17, 0.49)	0.34
ADA 40 mg EOW	326	34.5 (47.1)	291	0.0 (2.4)	-0.72 (-1.1, -0.39)	<0.001		
Erosion Score								
Placebo	649	17.0 (27.4)	588	0.4 (1.7)				
UPA 15 mg QD	644	16.5 (26.4)	588	0.0 (0.7)	-0.35 (-0.49, -0.21)	<0.001	0.01 (-0.16, 0.18)	0.91
ADA 40 mg EOW	326	15.4 (23.1)	291	0.0 (0.9)	-0.36 (-0.53, -0.18)	<0.001		
Joint Space Narrowing								
Placebo	649	18.9 (26.1)	588	0.4 (2.1)				
UPA 15 mg QD	644	17.5 (25.1)	588	0.1 (1.0)	-0.28 (-0.46, -0.1)	0.002	0.08 (-0.13, 0.3)	0.45
ADA 40 mg EOW	326	19.2 (25.8)	291	0.0 (0.9)	-0.36 (-0.58, -0.15)	0.001		

1: Estimated difference in the adjusted mean change from baseline compared to placebo, respective 95% CI, and p-values were based on a linear regression fit to the change from baseline in mTSS adjusting for treatment, baseline mTSS, key stratification factor.
2: Comparison is made with respect to placebo arm
3: Comparison is made with adalimumab 40 mg EOW
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; mTSS=modified total Sharpe score; N=total number of patients at baseline with measurement; n=total number of patients at Week 26 with observed x-ray
[Source: Statistical Reviewer]

The proportions of patients with no radiographic progression, as measured by an observed change from baseline ≤0, at Week 26 (Table 28). These results were consistent with the key radiograph results based on the treatment policy estimand, providing additional supportive evidence of the efficacy of both dosing regimens of upadacitinib.

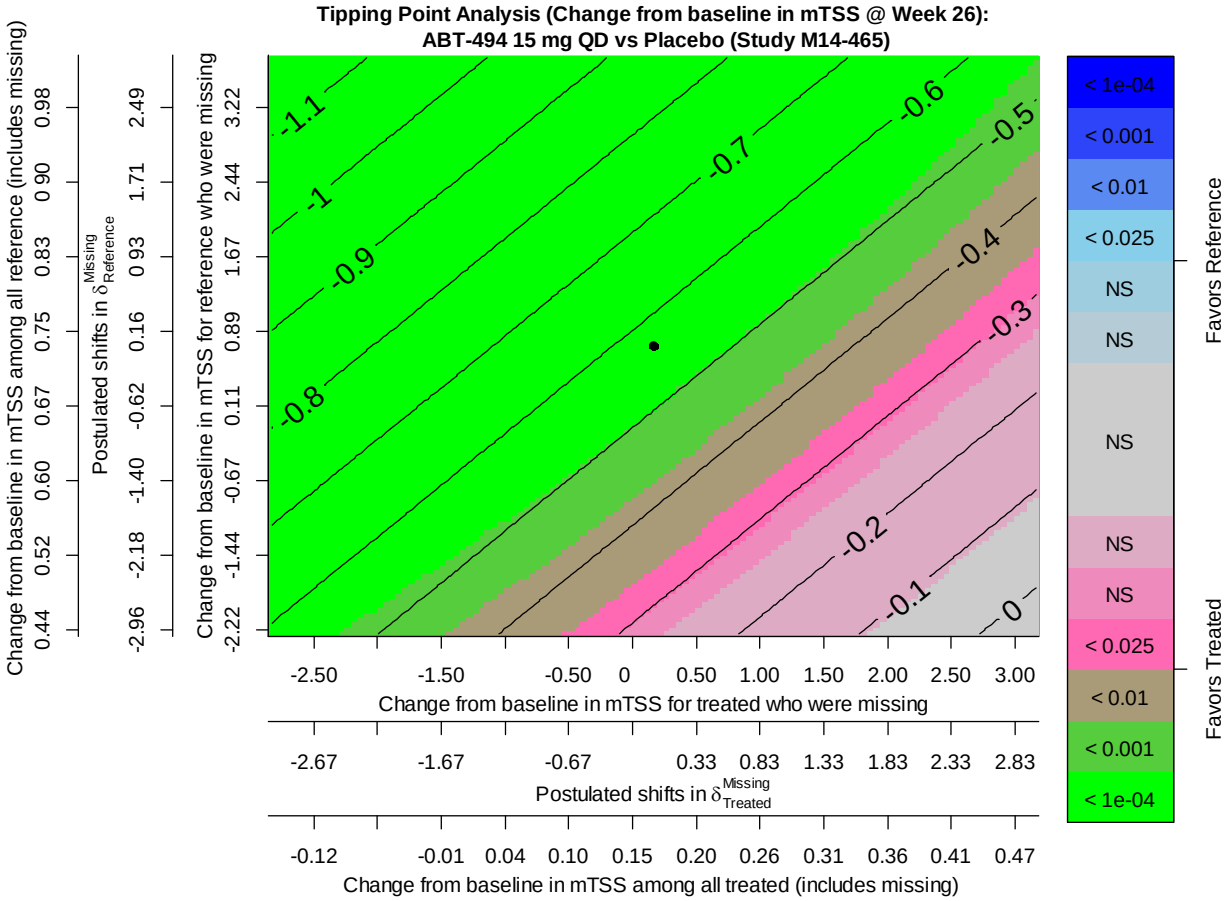
Table 28 Proportion of Patients with No Change in mTSS Score at Week 26, Study M14-465

Treatment Arm	Excluding Data after Rescue (LE)				Using All Data Collected after Rescue			
	n	Counts (%)[1]	Est Difference (95% CI) [2]	P-value[2]	n	Counts (%)[1]	Est Difference (95% CI) [2]	P-value[2]
mTSS								
Placebo	599	455 (76%)			587	451 (77%)		
UPA 15 mg QD	593	495 (83%)	7.5% (3.0% - 12.1%)	0.001	582	483 (83%)	6.2% (1.6% - 10.7%)	0.009
ADA 40 mg EOW	296	257 (87%)	10.9% (5.7% - 16.0%)	<0.001	290	251 (87%)	9.7% (4.5% - 14.9%)	<0.001
Erosion Score								
Placebo	599	498 (83%)			587	491 (84%)		
UPA 15 mg QD	593	536 (90%)	7.2% (3.4% - 11.1%)	<0.001	582	526 (90%)	6.7% (2.9% - 10.6%)	<0.001
ADA 40 mg EOW	296	269 (91%)	7.7% (3.3% - 12.2%)	0.002	290	265 (91%)	7.7% (3.3% - 12.1%)	0.002
Joint Space Narrowing								
Placebo	599	500 (83%)			587	488 (83%)		
UPA 15 mg QD	593	528 (89%)	5.6% (1.7% - 9.5%)	0.005	582	516 (89%)	5.5% (1.5% - 9.5%)	0.007
ADA 40 mg EOW	296	273 (92%)	8.8% (4.5% - 13.0%)	<0.001	290	265 (91%)	8.2% (3.8% - 12.7%)	0.001

1: Counts and percentages relative to n in parenthesis are reported for the proportion with no radiographic change from baseline.
2: Difference in the proportions, respective 95% CI using normal approximation to difference in binomial proportions, p-values from the Cochran Mantel Haenszel test were reported.
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; mTSS=modified total sharp score
[Source: Statistical Reviewer]

In the tipping point sensitivity analysis targeting the treatment policy estimand, there was further supportive evidence of the robustness of evidence for upadacitinib 15 mg QD dosing relative to placebo arm (Figure 5). Under various scenarios where placebo patients with truly missing data were assumed to have slower radiograph progression, as measured by smaller average change from baseline, and upadacitinib patients with truly missing data were assumed to have larger average change, the conclusions of the above findings remained statistically significant. These sensitivity analyses indicated that results did not tip in favor of the placebo arm even under implausible assumptions (such as an assumption of more than 2-point average worsening among upadacitinib dropouts and on average no worsening among placebo arm dropouts).

Figure 5 Tipping Point Sensitivity Analyses for Week 26 Observed Data regardless of any Escape Comparing Upadacitinib 15 mg QD (ABT-494/Treated) with Placebo Arm (Reference)



Abbreviations: mTSS=modified total Sharp score; QD=once daily;
[Source: Statistical Reviewer]

Using linear extrapolation to address the hypothetical estimand in the absence of escape, there were statistically significant evidence that upadacitinib delays radiograph progression (Table 64). However, based on statistical limitations identified in 3.2.3.3, I included the results based on the random coefficient model due to the statistical issues identified in Section 3.2.4. In these analyses, data after placebo patients were rescued to a different drug were assumed to be missing. Using the random coefficient model, the estimated adjusted mean (linear) rate of mTSS progression over 26 weeks among placebo patients in the absence of rescue is 0.78 units (Table 29). The estimated adjusted mean rate of mTSS progression over 26 weeks among upadacitinib patients in the absence of rescue is 0.15 units. The estimated difference in the adjusted mean rate of mTSS progression over 26 weeks comparing upadacitinib arm vs placebo arm is -0.63 (95% CI: -0.92, -0.34; p<0.001). These results were consistent with the applicant's results based on linear extrapolation. The estimated difference in the adjusted mean rate of progression in the individual components of mTSS, i.e., the erosion score and joint space narrowing score, over 26 weeks were significantly lower in the upadacitinib arm compared to the placebo arm.

Table 29 Estimated Rate of mTSS Progression over 26 weeks, Study M14-465

	Est Rate of mTSS Progression over 26 weeks		Est Difference in Rate (95% CI)	P-value
	Placebo Est Rate (SE)	Upadacitinib 15 mg QD Est Rate (SE)		
Primary Analysis using LE	NA	NA	-0.67 (-0.97, -0.37)	<0.001
FDA's analysis: mTSS	0.78 (0.14)	0.15 (0.06)	-0.63 (-0.92, -0.34)	<0.001
FDA's analysis: Erosion Score	0.43 (0.08)	0.04 (0.03)	-0.40 (-0.56, -0.23)	<0.001
FDA's analysis: JSN	0.44 (0.10)	0.11 (0.04)	-0.34 (-0.55, -0.14)	0.001

Random coefficient model fit to the mTSS value adjusting for time, treatment group, prior bDMARDs use, treatment group-by-time interaction with random slopes and random intercept. Covariance structure allowed for heterogeneity between treatment groups.
Abbreviations: PBO=placebo; UPA=upadacitinib; mTSS=modified total sharp score; JSN=joint space narrowing; bDMARDs=biologic disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

The random coefficient analysis model was repeated on all data collected regardless of escape therapy. The estimated difference in the adjusted mean rate of mTSS progression over 26 weeks comparing upadacitinib over placebo is -0.52 (95% CI: -0.75, -0.28; p-value <0.001) and were consistent with the findings based on all ANCOVA model pre-specified by the applicant (Table 30).

Table 30 Estimated Rate of mTSS Progression over Time Based on all Observed Data Regardless of Adherence or Use of Rescue, Study M14-465

	Est Rate of mTSS Progression at Week 26		Est Difference in Rate (95% CI)	P-value
	Placebo Est Rate (SE)	Upadacitinib 15 mg QD Est Rate (SE)		
Analysis using Observed Data	NA	NA	-0.56 (-0.83, -0.29)	<0.001
FDA's analysis: mTSS	0.67 (0.11)	0.16 (0.05)	-0.52 (-0.75, -0.28)	<0.001
FDA's analysis: Erosion Score	0.36 (0.06)	0.05 (0.03)	-0.31 (-0.45, -0.18)	<0.001
FDA's analysis: JSN	0.37 (0.08)	0.11 (0.04)	-0.25 (-0.43, -0.08)	0.004

Random coefficient model fit to the mTSS value adjusting for time, treatment group, prior bDMARDs use, treatment group-by-time interaction with random slopes and random intercept. Covariance structure allowed for heterogeneity between treatment groups.
Abbreviations: PBO=placebo; UPA=upadacitinib; mTSS=modified total sharp score; JSN=joint space narrowing; bDMARDs=biologic disease modifying anti-rheumatic drugs; NA=not applicable
[Source: Statistical Reviewer]

In summary, when the treatment policy estimand is of interest, analyses at Week 26 based on all data collected after treatment discontinuation or escape, additional supportive analyses (based on the random coefficient regression model), and tipping point analyses conducted provide

statistically significant evidence of delay in radiograph progression for upadacitinib 15 mg QD relative to placebo and the results were robust to missing data assumptions. In addition, when the estimand in the absence of any crossover to experimental treatment or active comparator is of interest, results from the applicant's analyses and the supportive analyses that I conducted were consistent and supportive of the findings that upadacitinib 15 mg QD statistically significantly delayed radiograph progression over 26 weeks.

3.2.6.6.2 Study M13-545

At Week 24, 87% of the randomized patients who remained in the study had an evaluable radiograph. The proportion of patients with radiograph evaluation at Week 24 were similar across arms.

Table 31 Disposition of Patients with Radiographs Evaluated up to Week 24, Study M13-545

	MTX (N=314)	UPA 15 mg QD (N=317)	UPA 30 mg QD (N=314)	All (N=945)
Baseline				
Patients with evaluable mTSS	309 (98%)	309 (97%)	309 (98%)	927 (98%)
Patients missing mTSS at baseline	5 (2%)	8 (3%)	5 (2%)	18 (2%)
Week 12				
Patients with evaluable mTSS	2 (1%)	2 (1%)	1 (<1%)	5 (1%)
Patients without baseline mTSS	1	1	1	3
On randomized treatment (prior to rescue)	2 (1%)	2 (1%)	1 (<1%)	5 (1%)
Discontinued randomized treatment [a]	-	-	-	-
Week 24				
Patients with evaluable mTSS	268 (85%)	281 (89%)	273 (87%)	822 (87%)
Patients without baseline mTSS	1	1	1	3
On randomized treatment	241 (77%)	265 (84%)	266 (85%)	772 (82%)
Had rescue treatment	21 (7%)	14 (4%)	4 (1%)	39 (4%)
Discontinued randomized treatment [b]	6 (2%)	2 (1%)	3 (1%)	11 (1%)
Patients without mTSS	46 (15%)	36 (11%)	41 (13%)	123 (13%)
Discontinued rescue/randomized treatment[c]	3 (1%)	1 (<1%)	5 (2%)	9 (1%)
Discontinued study participation	37 (12%)	24 (8%)	23 (7%)	84 (9%)
Ongoing	-	-	1 (<1%)	1 (<1%)
On either rescue/randomized treatment	6 (2%)	11 (3%)	12 (4%)	29 (3%)

Counts and percentages relative to N are presented
a: Some patients had Week 12 radiograph assessments.
b: These patients were excluded from the primary analysis because Week 12 radiographs were not collected.
c: Patients remained in the study but had discontinued either randomized treatment or rescue treatment.
Abbreviations: mTSS=modified total sharp score; UPA=upadacitinib; QD=once daily; MTX=methotrexate
[Source: Statistical Reviewer]

There was statistically significantly evidence of delay in radiograph progression comparing upadacitinib monotherapy relative to MTX monotherapy at Week 24 based on all observed data collected regardless of escape or treatment discontinuation (Table 32). Results for the components of mTSS, namely, erosion and JSN scores, were also statistically significant.

Table 32 Analysis for the Mean Change from Baseline in mTSS and Components at Week 24, including All Observed Data at Week 24, Study M13-545

Treatment Arm	N	Baseline Mean (SD)	n	Mean Change from baseline (SD)	Est Diff (95% CI)[1]	P-value
mTSS						
MTX	309	13.3 (30.5)	267	0.6 (2.8)		
UPA 15 mg QD	309	18.1 (38.2)	280	0.1 (1.4)	-0.52 (-0.8, -0.2)	0.001
UPA 30 mg QD	309	17.2 (38.3)	272	0.1 (1.2)	-0.58 (-0.9, -0.3)	<0.001
Erosion Score						
MTX	309	6.1 (15.5)	267	0.3 (1.4)		
UPA 15 mg QD	309	8.6 (19.3)	280	0.0 (0.5)	-0.29 (-0.4, -0.1)	<0.001
UPA 30 mg QD	309	8.0 (18.9)	272	0.0 (0.3)	-0.33 (-0.5, -0.2)	<0.001
Joint Space Narrowing						
MTX	309	7.2 (16.1)	267	0.3 (1.7)		
UPA 15 mg QD	309	9.6 (20.1)	280	0.1 (0.9)	-0.23 (-0.4, -0.02)	0.03
UPA 30 mg QD	309	9.3 (20.3)	272	0.1 (1.1)	-0.24 (-0.4, -0.03)	0.03

1: Estimated difference in the adjusted mean change from baseline compared to MTX, respective 95% CI, and p-values were based on a linear regression fit to the change from baseline in mTSS component adjusting for treatment, baseline mTSS component, key stratification factor.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate; mTSS=modified total Sharpe score
[Source: Statistical Reviewer]

When targeting the hypothetical estimand, in the absence of escape or discontinuation of randomized treatment, there was also statistically significant evidence in delay in radiograph progression at Week 24 comparing upadacitinib monotherapy relative to MTX monotherapy (Table 33). However, in this analysis, I note that there were no patients who were linearly extrapolated.

Table 33 Applicant's Primary Analysis for the Mean Change from Baseline in mTSS and Components at Week 24, Study M13-545

Treatment Arm	N	Baseline Mean (SD)	n	Mean Change from Baseline (SD)	Est Diff (95% CI)[1]	P-value
mTSS						
MTX	309	13.3 (30.5)	264	0.7 (2.8)		
UPA 15 mg QD	309	18.1 (38.2)	279	0.1 (1.4)	-0.53 (-0.8, -0.2)	0.001
UPA 30 mg QD	309	17.2 (38.3)	270	0.1 (1.2)	-0.59 (-0.9, -0.3)	<0.001
Erosion Score						
MTX	309	6.1 (15.5)	264	0.3 (1.5)		
UPA 15 mg QD	309	8.6 (19.3)	279	0.0 (0.5)	-0.29 (-0.4, -0.1)	<0.001
UPA 30 mg QD	309	8.0 (18.9)	270	0.0 (0.3)	-0.34 (-0.5, -0.2)	<0.001
Joint Space Narrowing						
MTX	309	7.2 (16.1)	264	0.3 (1.7)		
UPA 15 mg QD	309	9.6 (20.1)	279	0.1 (0.9)	-0.23 (-0.4, -0.02)	0.03
UPA 30 mg QD	309	9.3 (20.3)	270	0.1 (1.1)	-0.24 (-0.5, -0.03)	0.02

1: Estimated difference in the mean change from baseline, respective 95% CI, and p-values were based on a linear regression fit to the change from baseline in mTSS component adjusting for treatment, baseline mTSS component, key stratification factor.
Patients who discontinued randomized treatment after Week 16 had their radiograph at Week 24 linearly extrapolated if Week 12 radiographs were collected. However, no such patients were linearly extrapolated.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate; mTSS=modified total Sharpe score
[Source: Statistical Reviewer]

The proportions of patients with no radiographic progression, as measured by an observed change from baseline ≤0, at Week 24 (Table 34). These results were consistent with the key radiograph results based on the treatment policy estimand, providing additional supportive evidence of the efficacy of both dosing regimens of upadacitinib.

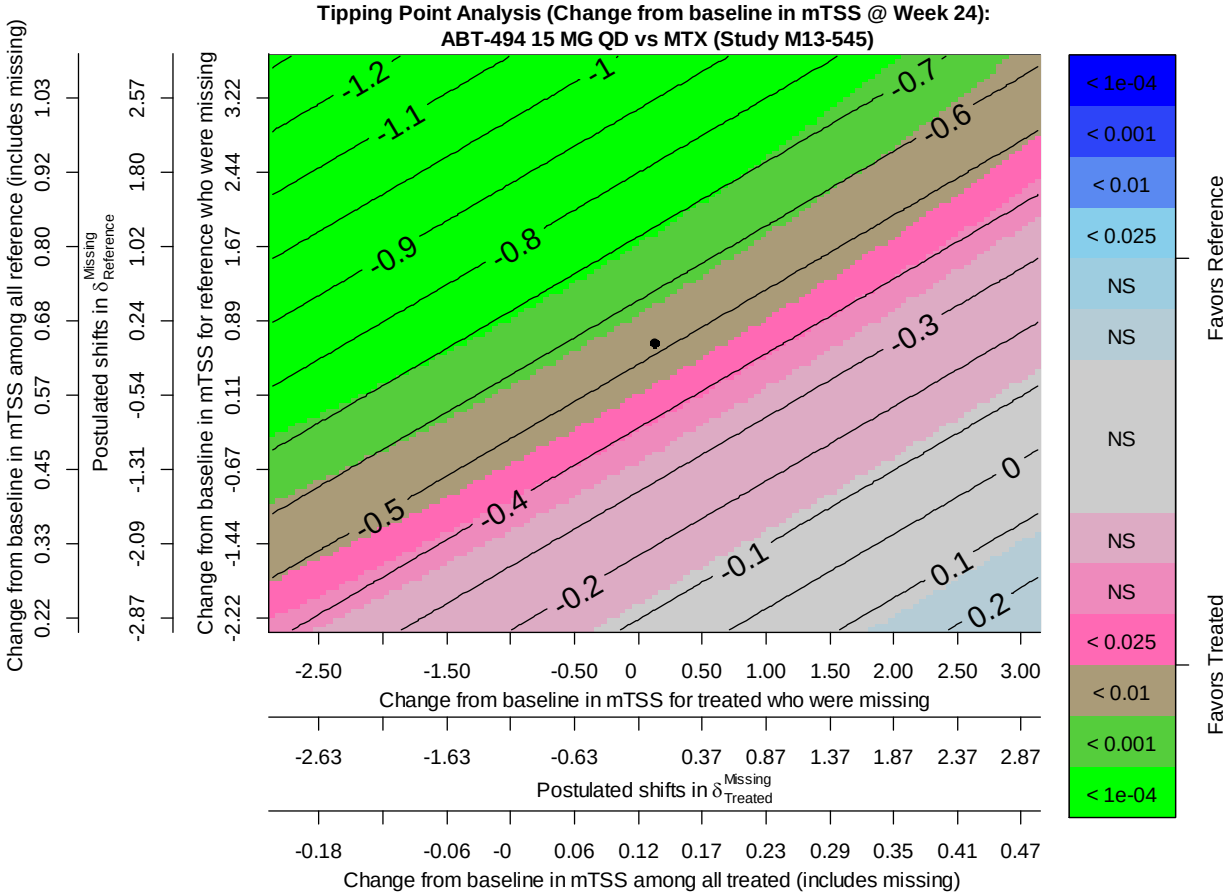
Table 34 Proportion of Patients with No Change in mTSS at Week 24, Study M13-545

Arms	Excluding Data after Rescue (LE)				Using All Data Collected after Rescue			
	n	Counts (%)[1]	Est Difference (95% CI) [2]	P-value[2]	n	Counts (%)[1]	Est Difference (95% CI) [2]	P-value[2]
mTSS								
MTX	264	205 (78%)			267	208 (78%)		
UPA 15 mg QD	279	244 (87%)	9.8% (3.5%, 16.2%)	0.002	280	245 (88%)	9.6% (3.3%, 15.9%)	0.003
UPA 30 mg QD	270	241 (89%)	11.6% (5.4%, 17.8%)	<0.001	272	243 (89%)	11.4% (5.3%, 17.6%)	<0.001
Erosion Score								
MTX	264	217 (82%)			267	220 (82%)		
UPA 15 mg QD	279	259 (93%)	10.6% (5.1%, 16.2%)	<0.001	280	260 (93%)	10.5% (5.0%, 15.9%)	<0.001
UPA 30 mg QD	270	257 (95%)	13.0% (7.7%, 18.3%)	<0.001	272	259 (95%)	12.8% (7.6%, 18.0%)	<0.001
Joint Space Narrowing								
MTX	264	230 (87%)			267	233 (87%)		
UPA 15 mg QD	279	255 (91%)	4.3% (-0.9%, 9.5%)	0.11	280	256 (91%)	4.2% (-1.0%, 9.3%)	0.12
UPA 30 mg QD	270	248 (92%)	4.7% (-0.5%, 9.9%)	0.07	272	250 (92%)	4.6% (-0.5%, 9.8%)	0.08

1: Counts and percentages relative to n in parenthesis are reported for the proportion with no radiographic change from baseline.
2: Difference in the proportions, respective 95% CI using normal approximation to difference in binomial proportions, p-values from the Cochran Mantel Haenszel test were reported.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total randomized; QD=once daily; n=number of patients with observed data; MTX=methotrexate
[Source: Statistical Reviewer]

In the tipping point sensitivity analysis targeting the treatment policy estimand, there was further supportive evidence of the robustness of evidence for both upadacitinib monotherapy dosing regimens relative to MTX monotherapy (Figure 6) in the setting when patients had early signs and symptoms of RA and some evidence of radiograph progression. Under various scenarios where MTX monotherapy patients with truly missing data were assumed to have slower radiograph progression, as measured by smaller average change from baseline, and upadacitinib patients with truly missing data were assumed to have larger average change, the conclusions of the above findings remained statistically significant. These sensitivity analyses indicated that results did not tip in favor of the MTX monotherapy even under implausible assumptions (such as an assumption of more than 2-point average worsening among upadacitinib dropouts and on average no worsening among MTX dropouts).

Figure 6 Tipping Point Sensitivity Analyses for Week 24 Observed Data regardless of any Escape Comparing Upadacitinib 15 mg QD (ABT-494) Monotherapy with Methotrexate Monotherapy, Study M13-545



Abbreviations: mTSS=modified total Sharp score; QD=once daily;
[Source: Statistical Reviewer]

3.3 Evaluation of Safety

3.3.1 Statistical Methodologies

The reader is referred to the medical reviewer for a full comprehensive finding on safety which included review of the safety for the individual studies. I reported key findings important to facilitate the benefit-risk assessment that is proposed to be included in the product label. When possible integrated summary of safety is provided. In this review, I included treatment emergent AEs based on the applicant's definitions. For the placebo-controlled or active-controlled period, a treatment emergent adverse event is defined if the AE onset date is on or after the first dose of oral study drug and prior to Week 14 dose date during treatment period 1. If a patient discontinues study drug prematurely before the Week 14 dosing, an AE is counted if the onset of

the AE occurred within 30 days after the last dose of placebo or upadacitinib and 70 days (due to different half-life) for patients on adalimumab.

I reported the following AEs: infections, any serious infections, any opportunistic infections, any malignancies excluding non-melanoma skin cancer (NMSC), any lymphomas, any gastrointestinal perforations, any neutropenias, any anemia, any herpes zoster, any tuberculosis, adjudicated MACE, adjudicated VTEs.

During the review cycle, the review team considered the following strategy to present the safety data for labelling purposes to be useful given the complexity of the study design, different treatment arms, and different patient population. The following integrated analysis datasets were considered relevant for labelling.

1. Patients exposed to either placebo in combination with cDMARD or upadacitinib 15 mg QD or 30 mg QD in combination with cDMARD during the 12/14 weeks placebo-controlled period based on integrating studies M13-542 and M13-549.
2. Patients exposed to either MTX monotherapy or upadacitinib 15 mg QD or 30 mg QD monotherapy during the 12/14 weeks MTX-controlled period based on integrating studies M15-555 and M13-545.
3. Patients in study M14-465 exposed to either placebo in combination with cDMARD or upadacitinib 15 mg QD in combination with cDMARD or adalimumab 40 mg EOW in combination with cDMARD were reported separately for the placebo-controlled 14-week period.
4. Patients exposed to placebo in combination with cDMARD or upadacitinib 15 mg QD in combination with cDMARD during the 12/14 weeks placebo-controlled period based on integrating studies M13-542, M13-549, and M14-465.

Because the study designs had allowed crossover between treatment arms after Week 12/14, interpretation of long-term safety was limited without a randomized comparator. For example, in study M14-465, patients on adalimumab can potentially crossover to upadacitinib if they met rescue criteria or patients on upadacitinib can potentially crossover to adalimumab after Week 14. Further, the placebo arm was entirely crossed over to upadacitinib after Week 26 (and more than 40% had crossover to upadacitinib arm after Week 14 and prior to Week 26).

Given the above limitations, we found it reasonable to characterize, at minimal, rare AEs based only on the studied doses of upadacitinib as taken up to at most one year, in absence of a controlled comparator. These rare AEs included gastrointestinal perforations, malignancies, VTE, and MACE and were of interest to the clinical team. Therefore, we pooled all the studies except M14-465 that evaluated both doses of upadacitinib. We find it reasonable to pool these four studies to described potential dose-dependent AEs even though they were active-controlled studies and placebo-controlled studies. The interpretation of these rates from pooling the four studies are limited to the studied doses and are not directly comparable to rates evaluated during the placebo-controlled, or active controlled period, or any background rates in the available literature given the absence of a concurrent controlled comparator. Study M14-465 was not included in the pooled analysis since only one dose of upadacitinib was studied. Inclusion of the 15 mg QD arm into the pooled analysis could potentially introduce confounding.

To address whether these rates could have been affected by crossover from placebo to upadacitinib, I included an analysis based only on patients as originally randomized to the dose of upadacitinib. This analysis excluded patients who had crossover (either pre-randomized or met rescue criteria) to upadacitinib.

We focus on the exposure-adjusted incidence rates (EAIR) because it more appropriately reflects and account for different follow-up of patients between treatment groups. To compute the exposure-adjusted incidence rate for each study, the numerator was defined to be the total number of patients with an event of interest and the denominator is the total exposure time for the patients in the treatment arm of interest who is at risk of an initial occurrence of the event. Patients without an event will have the total exposure time under the treatment group until the end of the defined period.

We reported study adjusted EAIRs to account for differences between exposures across studies by using Mantel Haenszel (MH) weights. To compare between EAIRs across treatment arms in the pooled studies, 95% CI MH weights were reported for comparisons between upadacitinib and reference arm. For 1-year dose comparisons, similar 95% CI with MH weights were reported for comparisons between the dose groups of upadacitinib in the absence of a reference arm.

During the review, I noted that the applicant did not report MH-adjusted EAIR within the individual treatment arm. Therefore, to be consistent with the reporting of the 95% CI and to account for potential differences in treatment exposure across studies, I reported MH-adjusted EAIR within the individual treatment arm. I note that the applicant had used MH weights using the total exposure within each study and later presented integrated safety findings using study size weights and the results were similar. In this review, I presented safety findings based on the MH weights specified in the original submission.

3.3.2 Results and Conclusions

3.3.2.1 Placebo-Controlled or MTX-Controlled Period (12/14 Weeks)

During the 12/14-week placebo-controlled period, the MH-adjusted EAIR for treatment emergent adverse events related to infections SAEs were 136, 164, and 180 per 100 PY for placebo, upadacitinib 15 mg QD, and upadacitinib 30 mg QD respectively. The estimated difference in the MH-adjusted EAIR of infections were 28.4 per 100 PY (95% CI: -11.5, 68.4) and 44.3 per 100 PY (95% CI: 3, 85.7) for comparing upadacitinib 15 mg QD vs placebo and upadacitinib 30 mg QD vs placebo respectively (Table 35). The MH-adjusted EAIRs were largely similar comparing upadacitinib 15 mg QD vs placebo and the wide CIs provided some measure of uncertainty in estimating the risk differences.

During the 12/14-week placebo-controlled period, the MH-adjusted EAIR of serious infections was statistically significantly higher in the upadacitinib 30 mg QD arm compared to placebo with an estimated difference in MH-adjusted EAIR of 7.1 per 100 PY (95% CI: 0.6, 13.6). The MH-adjusted EAIRs were largely similar comparing upadacitinib 15 mg QD vs placebo and the wide CIs provided some measure of uncertainty in estimating the risk differences.

In summary, there was an observed increase in MH-adjusted EAIR for infections, serious infections, opportunistic infections, malignancies excluding NMSC, any hepatic disorders, lymphopenia, herpes zoster towards the higher upadacitinib dosing relative to the 15 mg QD.

Table 35 Exposure Adjusted Incidence Rates for Treatment Emergent Adverse Events of Special Interests During Placebo-Controlled Period 12/14 Weeks, Studies M13-542 and M13-549

	Placebo + cDMARD (N=390)	UPA 15 mg QD + cDMARD (N=385)	UPA 30 mg QD + cDMARD (N=384)	UPA 15 mg Vs Placebo	UPA 30 mg Vs Placebo
	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	Diff [95% CI] [2]	Diff [95% CI] [2]
Infections	99/73.0 (136)	118/71.8 (164)	126/69.9 (180)	28.4 [-11.5, 68.4]	44.3 [3.0, 85.7]
Serious Infections	1/85.4 (1.2)	2/86.6 (2.3)	7/84.8 (8.2)	1.1 [-2.8, 5.1]	7.1 [0.6, 13.6]
Opportunistic Infection	1/85.4 (1.2)	2/86.5 (2.3)	6/84.6 (7.1)	1.1 [-2.8, 5.1]	5.9 [-0.2, 12.0]
Malignancy	0/85.4 (-)	1/86.8 (1.2)	5/84.7 (5.9)	1.2 [-1.1, 3.4]	5.9 [0.7, 11.1]
Malignancy (excluding NMSC)	0/85.4 (-)	1/86.8 (1.2)	3/84.8 (3.5)	1.2 [-1.1, 3.4]	3.5 [-0.5, 7.5]
Lymphoma	0/85.4 (-)	0/86.8 (-)	1/84.9 (1.2)	NA	1.2 [-1.1, 3.5]
Any Hepatic Disorder	9/84.7 (10.6)	5/86.4 (5.8)	9/83.8 (10.8)	-4.8 [-13.4, 3.8]	0.1 [-9.7, 10.0]
Any GI perforation	0/85.4 (-)	0/86.8 (-)	0/85.0 (-)	NA	NA
Anemia	2/85.2 (2.3)	1/86.6 (1.2)	4/84.6 (4.7)	-1.2 [-5.1, 2.8]	2.4 [-3.3, 8.1]
Neutropenia	0/85.4 (-)	10/85.3 (11.7)	12/83.4 (14.4)	11.7 [4.5, 19.0]	14.4 [6.3, 22.6]
Lymphopenia	3/85.0 (3.5)	3/86.4 (3.5)	7/84.2 (8.3)	-0.1 [-5.7, 5.5]	4.8 [-2.6, 12.1]
Herpes Zoster	2/85.2 (2.3)	2/86.6 (2.3)	6/84.6 (7.1)	-0.0 [-4.6, 4.5]	4.7 [-1.8, 11.3]
Active/Latent TB	0/85.4 (-)	0/86.8 (-)	0/85.0 (-)	NA	NA
MACE	0/85.4 (-)	1/86.7 (1.2)	1/85.0 (1.2)	1.2 [-1.1, 3.4]	1.2 [-1.1, 3.5]
Adjudicated VTEs	0/85.4 (-)	1/86.8 (1.2)	0/85.0 (-)	1.2 [-1.1, 3.4]	0.0 [0.0, 0.0]

Treatment-emergent adverse event is defined as any adverse event with an onset date on or after the first dose of study drug in period 1 and prior to the week 14 dose date in period 1 or up to 30 days after the last dose of placebo or upadacitinib and 70 days for adalimumab, if subject discontinued study drug prematurely before week 14 dosing in period 1 of the study.
1: EAIR rates were stratified by study using MH weights using total exposure within each study.
2: Difference in EAIR comparing each dose of upadacitinib with placebo and respective 95% CI accounting for total exposure within each study via MH weights were reported
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; EOW=every other week; incidence rates; PY=person years; CI=confidence intervals
[Source: Statistical Reviewer]

In study M14-465, the EAIR for treatment emergent AEs related to infections were 75.5, 110, and 83.7 per 100 PY for placebo, upadacitinib 15 mg QD, and adalimumab 40 mg EOW respectively (Table 36). The MH-adjusted EAIR of infections comparing upadacitinib 15 mg QD vs placebo were statistically significantly higher (Est: 34.4 per 100 PY; 95% CI: 12.8, 56.1). There was an observed trend towards higher EAIR for treatment emergent AEs including serious

infections, any hepatic disorder, GI perforation, neutropenia, lymphopenia, herpes zoster comparing upadacitinib with placebo.

Table 36 Exposure Adjusted Incidence Rates for Treatment Emergent Adverse Events of Special Interests During Placebo-Controlled Period 14 Weeks, Study M14-465

	Placebo + cDMARD (N=652)	UPA 15 mg QD + cDMARD (N=650)	ADA 40 mg EOW + cDMARD (N=327)	UPA 15 mg QD Vs Placebo	ADA 40 mg EOW Vs Placebo
	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	Diff [95% CI] [2]	Diff [95% CI] [2]
Infections	117/154.9 (75.5)	165/150.1 (110)	64/76.5 (83.7)	34.4 [12.8, 56.1]	8.2 [-16.5, 32.8]
Serious Infections	5/171.1 (2.9)	10/171.7 (5.8)	4/85.6 (4.7)	2.9 [-1.5, 7.3]	1.7 [-3.5, 7.0]
Opportunistic Infection	2/171.4 (1.2)	3/171.9 (1.7)	1/85.8 (1.2)	0.6 [-2.0, 3.1]	-0.0 [-2.8, 2.8]
Malignancy	2/171.3 (1.2)	0/172.5 (-)	1/85.9 (1.2)	-1.2 [-2.8, 0.5]	-0.0 [-2.8, 2.8]
Malignancy (Excl NMSC)	1/171.4 (0.6)	0/172.5 (-)	0/86.0 (-)	-0.6 [-1.7, 0.6]	-0.6 [-1.7, 0.6]
Lymphoma	0/171.4 (-)	0/172.5 (-)	0/86.0 (-)	NA	NA
Any Hepatic Disorder	22/167.8 (13.1)	38/166.8 (22.8)	11/84.5 (13)	9.7 [0.6, 18.7]	-0.1 [-9.5, 9.4]
Any GI perforation	0/171.4 (-)	2/172.3 (1.2)	0/86.0 (-)	1.2 [-0.4, 2.8]	NA
Anemia	14/169.4 (8.3)	8/171.8 (4.7)	5/85.1 (5.9)	-3.6 [-9.0, 1.8]	-2.4 [-9.1, 4.3]
Neutropenia	2/171.3 (1.2)	9/171.1 (5.3)	1/85.9 (1.2)	4.1 [0.3, 7.9]	-0.0 [-2.8, 2.8]
Lymphopenia	8/169.9 (4.7)	11/171.0 (6.4)	2/85.5 (2.3)	1.7 [-3.3, 6.7]	-2.4 [-7.0, 2.2]
Herpes Zoster	1/171.2 (0.6)	5/172.0 (2.9)	1/86.0 (1.2)	2.3 [-0.5, 5.1]	0.6 [-2.0, 3.1]
Active/Latent TB	0/171.4 (-)	1/172.5 (0.6)	0/86.0 (-)	0.6 [-0.6, 1.7]	NA
MACE	3/171.3 (1.8)	0/172.5 (-)	1/86.0 (1.2)	-1.8 [-3.7, 0.2]	-0.6 [-3.6, 2.4]
Adjudicated VTEs	1/171.4 (0.6)	1/172.5 (0.6)	3/85.9 (3.5)	-0.0 [-1.6, 1.6]	2.9 [-1.2, 7.0]

Treatment-emergent adverse event is defined as any adverse event with an onset date on or after the first dose of study drug in period 1 and prior to the week 14 dose date in period 1 or up to 30 days after the last dose of placebo or upadacitinib and 70 days for adalimumab, if subject discontinued study drug prematurely before week 14 dosing in period 1 of the study.
1: EAIR rates were stratified by study using MH weights using total exposure within each study.
2: Difference in EAIR comparing each dose of upadacitinib with placebo and respective 95% CI accounting for total exposure within each study via MH weights were reported
NA: When there are no events in both groups, the difference and 95% CMH weighted CIs are not reported.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; EOW=every other week; incidence rates; PY=person years; CI=confidence intervals; NA=not applicable
[Source: Statistical Reviewer]

Pooled analysis integrating the placebo-controlled study period comparing 15 mg vs placebo were consistent with results observed trends observed in study M14-465 (Table 66). Risk difference for AEs related to any hepatic disorder were lower after integrating with data from M13-542 and M13-549.

During the first 12/14 weeks for the monotherapy studies, there was an observed trend towards higher MH-adjusted EAIR for the higher dose of upadacitinib monotherapy relative to the 15 mg dose (Table 37). The MH-adjusted EAIR for the various treatment emergent AEs for upadacitinib 15 mg QD monotherapy were generally consistent with MTX monotherapy the wide CIs provided some measure of large risk differences. Therefore, these clinical data alone cannot rule out the possibility of meaningful increases in the risk of important adverse events on upadacitinib 15 mg QD monotherapy.

Table 37 Exposure Adjusted Incidence Rates for Treatment Emergent Adverse Events of Special Interests During MTX-Controlled Period 14 Weeks, Studies M13-545 and M15-555

	MTX Monotherapy (N=530)	UPA 15 mg QD Monotherapy (N=534)	UPA 30 mg QD Monotherapy (N=529)	UPA 15 mg QD Vs MTX	UPA 30 mg QD Vs MTX
	E /Exposure (100PY) (IR /100 PY)[1]	E /Exposure (100PY) (IR /100 PY)	E /Exposure (100PY) (IR /100 PY)	Diff [95% CI][2]	Diff [95% CI]
Infections	127/106.4 (119)	104/113.4 (91.7)	128/111.2 (115)	-27.7 [-54.9, -0.4]	-4.3 [-33.1, 24.5]
Serious Infections	2/121.6 (1.6)	3/126.4 (2.4)	8/126.3 (6.4)	0.7 [-2.8, 4.2]	4.7 [-0.2, 9.7]
Opportunistic Infection	1/121.5 (0.8)	0/126.8 (-)	4/126.0 (3.2)	-0.8 [-2.4, 0.8]	2.3 [-1.2, 5.8]
Malignancy	2/121.6 (1.6)	3/126.8 (2.4)	0/126.7 (-)	0.7 [-2.8, 4.2]	-1.6 [-3.9, 0.6]
Malignancy (excluding NMSC)	1/121.6 (0.8)	3/126.8 (2.4)	0/126.7 (-)	1.5 [-1.6, 4.7]	-0.8 [-2.4, 0.8]
Lymphoma	0/121.7 (-)	1/126.8 (0.8)	0/126.7 (-)	0.8 [-0.8, 2.3]	0.0 [0.0, 0.0]
Any Hepatic Disorder	10/120.2 (8.3)	19/123.7 (15.4)	16/124.9 (12.8)	7.0 [-1.6, 15.7]	4.5 [-3.6, 12.6]
Any GI perforation	0/121.7 (-)	0/126.8 (-)	2/126.7 (1.6)	NA	1.6 [-0.6, 3.8]
Anemia	6/120.9 (5)	5/125.8 (4)	6/125.9 (4.8)	-1.0 [-6.3, 4.3]	-0.2 [-5.7, 5.3]
Neutropenia	3/121.3 (2.5)	8/125.6 (6.4)	14/124.4 (11.3)	3.9 [-1.3, 9.1]	8.8 [2.3, 15.3]
Lymphopenia	7/120.6 (5.8)	2/126.3 (1.6)	4/126.1 (3.2)	-4.2 [-9.0, 0.6]	-2.6 [-7.9, 2.7]
Herpes Zoster	2/121.4 (1.6)	6/126.0 (4.8)	8/126.1 (6.3)	3.1 [-1.3, 7.6]	4.7 [-0.3, 9.6]
Active/Latent TB	0/121.7 (-)	0/126.8 (-)	0/126.7 (-)	NA	NA
MACE	1/121.7 (0.8)	2/126.8 (1.6)	3/126.6 (2.4)	0.8 [-2.0, 3.5]	1.5 [-1.6, 4.7]
Adjudicated VTEs	0/121.7 (-)	1/126.7 (0.8)	1/126.7 (0.8)	0.8 [-0.8, 2.3]	0.8 [-0.8, 2.3]

1: EAIR rates were stratified by study using MH weights using total exposure within each study.
2: Difference in EAIR comparing each dose of upadacitinib and MTX monotherapy and respective 95% CI accounting for total exposure within each study via MH weights were reported
NA: When there are no events in both groups, the difference and 95% CMH weighted CIs are not reported.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; MH=Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; CI=confidence intervals; MTX=methotrexate
[Source: Statistical Reviewer]

3.3.2.2 Phase 3 Upadacitinib Doses (1 year)

In absence of a placebo or active control arm, there are limitations to comparison of upadacitinib AE rates beyond Week 12. Despite that, it is reasonable to characterize EAIR rates for rare AEs to describe the dose-response relationship for upadacitinib. In general, there was a nominally, statistically significant dose dependent increase in AEs related to serious infections, herpes zosters, and neutropenia towards the higher dose of upadacitinib relative to the low dose of upadacitinib (Table 38). After excluding patients who had crossover from other treatment groups, these AEs remained significantly higher for the 30 mg upadacitinib. There was a trend towards higher MH-adjusted EAIR of AEs related to TB reported in the lower dose of upadacitinib.

Table 38 Exposure Adjusted Incidence Rates Emergent Adverse Events of Special Interests Up to 1-yr Exposure for the Two Doses of Upadacitinib, Studies M13-542, M13-549, M15-555, and M13-545

	Any UPA 15 mg QD (N=1213)	Any UPA 30 mg QD (N=1233)	UPA 30 mg QD Vs UPA 15 mg QD	As Randomized UPA 15 mg QD (N=919)	As Randomized UPA 30 mg QD (N=913)	UPA 30 mg QD vs UPA 15 mg QD
	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	Diff [95% CI] [2]	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	Diff [95% CI] [2]
Infections	615/733.6 (83.8)	674/677.7 (99.7)	15.9 [5.9, 25.9]	479/546.6 (87.4)	517/504.3 (103)	15.4 [3.6, 27.2]
Serious Infections	38/1081.6 (3.5)	59/1049.1 (5.6)	2.1 [0.3, 3.9]	25/817.4 (3.1)	46/794.4 (5.8)	2.8 [0.7, 4.8]
Opportunistic Infection	7/1088.9 (0.6)	15/1052.7 (1.4)	0.8 [-0.1, 1.6]	6/821.9 (0.7)	13/796.0 (1.6)	0.9 [-0.2, 2.0]
Malignancy	17/1089.0 (1.6)	22/1058.1 (2.1)	0.5 [-0.6, 1.7]	13/822.1 (1.6)	18/800.9 (2.3)	0.7 [-0.7, 2.0]
Malignancy (excluding NMSC)	13/1090.2 (1.2)	14/1060.7 (1.3)	0.1 [-0.8, 1.1]	10/823.0 (1.2)	12/802.7 (1.5)	0.3 [-0.9, 1.4]
Lymphoma	1/1091.6 (0.1)	1/1061.3 (0.1)	0.0 [-0.3, 0.3]	1/824.0 (0.1)	1/803.2 (0.1)	0.0 [-0.3, 0.3]
Any Hepatic Disorder	84/1048.9 (8.0)	87/1019.9 (8.5)	0.5 [-2.0, 3.0]	64/790.3 (8.1)	70/771.5 (9.1)	1.0 [-1.9, 3.9]
Any GI perforation	1/1091.4 (0.1)	5/1061.5 (0.5)	0.4 [-0.1, 0.8]	1/823.9 (0.1)	5/803.4 (0.6)	0.5 [-0.1, 1.1]
Anemia	45/1070.2 (4.2)	62/1031.2 (6.0)	1.8 [-0.1, 3.8]	34/809.0 (4.2)	45/780.8 (5.8)	1.6 [-0.6, 3.7]
Neutropenia	37/1065.7 (3.5)	65/1022.5 (6.4)	2.9 [1.0, 4.8]	27/803.2 (3.3)	54/771.2 (7.0)	3.7 [1.4, 5.9]
Lymphopenia	13/1083.0 (1.2)	27/1045.7 (2.6)	1.4 [0.2, 2.6]	10/817.4 (1.2)	22/790.5 (2.8)	1.6 [0.2, 3.0]
Herpes Zoster	41/1073.0 (3.8)	64/1033.9 (6.2)	2.4 [0.5, 4.3]	31/810.2 (3.8)	46/782.9 (5.9)	2.1 [-0.1, 4.2]
Active/Latent TB	27/1089.2 (2.5)	19/1059.0 (1.8)	-0.7 [-1.9, 0.5]	24/822.5 (2.9)	15/802.1 (1.9)	-1.1 [-2.6, 0.4]
MACE	9/1090.1 (0.8)	12/1059.5 (1.1)	0.3 [-0.5, 1.1]	5/823.3 (0.6)	9/801.4 (1.1)	0.5 [-0.4, 1.4]
Adjudicated VTEs	5/1090.8 (0.5)	4/1060.7 (0.4)	-0.1 [-0.6, 0.5]	3/824.1 (0.4)	3/803.3 (0.4)	0.0 [-0.6, 0.6]

1: EAIR rates were stratified by study using MH weights using total exposure within each study.
2: Difference in EAIR comparing upadacitinib 30 mg with upadacitinib 15 mg and respective 95% CI accounting for total exposure within each study via MH weights were reported.
NA: When there are no events in both groups, the difference and 95% CMH weighted CIs are not reported.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; CI=confidence intervals; MH=Mantel Haenszel ; EAIR=exposure-adjusted incidence rates; E=number of patients with at least 1 event
[Source: Statistical Reviewer]

3.4 Benefit-Risk Assessment

The reader is deferred to the medical reviewer for a full comprehensive risk:benefit assessment.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Demographic Subgroups

The sample estimates of treatment effect in ACR20 response at Week 12/14 among subgroups (specifically, age, gender, race, and region) were reported using the difference in proportions analysis in each study. There were some random highs and random lows in sample estimates of subgroup treatment effects due to small sample size and large variability for some subgroups within each study. To facilitate review, we considered it reasonable to pool the cDMARD add-on studies (M13-542, M13-549, and M13-465) for the comparison between upadacitinib 15 mg QD with placebo groups to obtain a bigger sample size within each subgroup category. Despite pooling the three studies, the sample size was not sufficient to conduct multi-way subgroup analyses. For monotherapy studies, we did not consider it reasonable to pool these studies as the patient characteristics were different. We only conducted comparisons between upadacitinib 15 mg QD with the reference arm using shrinkage analysis. Comparisons for 30 mg QD was not reported since this was not the proposed dosing of interest. Shrinkage analysis was only conducted for monotherapy study M15-555 for the key demographic subgroups in this section.

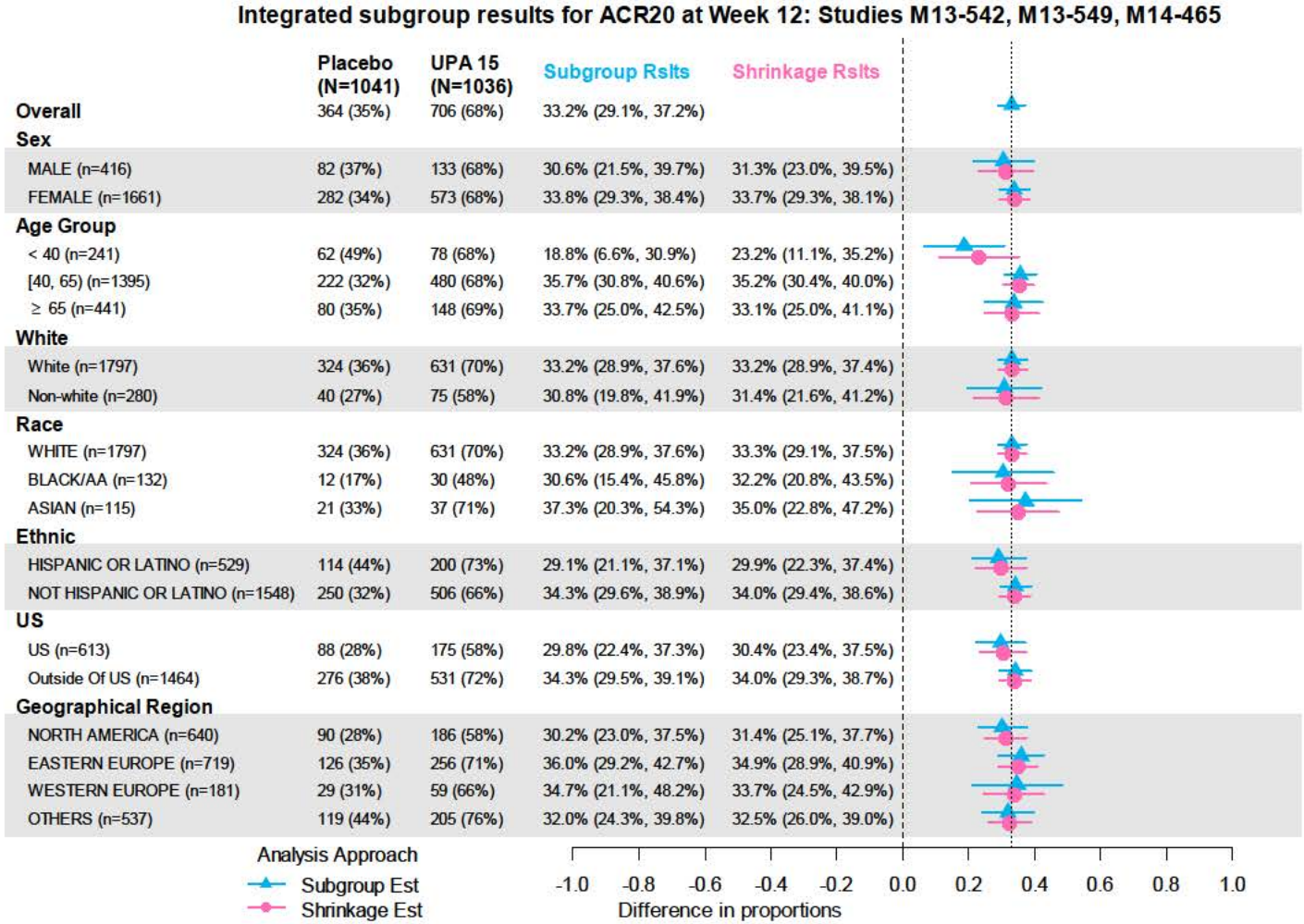
We included shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model based on summary sample estimates. The total variability in the sample estimates for the pooled studies were derived using CMH study weights. A shrinkage estimate of the subgroup treatment effect, which borrows information from the other subgroups while estimating the treatment effect for a specific subgroup, is a "weighted" average of the sample estimate and overall estimate. We used the same flat prior to derive shrinkage estimates for all subgroups. The Bayesian hierarchical model assumptions are:

For $i = 1, \ldots, k$, Y_i represents the estimated difference in ACR20 response in a subgroup level i, assume Y_i approximately $N(\mu_i, \sigma_i^2)$ where

- σ_i^2 are the estimated CMH weighted variance of the difference in ACR20 in subgroup level i
- $\mu_i \sim N(\mu, \tau^2)$
- $\mu \sim N(0, \omega^2)$, $1/\tau^2 \sim$ Gamma(0.001, 0.001)
- ω was chosen to be 4

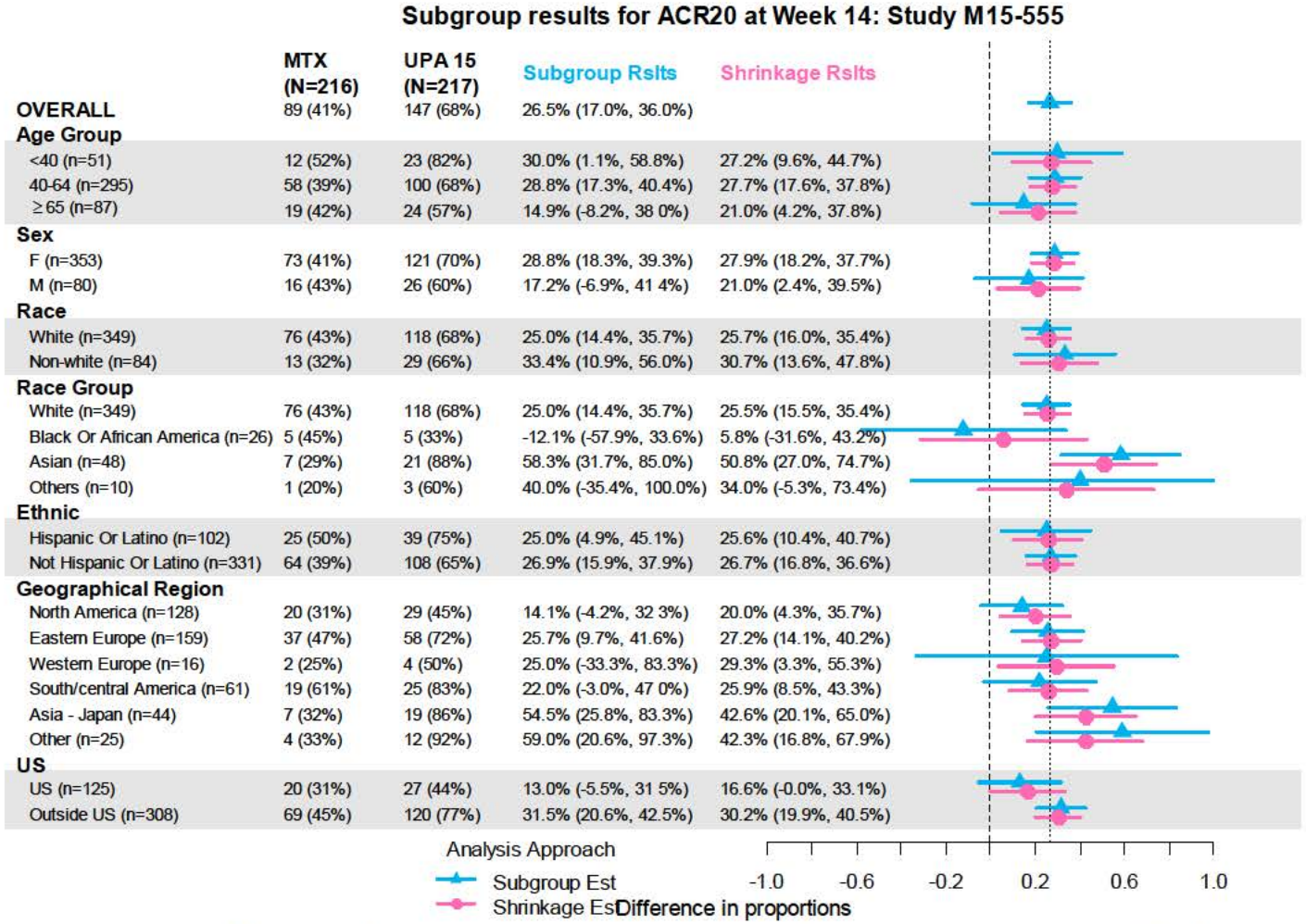
Bayesian hierarchical model was not evaluated for subgroups with small numbers. In summary, results from integrated subgroup analyses (studies M13-542, M13-549, M14-465) by demographics (age, sex, race, ethnicity, region, US only) were consistent with findings in the overall population, based on probability of ACR20 response at Week 12 and remaining on study and on randomized treatment comparing patients randomized to upadacitinib 15 mg QD vs placebo (Figure 7). Similarly, for the monotherapy study M15-555, the findings were also consistent with findings in the overall population (Figure 8).

Figure 7 Demographic Subgroup Results for the Probability of ACR20 Response at Week 12, Studies M13-542, M13-549, and M14-465



Estimates for the difference in binomial proportions are represented as circles or triangles. Frequentist 95% CI and shrinkage 95% PI are shown as blue and pink lines for each subgroup category respectively. Shrinkage results for race subgroups were not reported since there were insufficient numbers in Others category.
Abbreviations: ACR=American College of Rheumatology; UPA=upadacitinib; BMI body mass index; Est=estimate; CI=confidence intervals; PI=prediction intervals; Rslts=results; N=total number of patients randomized; n=number of patients in sub-category; AA=African American
[Source: Statistical Reviewer]

Figure 8 Demographic Subgroup Results for the Probability of ACR20 response at Week 14, Study M15-555



Point estimates for the difference in binomial proportions are represented as circles or triangles. Frequentist 95% CI and shrinkage 95% PI are show in blue and pink lines for each subgroup category respectively.

Abbreviations: ACR=American College of Rheumatology; UPA=upadacitinib; BMI body mass index; Est=estimate; CI=confidence intervals; PI=prediction intervals; Rslts=results; N=total number of patients randomized; n=number of patients in sub-category; AA=African American

[Source: Statistical Reviewer]

4.2 Subgroup Results by Individual Studies

Shrinkage analysis was not reported for the individual studies. When possible, I reported comparisons between upadacitinib 30 mg QD with the reference arm.

In summary, subgroup analyses by demographics characteristics were consistent with the findings in the overall population based on the probability of ACR20 response at primary efficacy time point in all studies except M13-545 (Figure 22, Figure 23, Figure 24, Figure 25). In study M13-545, subgroup analyses by demographic subgroups (age, sex, race, ethnicity, region, US only) were consistent with findings in the overall population, based on probability of ACR50 response at Week 12 and remaining on study and on randomized treatment comparing patients randomized to upadacitinib 15 mg QD monotherapy vs MTX monotherapy (Figure 26).

Subgroup analyses by key baseline disease characteristics were consistent with the findings in the overall population based on the probability of ACR20 response at primary efficacy time point for all studies except M13-545 (Figure 27, Figure 28, Figure 29, Figure 30). In study M13-545, subgroup analyses by key baseline disease characteristics subgroups were consistent with findings in the overall population, based on probability of ACR50 response at Week 12 and remaining on study and on randomized treatment comparing patients randomized to upadacitinib 15 mg QD monotherapy vs MTX monotherapy (Figure 31).

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

During the review, we identified the key statistical issues.

- Primary multi-component endpoint ACR20 (or ACR50 for study M13-545) results

Comparisons of proportions of responders were affected by the proportions of patients discontinuing from the randomized treatment prior to the primary endpoint visit because those discontinued were considered as non-responders. Therefore, it is difficult to determine whether observed treatment differences using NRI are due to treatment effects on patient signs and symptoms or due to differences in the proportion of patients on originally randomized treatment. Nevertheless, we note that the discontinuation rates were low across arms in all 5 studies and the observed proportions of responders in treatment arms were much larger than in the placebo arm. Supportive tipping point analyses conducted for ACR20 at the primary timepoint provided reassurance that the results were convincing despite the assumptions about the missing data. These results based on ACR50 at Week 12 for study M13-545 were similar.

Furthermore, because ACR itself is a multi-component endpoint, it is also vital to understand whether the results were consistent across the individual components (such as CRP), and whether the individual component results were affected by the presence of missing data. Results for the individual components of ACR at primary timepoint were in general consistently in favor of the upadacitinib arms. Furthermore, there was convincing evidence of benefit for the important supportive patient-reported outcome measure of functional ability, HAQ-DI score, in all studies, and tipping point analyses for this endpoint indicated that results were robust to missing data assumptions.

- Missing data

The presence of missing data can affect the interpretation of the study results. In general, the amount of missing data, i.e., due to lost-to-follow-up, withdrawal of informed consent to continue further participation, in these studies were small and balanced across arms within each study at primary efficacy timepoint. Tipping point sensitivity analyses provided reassurance of the robustness of the applicant's results to violations in assumptions about the missing data.

- Radiographic progression based on treatment policy estimand.

Two key estimands of interest are considered in this review, namely, the treatment policy estimand, i.e., the difference in progression regardless of adherence or escape, and an alternative estimand, the difference in progression in the absence of any placebo or upadacitinib crossover.

In study M14-465, I conducted supportive analyses based on all observed data within the Week 26 visit window, analyses targeting the proportion of patients with no radiographic progression at Week 26, and supportive analyses that target the rate of change in radiograph progression over the 26-week period. These additional results were convincing and provide statistically significant evidence conferring benefit for upadacitinib 15 mg QD relative to placebo.

In study M14-465, 10% of the patients did not have Week 26 radiograph evaluation. Because violations in assumptions about missing data could potentially impact the observed significant results, supportive tipping point sensitivity analyses were conducted. The results were found to be robust in various tipping point analyses, thereby providing further assurance of the applicant's significant findings.

The applicant's pre-specified linear extrapolation analysis might be considered to target the alternative estimand, i.e., the difference in progression in the absence of escape on placebo. However, I have concerns with the use of linear extrapolation because it relies on a strong and unverifiable assumption and does not appropriately account for uncertainty. Therefore, I conducted additional analyses comparing slopes of progression up to Week 26 in the absence of escape for placebo; these analyses did not rely on the single imputation step in the linear extrapolation approach and further excluded data after placebo or upadacitinib patients crossed over to other treatments.

In study M13-545, both doses of upadacitinib as a monotherapy were evaluated relative to MTX monotherapy. I conducted analyses based on all observed data within the Week 24 visit window and analyses targeting the proportion of patients with no radiographic progression at Week 24. These additional results were convincing and provide statistically significant evidence conferring benefit for either doses of upadacitinib as a monotherapy relative to MTX monotherapy.

In summary, results from analyses targeting the treatment policy estimand and the estimand in the absence of escape on placebo, as well as tipping point sensitivity analyses, were convincing to conclude that there is sufficient evidence of delay in radiographic progression with upadacitinib 15 mg QD.

- Evidence to support upadacitinib 15 mg QD

Based on placebo-controlled studies (M13-542, M13-549, and M14-465), there is statistical evidence of efficacy for the proposed upadacitinib 15 mg QD over placebo based on the primary, secondary endpoints findings, as well as the various sensitivity analyses conducted. Although

the higher dose of upadacitinib was not used to seek approval, we included description of the dose response between the two dosing regimens based on study M13-542 and M13-549. Comparisons between the two dosing regimens did not suggest a noticeable and consistent trend in favor of a dose. As supportive data, the applicant included study M13-545 in the submission to provide additional efficacy support. In the four studies that evaluated two doses, there was no consistent dose response relationship for 15 mg QD vs 30 mg QD based on review of the efficacy endpoints.

Based on the integrated safety analyses during the placebo-controlled or MTX-controlled period, there was an observed dose response relationship for key treatment emergent AEs such as infections, and serious infections, common to the JAK class. In absence of an active comparator, we also observed higher adjusted EAIR for infections, serious infections, herpes zosters, and neutropenia towards the higher dose of upadacitinib relative to the low dose when evaluating exposure up to one year. There was an observed higher EAIR of patients developing TB in the 15 mg dose relative to the higher dose.

The benefit-risk assessment would need to include a more comprehensive safety assessment (discussed in the clinical), as well as pharmacokinetic exposure not considered in this review. In summary, from an efficacy perspective and safety analyses limited to the discussion presented in 3.3, the applicant's proposed 15 mg QD dose is reasonable.

- (b) (4)

(b) (4)

- Evidence to support the indication (b) (4)

The applicant proposed (b) (4) Based on study M13-545, there was evidence of efficacy to support the use of upadacitinib as a potential

monotherapy for treatment of MTX-naïve patients. The primary efficacy and key secondary endpoints were further robust to missing data assumptions.

In summary, the results from studies M13-542, M13-549, M14-465 and M15-555 generally supported the use of upadacitinib 15 mg QD as a treatment in RA patients who are inadequate responders to MTX.

5.2 Collective Evidence

The collective evidence from cDMARD add-on studies M13-542, M13-549, and M14-465 supports the efficacy of upadacitinib 15 mg QD for the treatment of adult patients with RA patients who are inadequate response or intolerance to MTX. In addition, the results for upadacitinib 30 mg QD were also supportive of the efficacy profile. Results from additional sensitivity analyses conducted, and results for additional important endpoints such as patient-reported outcomes, were supportive of the findings of effectiveness for the upadacitinib 15 mg QD.

5.3 Conclusions and Recommendations

There is substantial evidence of efficacy based on the primary and secondary endpoint results for upadacitinib 15 mg QD dosing regimen in placebo-controlled studies, supporting the indication of adult patients with moderately to severely active rheumatoid arthritis who have had an inadequate response or are intolerant to MTX. The results and conclusions from the monotherapy studies provided further support of the efficacy of upadacitinib 15 mg QD.

The safety profile based on the placebo-controlled comparisons were consistent with known safety signals associated with JAK inhibitors. However, the lack of long term active-controlled comparison, in absence of crossing over between active and investigational product, and the limitations of the study designs provided limited interpretation of controlled safety comparisons to characterize AE rates beyond Week 14. These efficacy findings should be incorporated with safety results to evaluate the benefit-risk of upadacitinib for treatment of RA.

Based on efficacy findings of this review and discussion with the clinical review team, I recommend approval of upadacitinib 15 mg QD for treatment of adult patients with moderately to severely active rheumatoid arthritis in patients who have had an inadequate response or intolerance to methotrexate.

5.4 Labeling Recommendations

We included the following labelling recommendations in the proposed label.

Section 6

- We recommend reporting of exposure adjusted incidence rates stratified by study for all AEs reported in the label per described in section 3.3.2

- We recommend presenting the AE rates by integrating the 15 mg dose with placebo in the placebo-controlled studies. As supportive information, we recommend presenting the AE rates for placebo, 15 mg, 30 mg based on integrating studies M13-542 and M13-549.

- We recommend presenting the AE rates for MTX-controlled period for MTX monotherapy, 15 mg monotherapy, 30 mg monotherapy based on integrating studies M13-545 and M15-555.

- Present the 12-month exposure controlled by integrating studies M13-542, M13-545, M13-549, and M15-555 to describe dose related AE comparison between 15 and 30 mg.

Section 14

- The presentation of the study designs in a table is useful but it does not include information on whether patients with inadequate response crossover to other treatments. We therefore recommend that a shorter summary of such information be included instead of the table.

- We recommend removing these statements: (b) (4)

- We recommend presentation of data for the proportion of patients with DAS28CRP < 2.8 with active joints at the primary efficacy timepoint to be consistent with the primary endpoint.

- Removal (b) (4) in Table 7, 8, 9

- Removal (b) (4) in Section 14.

- Recommend reporting of the random coefficient model for the radiographic endpoint for study M14-465. If study M13-545 is included, we recommend providing a conclusion that the results were consistent with study M14-465.

- We do not recommend including study M13-545 in the label because the indication is not approved for the MTX naïve patient population. If included, a disclaimer should be included that there is limitation of use in this patient population due to unestablished risk:benefit profile available in this review.

6 APPENDICES

6.1 Brief Description of Study Endpoints Evaluated in this Review

An ACR20 response is defined as a reduction of at least 20% in the number of swollen joints (out of 66), a reduction of at least 20% in the number of tender joints (out of 68), and a reduction of 20% in three of the following five remaining ACR Core Set measures. The specific components of ACR are physician global assessment, patient global assessment, patient assessment of pain, C-reactive protein or erythrocyte sedimentation rate, and health assessment questionnaire disability index (HAQ-DI). Fifty and seventy percent improvement (ACR50 and ACR70) are similarly calculated using these higher levels of improvement.

The HAQ-DI is a self-administered questionnaire that measures the degree of difficulty a person has in accomplishing tasks in 8 functional areas (dressing, arising, eating, walking, hygiene, reaching, gripping, and errands and chores) over the past week. Responses on a four-level difficulty scale in each functional area were scored from 0, indicating no difficulty, to 3, indicating inability to perform a task in that area.

The DAS28 is a composite index of RA disease activity which incorporates the number of tender and swollen joints (out of 28 possible), a patient global assessment of disease activity (0-100 mm visual analog scale), and CRP results. An alternative equation is available for use with ESR results. These variables are summed[1] and weighted mathematically into a single numerical value ranging from 0 to 10.

Comparing the DAS28 and the ACR response criteria, beyond the differences in number of maximum tender or swollen joints counted (e.g. DAS28 does not include the joints of the feet), additional variables of physician global assessment, patient pain, and HAQ-DI score are incorporated into the ACR response criteria. The DAS28 has additional utility in measuring the level of disease activity at a given time point, whereas the ACR response is calculated as improvement in the variables over a set period of time. A DAS28 score >5.1 at a time-point is used to describe high disease activity, and <3.2 of low disease activity. A score of <2.6 has been used to describe an even lower threshold of disease activity.

The clinical disease activity index (CDAI) is composite index that can be computed by linearly summing up tender joint counts (out of 28), swollen joint counts (out of 28), physician global assessment, and patient global assessment. The simplified disease activity index (SDAI) is an alternative composite index which incorporates similar components in CDAI by summing them linearly into a single numerical value.[2] The ACR/EULAR Boolean remission is defined when a patient satisfies the following criteria at a specific visit: (1) Tender joint count ≤ 1 (based on 28 joints); (2) Swollen joint count ≤ 1 (based on 28 joints); (3) CRP ≤ 1 mg/dL; and (4) Patient global assessment of disease activity ≤ 10 (mm). We note that CDAI, SDAI or ACR/EULAR described an alternative summary measure of components that are captured in DAS28 and ACR. Therefore, for review purposes, I did not include them in the review.

---

[1] DAS28(CRP) is a continuous outcome derived by combining tender joints (28 joints) or TJC28, swollen joints (28 joints) or SJC28, CRP (mg/L), and patient's global assessment of disease activity [GH] by the following equation: DAS28(CRP) = 0.56 ×√TJC28 + 0.28 ×√SJC28 +0.36 × logarithm (CRP + 1) + 0.014 × GH + 0.96.

[2] SDAI is calculated via the following equation: TJC28 + SJC28 + PtGA (cm) + PhGA (cm) + hsCRP (mg/dL).

The van der Heijde-modified total sharp score (mTSS) is used to evaluate radiographic progression of structural joint damage. This instrument scores the presence of erosions in the joints of the hands and feet, and the presence of joint space narrowing (JSN) in the hands, wrists, and feet. The scores for each feature for the individual joints are summed. Erosions are assessed at 16 locations in each hand and wrists and 12 locations in each foot, using a 6-point scale from 0 to 5. Scores are derived based on the number and size of discrete erosions in each location, but are summed to a maximum of 5. Thus, the maximum erosion score for the hands/wrists is 160, and the maximum erosion score for the feet is 120, for a maximum total erosion score of 280. JSN scores are based on 15 locations in each hand and wrist and 6 locations in each foot, scored using a 5-point scale from 0 to 4: 0=normal; 1=focal or minimal and generalized narrowing; 2=generalized narrowing <50%; 3=generalized narrowing >50% or subluxation; and 4=ankylosis or complete dislocation. The maximum total JSN for the hands/wrists is 120, and the maximum total JSN for the feet is 48, for a maximum total JSN score of 168. Therefore, the theoretical maximum mTSS is 448, although the actual clinical range in RA drug development trials is typically much lower because a given individual typically only has a fraction of his or her joints affected by radiographically evident damage.

The short form health survey (SF-36) is an instrument used to measure health-related quality of life or general health status. It consists of eight subscales that are scored individually: physical functioning (10 items), role-physical (four items), bodily pain (two items), general health (five items), vitality (four items), social functioning (two items), role-emotional (three items), and mental health (five items). Two summary scores, the PCS (based on physical functioning, role-physical, bodily pain, and general health) and the mental component summary (MCS) (based on vitality, social functioning, role-emotional, and mental health), are computed.

Functional assessment of chronic illness therapy-fatigue (FACIT-F) was evaluated in study M13-549, M14-465, and M13-545. The FACIT-F is a 13-item questionnaire that measures an individual's level of fatigue over the past week on a four-point Likert scale. The range of the score is from 0 to 52 and the higher scores, the better the quality of life. The scores are added together and pro-rated according to the number of items answered. If less than 7 items are answered, the score will not be computed. The interpretation of FACIT-Fatigue scores is such that a higher score indicates less fatigue, with a range of possible scores of 0-52, with 0 being the worst possible score and 52 the best. To obtain the 0-52 score, 11 of the 13 items require recoding such that the response value of 0 represents greatest fatigue and the response value of 4 represents no fatigue. The response values of all 13 items are added with equal weight to obtain the total score.

6.2 Tipping Point Analysis Methodology[3]

The goal is to evaluate the potential effect of violations in assumptions about missing data on the reliability of conclusions. Suppose that outcomes Y are independently distributed on the control and test drug arms. The parameter of interest is the difference in means θ. Consider the following parameterization and notation to describe the probabilities of completing the study (non-missingness), the true means in completers and dropouts, and the numbers of completers and total patients on the two treatment arms:

[3] Source: Gregory Levin

Table 39 Parameters and Notation for Tipping Point Analysis in Presence of Missing Data

Arm	Probability of non-missing	Mean among completers	Mean among dropouts	Number of completers	Sample size per arm
Placebo	π_c	μ_c	$\mu_c + \delta_c$	N_c	n_c
Treated	π_t	μ_t	$\mu_t + \delta_t$	N_t	n_t

Given this parameterization, the target of inference is
$\theta = [\pi_t\mu_t + (1 - \pi_t)(\mu_t + \delta_t)] - [\pi_c\mu_c + (1 - \pi_c)(\mu_c + \delta_c)] \equiv \mu_t + (1 - \pi_t)\delta_t - [\mu_c + (1 - \pi_c)\delta_c]$ An analysis based on completers will provide reliable inference on θ if the missing-at random assumption, i.e., the assumption that $\delta_c = \delta_t = 0$, is valid. We will perform sensitivity analyses that allow for the possibility that outcomes among dropouts are not missing-at-random by performing inference under different assumed values of the parameters δ_c and δ_t.

Denote M_{ij} to be an indicator that patient j on treatment i is a completer, i.e., his or her outcome is observed where $i = c,t$, and $j = 1,\hat{a}\breve{<}, n_i$. By assuming fixed values of sensitivity parameters δ_c and δ_t, an estimator of θ can be represented by

$$\hat{\theta} = \hat{\mu}_t + (1 - \hat{\pi}_t)\delta_t - [\hat{\mu}_c + (1 - \hat{\pi}_c)\delta_c]$$

where $\hat{\mu}_i = \frac{1}{N_i}\sum_{k=1}^{n_i} Y_{ik}|M_{ik} = 1$ is the sample mean in the completers and $\hat{\pi}_i = \frac{N_i}{n_i} \equiv \sum_{k=1}^{n_i} M_{ik}/n_i$ is the sample proportion of completers on the treatment arm i, with i taking values c or t.

The test statistic can be constructed as follows:

$$\frac{\hat{\theta} - \theta}{\sqrt{\frac{s_t^2}{N_t} + \frac{s_c^2}{N_c} + \frac{\delta_t^2\hat{\pi}_t(1 - \hat{\pi}_t)}{n_t} + \frac{\delta_c^2\hat{\pi}_c(1 - \hat{\pi}_c)}{n_c}}}$$

where s_i^2 is the sample variance of the outcome. Under suitable conditions, the sampling test statistic is asymptotically normal with mean 0 and standard deviation 1.

The Wald-based 100 (1-α) % confidence interval of the form
$\hat{\theta} \pm z_{1-\alpha/2}\sqrt{\frac{s_t^2}{N_t} + \frac{s_c^2}{N_c} + \frac{\delta_t^2\hat{\pi}_t(1 - \hat{\pi}_t)}{n_t} + \frac{\delta_c^2\hat{\pi}_c(1 - \hat{\pi}_c)}{n_c}}$ can be constructed where z_q is the q quantile of the standard normal distribution.

6.3 Additional Tables

Table 40 Baseline Characteristics for the Components of ACR[1] and DAS28(CRP)[2], Study M13-542

		PBO + cDMARD (N=169)	UPA 15 mg + cDMARD (N=164)	UPA 30 mg + cDMARD (N=165)
SJC [2] (Out of 28)	*Mean (SD)*	11.4 (5.6)	11.9 (5.8)	11.6 (5.7)
	Min -Max	2 - 27	0 - 26	0 - 28
	n	(n=169)	(n=164)	(n=165)
SJC [1] (Out of 66)	*Mean (SD)*	16.3 (9.6)	17.0 (10.8)	17.2 (11.4)
	Min -Max	4 - 57	6 - 62	6 - 64
	n	(n=169)	(n=164)	(n=165)
TJC [2] (Out of 28)	*Mean (SD)*	15.9 (7.3)	16.1 (7.3)	15.4 (6.5)
	Min -Max	2 - 28	2 - 28	1 - 28
	n	(n=169)	(n=164)	(n=165)
TJC [1] (Out of 68)	*Mean (SD)*	28.5 (15.3)	27.8 (16.3)	27.3 (15.2)
	Min -Max	6 - 66	6 - 68	6 - 66
	n	(n=169)	(n=164)	(n=165)
HAQ-DI[1]	*Mean (SD)*	1.6 (0.6)	1.7 (0.6)	1.6 (0.6)
	Min -Max	0 - 3	0 - 3	0 - 3
	n	(n=166)	(n=163)	(n=161)
Physician Global Assessment[1]	*Mean (SD)*	66.9 (16.9)	68.7 (16.6)	66.4 (15.6)
	Min -Max	13 - 100	15 - 100	21 - 100
	n	(n=161)	(n=157)	(n=157)
Patient Global Assessment [1,2]	*Mean (SD)*	66.3 (22.7)	67.2 (19.6)	64.7 (21.0)
	Min -Max	9 - 100	16 - 100	16 - 100
	n	(n=166)	(n=163)	(n=163)
Patient Pain VAS[1]	*Mean (SD)*	68.9 (21.0)	68.2 (19.8)	65.3 (20.7)
	Min -Max	13 - 100	4 - 100	13 - 100
	n	(n=166)	(n=163)	(n=161)
CRP (mg/L) [1,2]	*Mean (SD)*	16.3 (21.1)	16.2 (18.6)	16.0 (21.2)
	Min -Max	0.3 - 150.0	0.2 - 129.0	0.6 - 137.0
	n	(n=169)	(n=164)	(n=165)

Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP = C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI = Health Assessment Questionnaire - Disability Index; DAS28=disease activity score using 28 joints based on CRP; cDMARD=conventional disease modifying anti-rheumatic drugs

[Source: Statistical Reviewer]

Table 41 Baseline Characteristics for the Components of ACR[1] and DAS28(CRP)[2], Study M13-549

		PBO + cDMARD (N=221)	UPA 15 mg + cDMARD (N=221)	UPA 30 mg + cDMARD (N=219)
SJC [2] (Out of 28)	*Mean (SD)*	10.8 (5.1)	11.0 (5.4)	11.2 (5.5)
	Min -Max	1 - 27	1 - 27	0 - 28
	n	(n=221)	(n=221)	(n=219)
SJC [1] (Out of 66)	*Mean (SD)*	15.4 (9.2)	16.0 (10.0)	16.2 (10.6)
	Min -Max	3 - 62	1 - 61	6 - 60
	n	(n=221)	(n=221)	(n=219)
TJC [2] (Out of 28)	*Mean (SD)*	14.2 (6.6)	14.3 (6.7)	15.0 (6.8)
	Min -Max	1 - 28	0 - 28	1 - 28
	n	(n=221)	(n=221)	(n=219)
TJC [1] (Out of 68)	*Mean (SD)*	24.7 (15.0)	25.2 (13.8)	26.2 (14.3)
	Min -Max	2 - 66	6 - 64	6 - 68
	n	(n=221)	(n=221)	(n=219)
HAQ-DI[1]	*Mean (SD)*	1.4 (0.6)	1.5 (0.6)	1.5 (0.6)
	Min -Max	0 - 3	0 - 3	0 - 3
	n	(n=221)	(n=216)	(n=219)
Physician Global Assessment[1]	*Mean (SD)*	64.4 (17.7)	64.3 (16.2)	63.0 (18.0)
	Min -Max	8 - 100	26 - 100	13 - 100
	n	(n=211)	(n=209)	(n=213)
Patient Global Assessment [1,2]	*Mean (SD)*	60.3 (20.5)	63.1 (21.9)	62.8 (20.3)
	Min -Max	2 - 100	5 - 100	0 - 100
	n	(n=221)	(n=217)	(n=219)
Patient Pain VAS[1]	*Mean (SD)*	61.5 (20.8)	64.1 (19.5)	64.0 (19.8)
	Min -Max	0 - 100	8 - 100	6 - 100
	n	(n=221)	(n=217)	(n=219)
CRP (mg/L) [1,2]	*Mean (SD)*	12.6 (14.0)	16.6 (19.2)	14.8 (16.9)
	Min -Max	0.3 - 104.0	0.2 - 108.0	0.8 - 122.0
	n	(n=221)	(n=221)	(n=219)

Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP = C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI = Health Assessment Questionnaire - Disability Index; DAS28=disease activity score using 28 joints based on CRP; cDMARD=conventional disease modifying anti-rheumatic drugs

[Source: Statistical Reviewer]

Table 42 Baseline Characteristics for the Components of ACR[1] and DAS28(CRP)[2], Study M14-465

		PBO + cDMARD (N=651)	UPA 15 mg + cDMARD (N=651)	ADA 40 mg EOW + cDMARD (N=327)
SJC [2] (Out of 28)	*Mean (SD)*	11.8 (5.3)	11.4 (5.6)	11.7 (5.5)
	Min -Max	0 - 28	0 - 28	1 - 28
	n	(n=651)	(n=651)	(n=327)
SJC [1] (Out of 66)	*Mean (SD)*	16.2 (9.0)	16.6 (10.3)	16.3 (9.2)
	Min -Max	3 - 55	1 - 61	4 - 64
	n	(n=651)	(n=651)	(n=327)
TJC [2] (Out of 28)	*Mean (SD)*	15.4 (6.8)	15.0 (6.9)	15.1 (7.0)
	Min -Max	0 - 28	0 - 28	0 - 28
	n	(n=651)	(n=651)	(n=327)
TJC [1] (Out of 68)	*Mean (SD)*	26.0 (14.3)	26.4 (15.1)	26.4 (15.2)
	Min -Max	6 - 68	5 - 68	6 - 68
	n	(n=651)	(n=651)	(n=327)
HAQ-DI[1]	*Mean (SD)*	1.6 (0.6)	1.6 (0.6)	1.6 (0.6)
	Min -Max	0 - 3	0 - 3	0 - 3
	n	(n=649)	(n=646)	(n=325)
Physician Global Assessment[1]	*Mean (SD)*	66.0 (18.2)	65.6 (17.1)	65.1 (17.6)
	Min -Max	4 - 100	7 - 100	12 - 99
	n	(n=620)	(n=616)	(n=305)
Patient Global Assessment [1,2]	*Mean (SD)*	63.8 (21.5)	64.3 (21.8)	65.8 (21.1)
	Min -Max	0 - 100	1 - 100	0 - 100
	n	(n=649)	(n=647)	(n=324)
Patient Pain VAS[1]	*Mean (SD)*	65.0 (20.7)	65.7 (21.0)	66.2 (20.5)
	Min -Max	0 - 100	1 - 100	0 - 100
	n	(n=649)	(n=647)	(n=325)
CRP (mg/L) [1,2]	*Mean (SD)*	18.0 (21.5)	17.9 (22.5)	19.8 (21.5)
	Min -Max	0.2 - 198.0	0.2 - 174.0	0.2 - 142.0
	n	(n=651)	(n=651)	(n=327)

Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP = C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI = Health Assessment Questionnaire - Disability Index; DAS28=disease activity score using 28 joints based on CRP; cDMARD=conventional disease modifying anti-rheumatic drugs

[Source: Statistical Reviewer]

Table 43 Baseline Characteristics for the Components of ACR[1] and DAS28(CRP)[2], Study M15-555

		MTX Monotherapy (N=216)	UPA 15 mg Monotherapy (N=217)	UPA 30 mg Monotherapy (N=215)
SJC [2] (Out of 28)	*Mean (SD)*	11.3 (6.0)	11.1 (5.8)	11.6 (5.7)
	Min -Max	0 - 28	2 - 28	0 - 28
	n	(n=216)	(n=217)	(n=215)
SJC [1] (Out of 66)	*Mean (SD)*	16.9 (11.5)	16.4 (10.9)	16.9 (10.2)
	Min -Max	6 - 66	6 - 64	3 - 59
	n	(n=216)	(n=217)	(n=215)
TJC [2] (Out of 28)	*Mean (SD)*	14.5 (7.4)	14.2 (6.8)	14.2 (7.5)
	Min -Max	0 - 28	1 - 28	0 - 28
	n	(n=216)	(n=217)	(n=215)
TJC [1] (Out of 68)	*Mean (SD)*	25.2 (16.0)	24.5 (15.1)	24.8 (15.2)
	Min -Max	6 - 68	6 - 68	3 - 67
	n	(n=216)	(n=217)	(n=215)
HAQ-DI[1]	*Mean (SD)*	1.5 (0.7)	1.5 (0.7)	1.5 (0.7)
	Min -Max	0 - 3	0 - 3	0 - 3
	n	(n=216)	(n=216)	(n=215)
Physician Global Assessment[1]	*Mean (SD)*	62.1 (17.5)	65.7 (18.5)	62.6 (17.8)
	Min -Max	17 - 100	17 - 100	23 - 100
	n	(n=200)	(n=209)	(n=202)
Patient Global Assessment [1,2]	*Mean (SD)*	59.6 (21.8)	62.2 (22.3)	59.4 (22.8)
	Min -Max	2 - 100	3 - 100	0 - 100
	n	(n=216)	(n=216)	(n=215)
Patient Pain VAS[1]	*Mean (SD)*	62.5 (21.3)	62.3 (22.5)	61.9 (22.1)
	Min -Max	8 - 100	7 - 100	4 - 100
	n	(n=216)	(n=216)	(n=215)
CRP (mg/L) [1,2]	*Mean (SD)*	14.5 (17.3)	14.0 (16.5)	16.3 (20.8)
	Min -Max	0.2 - 128.0	0.2 - 122.0	0.4 - 136.0
	n	(n=216)	(n=217)	(n=215)

Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP = C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI = Health Assessment Questionnaire - Disability Index; DAS28=disease activity score using 28 joints based on CRP
[Source: Statistical Reviewer]

Table 44 Baseline Characteristics for the Components of ACR[1] and DAS28(CRP)[2], Study M13-545

		MTX Monotherapy (N=314)	UPA 15 mg Monotherapy (N=317)	UPA 30 mg Monotherapy (N=317)
SJC [2] (Out of 28)	*Mean (SD)*	11.8 (6.0)	11.8 (5.8)	11.4 (5.7)
	Min -Max	0 - 28	1 - 28	0 - 28
	n	(n=314)	(n=317)	(n=314)
SJC [1] (Out of 66)	*Mean (SD)*	16.9 (10.6)	16.9 (10.3)	15.7 (9.7)
	Min -Max	5 - 60	6 - 58	0 - 59
	n	(n=314)	(n=317)	(n=314)
TJC [2] (Out of 28)	*Mean (SD)*	15.3 (7.2)	15.0 (6.9)	14.8 (7.2)
	Min -Max	1 - 28	1 - 28	1 - 28
	n	(n=314)	(n=317)	(n=314)
TJC [1] (Out of 68)	*Mean (SD)*	26.4 (16.1)	25.4 (14.4)	25.2 (15.0)
	Min -Max	3 - 68	4 - 68	2 - 68
	n	(n=314)	(n=317)	(n=314)
HAQ-DI[1]	*Mean (SD)*	1.6 (0.7)	1.6 (0.7)	1.5 (0.7)
	Min -Max	0 - 3	0 - 3	0 - 3
	n	(n=314)	(n=317)	(n=311)
Physician Global Assessment[1]	*Mean (SD)*	68.7 (16.4)	67.1 (17.0)	65.3 (16.6)
	Min -Max	12 - 100	8 - 100	11 - 100
	n	(n=299)	(n=301)	(n=304)
Patient Global Assessment [1,2]	*Mean (SD)*	65.8 (21.5)	66.6 (22.0)	64.9 (21.6)
	Min -Max	0 - 100	2 - 100	0 - 100
	n	(n=314)	(n=317)	(n=311)
Patient Pain VAS[1]	*Mean (SD)*	65.7 (21.5)	68.4 (20.6)	65.3 (21.5)
	Min -Max	1 - 100	6 - 100	0 - 100
	n	(n=314)	(n=317)	(n=311)
CRP (mg/L) [1,2]	*Mean (SD)*	21.2 (22.1)	23.0 (27.4)	19.4 (22.6)
	Min -Max	0.2 - 149.0	0.2 - 207.0	0.2 - 159.0
	n	(n=314)	(n=317)	(n=314)

Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; CRP = C-reactive protein; SJC=swollen joint counts; TJC=tender joint counts; VAS=visual analogue scale; HAQ-DI = Health Assessment Questionnaire - Disability Index; DAS28=disease activity score using 28 joints based on CRP
[Source: Statistical Reviewer]

Table 45 Baseline Radiograph Scores and Components, Study M14-465

		Placebo + cDMARD (N=651)	UPA 15 MG QD + cDMARD (N=651)	ADA 40 MG EOW + cDMARD (N=327)
mTSS	*Mean (SD)*	35.9 (51.7)	34.0 (50.1)	34.5 (47.1)
	Min -Max	0.0 - 324.5	0.0 - 304.5	0.0 - 249.5
	n	(n=649)	(n=644)	(n=326)
Erosion Score	*Mean (SD)*	17.0 (27.4)	16.5 (26.4)	15.4 (23.1)
	Min -Max	0.0 - 192.0	0.0 - 168.0	0.0 - 129.0
	n	(n=649)	(n=644)	(n=326)
Joint Space Narrowing	*Mean (SD)*	18.9 (26.1)	17.5 (25.1)	19.2 (25.8)
	Min -Max	0.0 - 134.5	0.0 - 136.5	0.0 - 125.0
	n	(n=649)	(n=644)	(n=326)
Proportion with mTSS = 0	*Count (%)*	79 (12%)	90 (14%)	44 (13%)

Counts (%): Counts and percentages relative to N in parenthesis.
Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; mTSS=modified total sharp score; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

Table 46 Baseline Radiograph Scores and Components, Study M13-545

		MTX Monotherapy (N=314)	UPA 15 mg Monotherapy (N=317)	UPA 30 mg Monotherapy (N=317)
mTSS	*Mean (SD)*	13.3 (30.5)	18.1 (38.2)	17.2 (38.3)
	Min -Max	0.0 - 260.5	0.0 - 206.5	0.0 - 243.5
	n	(n=309)	(n=309)	(n=309)
Erosion Score	*Mean (SD)*	6.1 (15.5)	8.6 (19.3)	8.0 (18.9)
	Min -Max	0.0 - 147.0	0.0 - 123.0	0.0 - 126.5
	n	(n=309)	(n=309)	(n=309)
Joint Space Narrowing	*Mean (SD)*	7.2 (16.1)	9.6 (20.1)	9.3 (20.3)
	Min -Max	0.0 - 113.5	0.0 - 103.0	0.0 - 118.0
	n	(n=309)	(n=309)	(n=309)
Proportion with mTSS = 0	*Count (%)*	100 (32%)	108 (34%)	103 (33%)

Counts (%): Counts and percentages relative to N in parenthesis.
Abbreviations: N=total screened patients randomized; n=total randomized patients with non-missing baseline; Min=minimum; Max=maximum; SD=standard deviation; mTSS=modified total sharp score
[Source: Statistical Reviewer]

Table 47 Change from Baseline in Components of ACR, excluding HAQ-DI, Study M13-542

Treatment Arm	N	Baseline Mean (SD)	n	Week 12 Mean (SD)	Est Diff (95 % CI)[1]	P-value[1]
CRP						
Placebo	169	16.3 (21.1)	151	13.8 (17.3)		
UPA 15 mg QD	164	16.2 (18.6)	158	5.1 (14.1)	-8.8 (-11.7, -5.9)	<0.001
UPA 30 mg QD	165	16.0 (21.2)	152	4.8 (8.0)	-9.2 (-12.1, -6.3)	<0.001
SJC (Out of 66)						
Placebo	169	16.3 (9.6)	151	9.2 (10.2)		
UPA 15 mg QD	164	17.0 (10.8)	160	6.0 (8.2)	-3.8 (-5.6, -1.9)	<0.001
UPA 30 mg QD	165	17.2 (11.4)	154	7.7 (11.4)	-2.3 (-4.1, -0.4)	0.016
TJC (Out of 68)						
Placebo	169	28.5 (15.3)	151	18.5 (17.2)		
UPA 15 mg QD	164	27.8 (16.3)	160	11.4 (14.3)	-6.8 (-9.6, -4.0)	<0.001
UPA 30 mg QD	165	27.3 (15.2)	154	11.4 (13.9)	-6.4 (-9.3, -3.6)	<0.001
Physician Global						
Placebo	166	66.9 (16.9)	144	39.2 (25.1)		
UPA 15 mg QD	163	68.7 (16.6)	159	28.9 (22.4)	-10.9 (-16.1, -5.7)	<0.001
UPA 30 mg QD	161	66.4 (15.6)	150	25.9 (21.7)	-13.3 (-18.5, -8.0)	<0.001
Patient Global						
Placebo	161	66.3 (22.7)	150	53.6 (27.9)		
UPA 15 mg QD	157	67.2 (19.6)	160	39.9 (26.5)	-13.9 (-19.7, -8.1)	<0.001
UPA 30 mg QD	157	64.7 (21.0)	154	35.2 (27.1)	-17.8 (-23.6, -11.9)	<0.001
Patient Pain						
Placebo	166	68.9 (21.0)	150	55.2 (27.7)		
UPA 15 mg QD	163	68.2 (19.8)	160	40.9 (27.6)	-13.8 (-19.6, -8.1)	<0.001
UPA 30 mg QD	161	65.3 (20.7)	154	35.3 (26.2)	-18.9 (-24.7, -13.0)	<0.001

1: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.

Abbreviations: N=total randomized; n=total patients with observed data; SD=standard deviation; CI=confidence intervals; CRP=C-reactive protein; UPA=upadacitinib; QD=once daily; SJC=swollen joint counts; TJC=tender joint counts

[Source: Statistical Reviewer]

Table 48 Change from Baseline in Components of ACR, excluding HAQ-DI, Study M13-549

Treatment Arm	N	Baseline Mean (SD)	n	Week 12 Mean (SD)	Est Diff (95 % CI)[1]	P-value[1]
CRP						
Placebo	221	12.6 (14.0)	207	13.2 (15.6)		
UPA 15 mg QD	221	16.6 (19.2)	209	4.5 (9.6)	-9.7 (-11.9, -7.6)	<0.001
UPA 30 mg QD	219	14.8 (16.9)	201	5.3 (8.6)	-8.4 (-10.5, -6.2)	<0.001
SJC (Out of 66)						
Placebo	221	15.4 (9.2)	207	8.9 (9.8)		
UPA 15 mg QD	221	16.0 (10.0)	210	6.4 (9.6)	-2.8 (-4.2, -1.4)	<0.001
UPA 30 mg QD	219	16.2 (10.6)	201	4.6 (6.4)	-4.5 (-5.8, -3.1)	<0.001
TJC (Out of 68)						
Placebo	221	24.7 (15.0)	207	16.3 (17.0)		
UPA 15 mg QD	221	25.2 (13.8)	210	11.9 (14.4)	-4.9 (-7.2, -2.7)	<0.001
UPA 30 mg QD	219	26.2 (14.3)	201	9.5 (11.4)	-7.6 (-9.9, -5.3)	<0.001
Physician Global						
Placebo	221	64.4 (17.7)	201	40.9 (24.5)		
UPA 15 mg QD	216	64.3 (16.2)	205	26.1 (20.9)	-14.7 (-18.9, -10.6)	<0.001
UPA 30 mg QD	219	63.0 (18.0)	193	22.8 (18.9)	-18.3 (-22.5, -14.0)	<0.001
Patient Global						
Placebo	211	60.3 (20.5)	206	49.8 (25.7)		
UPA 15 mg QD	209	63.1 (21.9)	210	32.3 (23.9)	-18.4 (-22.9, -13.9)	<0.001
UPA 30 mg QD	213	62.8 (20.3)	200	30.2 (22.6)	-20.2 (-24.7, -15.7)	<0.001
Patient Pain						
Placebo	221	61.5 (20.8)	206	50.8 (25.6)		
UPA 15 mg QD	217	64.1 (19.5)	210	33.1 (24.3)	-18.8 (-23.3, -14.4)	<0.001
UPA 30 mg QD	219	64.0 (19.8)	200	30.1 (22.8)	-21.5 (-26.0, -17.0)	<0.001

1: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model. Slight differences in Week 12 because data from TP1 was used and some observed data collected after Week 12 were in TP2 and not reviewed.

Abbreviations: N=total randomized; n=total patients with observed data; SD=standard deviation; CI=confidence intervals; CRP=C-reactive protein; UPA=upadacitinib; QD=once daily; SJC=swollen joint counts; TJC=tender joint counts

[Source: Statistical Reviewer]

Table 49 Change from Baseline in Components of ACR, excluding HAQ-DI, Study M14-465

Treatment Arm	N	Baseline Mean (SD)	n	Week 12 Mean (SD)	Est Diff (95 % CI)[1]	P-value[1]
CRP						
Placebo	651	18.0 (21.5)	595	16.2 (19.8)		
UPA 15 mg QD	651	17.9 (22.5)	590	5.5 (10.9)	-10.7 (-12.4, -9.0)	<0.001
ADA 40 mg EOW	327	19.8 (21.5)	299	9.3 (15.8)	-7.3 (-9.4, -5.2)	<0.001
SJC (Out of 66)						
Placebo	651	16.2 (9.0)	625	9.3 (8.9)		
UPA 15 mg QD	651	16.6 (10.3)	624	5.2 (7.4)	-4.3 (-5.0, -3.5)	<0.001
ADA 40 mg EOW	327	16.3 (9.2)	314	5.9 (6.5)	-3.4 (-4.3, -2.5)	<0.001
TJC (Out of 68)						
Placebo	651	26.0 (14.3)	625	16.4 (14.6)		
UPA 15 mg QD	651	26.4 (15.1)	624	9.9 (12.6)	-6.8 (-8.0, -5.6)	<0.001
ADA 40 mg EOW	327	26.4 (15.2)	314	11.7 (12.7)	-4.8 (-6.3, -3.4)	<0.001
Physician Global						
Placebo	649	1.6 (0.6)	623	1.3 (0.7)		
UPA 15 mg QD	646	1.6 (0.6)	623	1.0 (0.7)	-0.3 (-0.4, -0.3)	<0.001
ADA 40 mg EOW	325	1.6 (0.6)	314	1.1 (0.7)	-0.2 (-0.3, -0.1)	<0.001
Patient Global						
Placebo	620	66.0 (18.2)	613	40.8 (24.7)		
UPA 15 mg QD	616	65.6 (17.1)	614	26.8 (21.4)	-14.0 (-16.5, -11.5)	<0.001
ADA 40 mg EOW	305	65.1 (17.6)	304	29.4 (20.5)	-11.2 (-14.3, -8.2)	<0.001
Patient Pain						
Placebo	649	63.8 (21.5)	623	48.3 (24.5)		
UPA 15 mg QD	647	64.3 (21.8)	623	33.4 (24.2)	-15.0 (-17.6, -12.4)	<0.001
ADA 40 mg EOW	324	65.8 (21.1)	314	41.0 (24.7)	-7.9 (-11.1, -4.7)	<0.001

1: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.

Abbreviations: N=total randomized; n=total patients with observed data; SD=standard deviation; CI=confidence intervals; CRP=C-reactive protein; UPA=upadacitinib; QD=once daily; ADA=adalimumab; SJC=swollen joint counts; TJC=tender joint counts

[Source: Statistical Reviewer]

Table 50 Change from Baseline in Components of ACR, excluding HAQ-DI, Study M15-555

Treatment Arm	N	Baseline Mean (SD)	n	Week 14 Mean (SD)	Est Diff (95 % CI)[1]	P-value[1]
CRP						
MTX	216	14.5 (17.3)	196	12.8 (21.7)		
UPA 15 mg QD	217	14.0 (16.5)	197	3.7 (7.8)	-9.1 (-11.9, -6.3)	<0.001
UPA 30 mg QD	215	16.3 (20.8)	199	5.0 (12.0)	-8.3 (-11.1, -5.4)	<0.001
SJC (Out of 66)						
MTX	216	16.9 (11.5)	197	9.4 (11.1)		
UPA 15 mg QD	217	16.4 (10.9)	199	5.7 (8.6)	-3.2 (-4.6, -1.9)	<0.001
UPA 30 mg QD	215	16.9 (10.2)	203	4.6 (7.0)	-4.8 (-6.1, -3.4)	<0.001
TJC (Out of 68)						
MTX	216	25.2 (16.0)	197	14.7 (15.7)		
UPA 15 mg QD	217	24.5 (15.1)	199	9.8 (12.5)	-4.4 (-6.6, -2.3)	<0.001
UPA 30 mg QD	215	24.8 (15.2)	203	7.5 (10.7)	-7.0 (-9.1, -4.8)	<0.001
Physician Global						
MTX	216	62.1 (17.5)	193	36.9 (24.1)		
UPA 15 mg QD	216	65.7 (18.5)	193	25.3 (20.4)	-12.8 (-17.1, -8.6)	<0.001
UPA 30 mg QD	215	62.6 (17.8)	192	21.9 (19.5)	-14.8 (-19.1, -10.5)	<0.001
Patient Global						
MTX	200	59.6 (21.8)	195	48.2 (25.7)		
UPA 15 mg QD	209	62.2 (22.3)	199	36.8 (27.3)	-12.3 (-17.2, -7.5)	<0.001
UPA 30 mg QD	202	59.4 (22.8)	201	29.7 (23.8)	-18.5 (-23.3, -13.7)	<0.001
Patient Pain						
MTX	216	62.5 (21.3)	195	48.1 (25.1)		
UPA 15 mg QD	216	62.3 (22.5)	199	35.4 (26.4)	-12.8 (-17.5, -8.1)	<0.001
UPA 30 mg QD	215	61.9 (22.1)	201	28.8 (23.7)	-19.3 (-24.0, -14.6)	<0.001

1: Estimated difference in adjusted mean change from baseline compared to MTX, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.

Abbreviations: N=total randomized; n=total patients with observed data; SD=standard deviation; CI=confidence intervals; CRP=C-reactive protein; MTX=methotrexate; UPA=upadacitinib; QD=once daily; SJC=swollen joint counts; TJC=tender joint counts

[Source: Statistical Reviewer]

Table 51 Change from Baseline in Components of ACR, excluding HAQ-DI, Study M13-545

Treatment Arm	N	Baseline Mean (SD)	n	Week 12 Mean (SD)	Est Diff (95 % CI)[1]	P-value[1]
CRP						
MTX	314	21.2 (22.1)	292	10.9 (14.9)		
UPA 15 mg QD	317	23.0 (27.4)	302	4.2 (8.8)	-6.2 (-8.3, -4.2)	<0.001
UPA 30 mg QD	314	19.4 (22.6)	298	5.0 (15.6)	-5.0 (-7.0, -2.9)	<0.001
SJC (Out of 66)						
MTX	314	16.9 (10.6)	291	6.4 (8.3)		
UPA 15 mg QD	317	16.9 (10.3)	303	4.5 (6.6)	-1.7 (-2.6, -0.8)	<0.001
UPA 30 mg QD	314	15.7 (9.7)	301	3.9 (5.5)	-2.0 (-2.9, -1.1)	<0.001
TJC (Out of 68)						
MTX	314	26.4 (16.1)	291	12.8 (14.6)		
UPA 15 mg QD	317	25.4 (14.4)	303	8.6 (11.6)	-3.8 (-5.3, -2.2)	<0.001
UPA 30 mg QD	314	25.2 (15.0)	301	7.6 (10.0)	-4.5 (-6.0, -2.9)	<0.001
Physician Global						
MTX	314	1.6 (0.7)	289	1.1 (0.7)		
UPA 15 mg QD	317	1.6 (0.7)	303	0.8 (0.7)	-0.3 (-0.4, -0.2)	<0.001
UPA 30 mg QD	311	1.5 (0.7)	301	0.7 (0.7)	-0.3 (-0.4, -0.2)	<0.001
Patient Global						
MTX	299	68.7 (16.4)	288	32.2 (22.4)		
UPA 15 mg QD	301	67.1 (17.0)	301	21.7 (18.6)	-9.5 (-12.6, -6.4)	<0.001
UPA 30 mg QD	304	65.3 (16.6)	294	20.2 (17.5)	-10.8 (-13.8, -7.7)	<0.001
Patient Pain						
MTX	314	65.8 (21.5)	289	41.6 (25.0)		
UPA 15 mg QD	317	66.6 (22.0)	303	31.2 (24.3)	-9.8 (-13.6, -6.1)	<0.001
UPA 30 mg QD	311	64.9 (21.6)	301	28.0 (25.1)	-12.6 (-16.4, -8.8)	<0.001

1: Estimated difference in adjusted mean change from baseline compared to MTX, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4. Wald-based p-values were reported from the same regression model.

Abbreviations: N=total randomized; n=total patients with observed data; SD=standard deviation; CI=confidence intervals; CRP=C-reactive protein; MTX=methotrexate; UPA=upadacitinib; QD=once daily; SJC=swollen joint counts; TJC=tender joint counts

[Source: Statistical Reviewer]

Table 52 Proportion of Patients Rescued at Visit Weeks, Study M13-545

	MTX Monotherapy (N=314)	UPA 15 mg Monotherapy (N=317)	UPA 30 mg Monotherapy (N=317)	All (N=945)
Rescued at Week 16	18 (6%)	10 (3%)	2 (1%)	30 (3%)
Rescued at Week 20	3 (1%)	6 (2%)	2 (1%)	11 (1%)
Rescued at Week 24	6 (2%)	1 (0%)	2 (1%)	9 (1%)
Total Rescued at or prior to Week 24	27 (9%)	17 (5%)	8 (3%)	50 (5%)

Counts and percentages relative to N in parenthesis.
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate
[Source: Statistical Reviewer]

Table 53 Summary of ACR at Visit Weeks through Week 14, M14-465

Visit Week	PBO (N=651) Counts (%) [1]	UPA 15 mg QD (N=651) Counts (%)[1]	ADA 40 mg EOW (N=327) Counts (%)[1]	UPA 15 mg QD vs PBO Diff (95% CI)[2]	ADA 40 mg EOW vs PBO Diff (95% CI)[2]	UPA vs ADA Diff (95% CI)[2]
			ACR20			
2	91 (14%)	219 (34%)	109 (33%)	19.7% (15.2% - 24.2%)	19.4% (13.6% - 25.1%)	0.3% (-6.0% - 6.6%)
4	170 (26%)	349 (54%)	152 (47%)	27.5% (22.4% - 32.6%)	20.4% (14.0% - 26.7%)	7.1% (0.5% - 13.8%)
8	232 (36%)	425 (65%)	191 (58%)	29.6% (24.5% - 34.8%)	22.8% (16.3% - 29.3%)	6.9% (0.4% - 13.3%)
12	237 (36%)	459 (71%)	206 (63%)	34.1% (29.0% - 39.2%)	26.6% (20.2% - 33.0%)	7.5% (1.2% - 13.8%)
14	273 (42%)	476 (73%)	209 (64%)	31.2% (26.1% - 36.3%)	22.0% (15.5% - 28.4%)	9.2% (3.0% - 15.4%)
			ACR50			
2	16 (2.5)	75 (11.5)	28 (8.6)	9.1% (6.3% - 11.8%)	6.1% (2.8% - 9.4%)	3.0% (-0.9% - 6.9%)
4	36 (5.5)	138 (21.2)	48 (14.7)	15.7% (12.1% - 19.3%)	9.1% (4.9% - 13.4%)	6.5% (1.6% - 11.5%)
8	78 (12.0)	227 (34.9)	97 (29.7)	22.9% (18.5% - 27.3%)	17.7% (12.1% - 23.2%)	5.2% (-1.0% - 11.4%)
12	97 (14.9)	294 (45.2)	95 (29.1)	30.3% (25.6% - 35.0%)	14.2% (8.5% - 19.8%)	16.1% (9.9% - 22.3%)
14	117 (18.0)	307 (47.2)	115 (35.2)	29.2% (24.3% - 34.0%)	17.2% (11.2% - 23.2%)	12.0% (5.5% - 18.4%)
			ACR70			
2	4 (0.6)	18 (2.8)	6 (1.8)	2.2% (0.8% - 3.5%)	1.2% (-0.4% - 2.8%)	0.9% (-1.0% - 2.9%)
4	10 (1.5)	61 (9.4)	16 (4.9)	7.8% (5.4% - 10.3%)	3.4% (0.8% - 5.9%)	4.5% (1.2% - 7.7%)
8	24 (3.7)	114 (17.5)	36 (11.0)	13.8% (10.6% - 17.1%)	7.3% (3.6% - 11.0%)	6.5% (2.0% - 11.0%)
12	32 (4.9)	162 (24.9)	44 (13.5)	20.0% (16.3% - 23.7%)	8.5% (4.5% - 12.6%)	11.4% (6.5% - 16.4%)
14	51 (7.8)	179 (27.5)	54 (16.5)	19.7% (15.7% - 23.7%)	8.7% (4.2% - 13.2%)	11.0% (5.7% - 16.3%)

Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.

1: Counts and percentages relative to N in parenthesis are reported for the probability of ACR20 response.

2: Difference in the probability of ACR response (%), respective 95% CI using normal approximation to difference in binomial proportions.

Note that comparative assessment between any of the arms at various timepoints are not controlled for multiplicity.

Abbreviations: ACR=American College of Rheumatology; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab, EOW=every other week

[Source: Statistical Reviewer]

Table 54 Summary of ACR at Visit Weeks through Week 24, M13-545

Visit Week	MTX monotherapy (N=314) Counts (%)[1]	UPA 15 mg QD monotherapy (N=317) Counts (%)[1]	UPA 30 mg QD monotherapy (N=314) Counts (%)[1]	UPA 15 mg QD Vs MTX Diff (95% CI)[2]	UPA 30 mg QD Vs MTX Diff (95% CI)[2]
			ACR20		
2	51 (16.2)	127 (40.1)	149 (47.5)	23.8% (17.1% - 30.6%)	31.2% (24.3% - 38.1%)
4	106 (33.8)	193 (60.9)	203 (64.6)	27.1% (19.6% - 34.6%)	30.9% (23.5% - 38.3%)
8	157 (50.0)	234 (73.8)	226 (72.0)	23.8% (16.5% - 31.2%)	22.0% (14.5% - 29.4%)
12	170 (54.1)	240 (75.7)	242 (77.1)	21.6% (14.3% - 28.8%)	22.9% (15.7% - 30.1%)
16	198 (63.1)	254 (80.1)	248 (79.0)	17.1% (10.2% - 24.0%)	15.9% (8.9% - 22.9%)
20	194 (61.8)	251 (79.2)	250 (79.6)	17.4% (10.4% - 24.4%)	17.8% (10.9% - 24.8%)
24	184 (58.6)	250 (78.9)	245 (78.0)	20.3% (13.2% - 27.3%)	19.4% (12.3% - 26.5%)
			ACR50		
2	9 (2.9)	48 (15.1)	59 (18.8)	12.3% (7.9% - 16.6%)	15.9% (11.2% - 20.6%)
4	23 (7.3)	102 (32.2)	101 (32.2)	24.9% (19.0% - 30.7%)	24.8% (18.9% - 30.8%)
8	59 (18.8)	140 (44.2)	148 (47.1)	25.4% (18.4% - 32.3%)	28.3% (21.3% - 35.4%)
12	89 (28.3)	165 (52.1)	177 (56.4)	23.7% (16.3% - 31.1%)	28.0% (20.6% - 35.4%)
16	103 (32.8)	173 (54.6)	188 (59.9)	21.8% (14.2% - 29.3%)	27.1% (19.6% - 34.6%)
20	119 (37.9)	196 (61.8)	197 (62.7)	23.9% (16.4% - 31.5%)	24.8% (17.3% - 32.4%)
24	105 (33.4)	191 (60.3)	206 (65.6)	26.8% (19.3% - 34.3%)	32.2% (24.8% - 39.6%)
			ACR70		
2	4 (1.3)	16 (5.0)	24 (7.6)	3.8% (1.1% - 6.5%)	6.4% (3.2% - 9.6%)
4	9 (2.9)	36 (11.4)	58 (18.5)	8.5% (4.5% - 12.4%)	15.6% (10.9% - 20.3%)
8	25 (8.0)	74 (23.3)	95 (30.3)	15.4% (9.8% - 20.9%)	22.3% (16.4% - 28.2%)
12	44 (14.0)	103 (32.5)	116 (36.9)	18.5% (12.1% - 24.9%)	22.9% (16.4% - 29.5%)
16	44 (14.0)	115 (36.3)	125 (39.8)	22.3% (15.7% - 28.8%)	25.8% (19.2% - 32.4%)
20	61 (19.4)	128 (40.4)	147 (46.8)	21.0% (14.0% - 27.9%)	27.4% (20.3% - 34.4%)
24	58 (18.5)	141 (44.5)	156 (49.7)	26.0% (19.1% - 33.0%)	31.2% (24.2% - 38.2%)

Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.

1: Counts and percentages relative to N in parenthesis are reported for the probability of ACR20 response.

2: Difference in the probability of ACR response (%), respective 95% CI using normal approximation to difference in binomial proportions.

Note that comparative assessment between any of the arms at various timepoints are not controlled for multiplicity.

Abbreviations: ACR=American College of Rheumatology; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab, EOW=every other week

[Source: Statistical Reviewer]

Table 55 Change from Baseline in CDAI at Primary Timepoint for each Study

Treatment Arm	N	Baseline Mean (SD)[1]	n	Visit Mean (SD)[1]	Adj Mean Diff (95% CI) [2]
cDMARD Add-on Studies					
M13-542					
Placebo	160	41.0 (13.3)	139	25.8 (15.9)	
UPA 15 MG QD	156	41.7 (13.3)	151	17.3 (13.6)	-9.1 (-12.2, -6.1)
UPA 30 MG QD	156	40.1 (12.2)	141	16.9 (14.6)	-8.7 (-11.8, -5.6)
M13-549					
Placebo	211	37.8 (11.8)	192	24.0 (15.5)	
UPA 15 MG QD	208	38.3 (11.9)	193	16.5 (14.7)	-8.0 (-10.6, -5.5)
UPA 30 MG QD	213	38.6 (12.7)	186	13.5 (11.8)	-11.1 (-13.6, -8.6)
M14-465					
Placebo	620	40.0 (12.7)	590	25.4 (15.6)	
UPA 15 mg QD	616	39.7 (12.9)	581	14.8 (12.7)	-10.4 (-11.9, -9.0)
ADA 40 mg EOW	304	39.8 (13.2)	285	17.8 (12.7)	-7.5 (-9.3, -5.7)
Monotherapy Studies					
M15-555 [3]					
MTX	200	37.8 (14.4)	178	23.0 (16.9)	
UPA 15 MG QD	209	38.0 (13.1)	187	15.4 (12.8)	-7.4 (-10.0, -4.9)
UPA 30 MG QD	202	38.4 (13.8)	180	12.8 (11.7)	-10.4 (-13.0, -7.8)
M13-545					
MTX	300	40.4 (13.4)	276	19.3 (14.8)	
UPA 15 MG QD	301	40.4 (13.3)	287	13.8 (12.3)	-5.5 (-7.4, -3.6)
UPA 30 MG QD	303	39.3 (13.5)	287	12.3 (11.2)	-6.6 (-8.5, -4.6)

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to reference arm (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in CDAI adjusting for treatment groups, baseline CDAI, and key stratification factor listed in Table 4.
3: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
Abbreviations: CDAI=clinical disease activity index; UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; cDMARD=conventional disease modifying anti-rheumatic drugs
[Source: Statistical Reviewer]

Table 56 Proportion of Patients with CDAI < 2.8 or < 10 and Remain on Randomized Treatment at the Primary Timepoint for all Studies

Treatment Arm	Total	CDAI < 2.8		CDAI < 10	
		Counts (%)	Diff (%) [95 % CI][2]	Counts (%)	Diff (%) [95 % CI][2]
cDMARD Add-on Studies					
M13-542					
Placebo	169	8 (5%)		24 (14%)	
UPA 15 MG QD	164	13 (8%)	3.2% (-2.0%, 8.4%)	52 (32%)	17.5% (8.7%, 26.4%)
UPA 30 MG QD	165	19 (12%)	6.8% (1.0%, 12.6%)	56 (34%)	19.7% (10.8%, 28.7%)
M13-549					
Placebo	221	7 (3%)		42 (19%)	
UPA 15 MG QD	221	20 (9%)	5.9% (1.5%, 10.3%)	89 (40%)	21.3% (13.0%, 29.5%)
UPA 30 MG QD	219	26 (12%)	8.7% (3.8%, 13.6%)	92 (42%)	23.0% (14.7%, 31.3%)
M14-465					
Placebo	651	20 (3%)		93 (30%)	
UPA 15 mg QD	651	87 (13%)	4.6% (1.4%, 7.7%)	154 (49%)	19.4% (11.9%, 26.9%)
ADA 40 mg EOW	327	25 (8%)	10.3% (7.4%, 13.2%)	147 (46%)	16.8% (9.3%, 24.2%)
Monotherapy Studies					
M15-555 [1]					
MTX	216	2 (1%)		53 (25%)	
UPA 15 MG QD	217	28 (13%)	12.0% (7.3%, 16.6%)	75 (35%)	10.0% (1.5%, 18.6%)
UPA 30 MG QD	215	41 (19%)	18.1% (12.7%, 23.5%)	100 (47%)	22.0% (13.2%, 30.8%)
M13-545					
MTX	314	20 (6%)		93 (30%)	
UPA 15 MG QD	317	51 (16%)	9.7% (4.9%, 14.6%)	147 (46%)	16.8% (9.3%, 24.2%)
UPA 30 MG QD	314	67 (21%)	15.0% (9.7%, 20.2%)	154 (49%)	19.4% (11.9%, 26.9%)

Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.
Counts (%): Counts and percentages relative to N in parenthesis are reported for the probability of response.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Difference in the probability of response compared to reference, respective 95% CI using normal approximation to difference in binomial proportions.
Abbreviations: CDAI=clinical disease activity index; UPA=upadacitinib; cDMARD=conventional disease modifying anti-rheumatic drugs; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval
[Source: Statistical Reviewer]

Table 57 Change from Baseline in SDAI at the Primary Timepoint for all Studies

Treatment Arm	N	Baseline Mean (SD) [1]	n	Visit Mean (SD) [1]	Adjusted Mean Diff (95 % CI) [2]
cDMARD Add-on Studies					
M13-542					
Placebo	160	42.6 (13.9)	138	27.1 (16.5)	
UPA 15 MG QD	156	43.3 (13.7)	149	17.8 (14.0)	-10.1 (-13.2, -6.9)
UPA 30 MG QD	156	41.7 (12.8)	138	17.3 (14.7)	-9.6 (-12.8, -6.4)
M13-549					
Placebo	211	39.0 (11.9)	192	25.4 (15.9)	
UPA 15 MG QD	208	39.9 (12.5)	192	16.9 (14.8)	-9.2 (-11.7, -6.6)
UPA 30 MG QD	213	40.0 (13.1)	186	14.0 (11.8)	-11.9 (-14.5, -9.3)
M14-465					
Placebo	620	41.8 (13.3)	563	27.0 (16.2)	
ADA 40 mg EOW	304	41.8 (13.7)	269	18.8 (13.4)	-8.0 (-9.9, -6.1)
UPA 15 mg QD	616	41.5 (13.5)	551	15.3 (12.8)	-11.5 (-13.0, -10.0)
Monotherapy Studies					
M15-555 [3]					
MTX	200	39.2 (14.6)	177	24.4 (17.7)	
UPA 15 MG QD	209	39.4 (13.4)	185	15.8 (13.0)	-8.3 (-11.0, -5.7)
UPA 30 MG QD	202	40.0 (14.3)	177	13.3 (11.8)	-11.4 (-14.1, -8.8)
M13-545					
MTX	300	42.5 (14.0)	275	20.5 (15.3)	
UPA 15 MG QD	301	42.7 (13.9)	286	14.3 (12.4)	-6.4 (-8.3, -4.4)
UPA 30 MG QD	303	41.3 (14.4)	284	12.8 (11.5)	-7.2 (-9.2, -5.2)

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to reference (placebo or MTX), respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
3: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
Abbreviations: SDAI=simple disease activity index; UPA=upadacitinib; cDMARD=conventional disease modifying anti-rheumatic drugs; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval
[Source: Statistical Reviewer]

Table 58 Proportion of Patients with SDAI < 3.3 or < 11 and Remained on Randomized Treatment at the Primary Timepoint for all Studies

Treatment Arm	N	SDAI < 3.3		SDAI < 11	
		Count (%)	Diff (%) [95% CI][2]	Counts (%)	Diff (%) [95% CI][2]
cDMARD Add-on Studies					
M13-542					
Placebo	169	8 (5%)		24 (14%)	
UPA 15 MG QD	164	14 (9%)	3.8% (-1.5%, 9.1%)	56 (34%)	19.9% (11.0%, 28.9%)
UPA 30 MG QD	165	18 (11%)	6.2% (0.4%, 11.9%)	58 (35%)	21.0% (12.0%, 29.9%)
M13-549					
Placebo	221	7 (3%)		42 (19%)	
UPA 15 MG QD	221	21 (10%)	6.3% (1.8%, 10.8%)	92 (42%)	22.6% (14.3%, 30.9%)
UPA 30 MG QD	219	27 (12%)	9.2% (4.2%, 14.1%)	99 (45%)	26.2% (17.8%, 34.6%)
M14-465					
Placebo	651	18 (3%)		99 (15%)	
UPA 15 mg QD	651	79 (12%)	4.6% (1.5%, 7.7%)	263 (40%)	14.8% (9.1%, 20.4%)
ADA 40 mg EOW	327	24 (7%)	9.4% (6.6%, 12.2%)	98 (30%)	25.2% (20.5%, 29.9%)
Monotherapy Studies					
M15-555 [1]					
MTX	216	2 (1%)		52 (24%)	
UPA 15 MG QD	217	30 (14%)	12.9% (8.1%, 17.7%)	81 (37%)	13.3% (4.7%, 21.9%)
UPA 30 MG QD	215	39 (18%)	17.2% (11.9%, 22.5%)	101 (47%)	22.9% (14.1%, 31.7%)
M13-545					
MTX	314	19 (6%)		93 (30%)	
UPA 15 MG QD	317	50 (16%)	9.7% (4.9%, 14.5%)	155 (49%)	19.3% (11.8%, 26.7%)
UPA 30 MG QD	314	69 (22%)	15.9% (10.6%, 21.2%)	163 (52%)	22.3% (14.8%, 29.8%)

Patients who had discontinued the study treatment prior to primary efficacy timepoint, were lost to follow-up, or had withdrawn from the study were imputed as non-response.
The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
Counts (%): Counts and percentages relative to N in parenthesis are reported for the probability of response.
1: The primary time-point in Study M15-555 is Week 14 while other studies are evaluated at Week 12.
2: Difference in the probability of response compared to reference, respective 95% CI using normal approximation to difference in binomial proportions.
Abbreviations: SDAI=simple disease activity index; UPA=upadacitinib; cDMARD=conventional disease modifying anti-rheumatic drugs; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval
[Source: Statistical Reviewer]

Table 59 Individual Components of SF-36, Study M13-542

Components	Treatment Arm	N	Baseline Mean (SD) [1]	n	Week 12 Mean (SD) [1]	Diff (95% CI) [2]	P-value
Physical Component Score							
Bodily Pain	Placebo	166	33.2 (6.4)	147	38.6 (8.7)		
	UPA 15 mg QD	163	33.1 (6.4)	157	41.7 (9.6)	3.1 (1.3, 4.9)	<0.001
	UPA 30 mg QD	162	34.5 (6.9)	149	43.2 (8.6)	4.2 (2.4, 6.0)	<0.001
General Health	Placebo	166	38.9 (9.7)	147	41.2 (9.4)		
	UPA 15 mg QD	163	37.1 (8.2)	157	42.3 (9.7)	1.8 (0.2, 3.4)	0.029
	UPA 30 mg QD	162	38.1 (9.1)	149	42.6 (9.5)	1.7 (0.1, 3.3)	0.039
Physical Functioning	Placebo	166	32.0 (8.9)	147	34.3 (10.0)		
	UPA 15 mg QD	163	30.6 (9.2)	157	36.0 (10.6)	2.5 (0.7, 4.2)	0.006
	UPA 30 mg QD	162	31.2 (8.1)	149	37.8 (9.8)	4.1 (2.4, 5.9)	<0.001
Role Physical	Placebo	166	34.2 (8.0)	147	36.8 (8.6)		
	UPA 15 mg QD	163	33.0 (8.7)	157	38.7 (10.5)	2.7 (1.0, 4.3)	0.001
	UPA 30 mg QD	162	34.9 (7.6)	149	39.9 (8.8)	2.9 (1.3, 4.6)	<0.001
Mental Component Score							
Mental Health	Placebo	166	43.6 (12.5)	147	47.0 (10.7)		
	UPA 15 mg QD	163	42.9 (11.3)	157	47.1 (12.1)	0.5 (-1.3, 2.4)	0.57
	UPA 30 mg QD	162	44.2 (11.4)	149	47.7 (10.9)	0.1 (-1.8, 1.9)	0.94
Role Emotional	Placebo	166	41.0 (12.6)	147	44.0 (10.9)		
	UPA 15 mg QD	163	39.2 (12.7)	157	44.6 (12.4)	1.6 (-0.4, 3.5)	0.11
	UPA 30 mg QD	162	41.5 (12.7)	149	44.9 (11.1)	0.6 (-1.4, 2.5)	0.56
Social Functioning	Placebo	166	39.5 (10.8)	147	43.2 (9.7)		
	UPA 15 mg QD	163	37.3 (10.4)	157	44.0 (10.6)	1.7 (-0.1, 3.6)	0.07
	UPA 30 mg QD	162	39.1 (10.8)	149	44.7 (9.8)	1.4 (-0.4, 3.3)	0.14
Vitality	Placebo	166	40.6 (9.3)	147	44.0 (9.8)		
	UPA 15 mg QD	163	37.7 (10.0)	157	45.0 (11.2)	2.6 (0.7, 4.5)	0.008
	UPA 30 mg QD	162	39.2 (9.4)	149	46.5 (10.1)	3.3 (1.4, 5.3)	<0.001

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total patients with baseline; QD=once daily; n=number of patients with observed data
[Source: Statistical Reviewer]

Table 60 Individual Components of SF-36, Study M13-549

Components	Treatment Arm	N	Baseline Mean (SD) [1]	n	Week 12 Mean (SD) [1]	Diff (95% CI) [2]	P-value
Physical Component Score							
Bodily Pain	Placebo	221	35.3 (6.9)	207	40.2 (8.9)		
	UPA 15 mg QD	219	35.5 (6.4)	211	44.8 (8.1)	4.6 (3.1, 6.1)	<0.001
	UPA 30 mg QD	217	34.6 (6.8)	199	45.8 (8.1)	5.8 (4.2, 7.3)	<0.001
General Health	Placebo	221	39.2 (9.0)	207	41.5 (9.4)		
	UPA 15 mg QD	219	38.9 (8.1)	211	44.9 (9.3)	3.6 (2.1, 5.0)	<0.001
	UPA 30 mg QD	217	39.1 (9.4)	199	44.8 (9.3)	3.1 (1.6, 4.5)	<0.001
Physical Functioning	Placebo	221	33.3 (9.4)	207	36.9 (10.2)		
	UPA 15 mg QD	219	33.5 (8.8)	211	40.7 (9.9)	3.8 (2.4, 5.3)	<0.001
	UPA 30 mg QD	217	32.9 (9.7)	199	40.7 (10.7)	3.9 (2.4, 5.4)	<0.001
Role Physical	Placebo	221	35.4 (8.1)	207	38.6 (8.9)		
	UPA 15 mg QD	219	35.5 (8.5)	211	41.9 (9.1)	3.5 (2.0, 5.0)	<0.001
	UPA 30 mg QD	217	34.7 (8.9)	199	41.9 (9.5)	3.6 (2.1, 5.1)	<0.001
Mental Component Score							
Mental Health	Placebo	221	43.8 (10.7)	207	46.9 (10.7)		
	UPA 15 mg QD	219	44.1 (9.9)	211	48.3 (9.5)	1.3 (-0.2, 2.9)	0.09
	UPA 30 mg QD	217	44.0 (11.5)	199	48.7 (10.1)	1.4 (-0.1, 3.0)	0.07
Role Emotional	Placebo	221	42.2 (11.8)	207	44.3 (11.1)		
	UPA 15 mg QD	219	41.1 (11.7)	211	46.4 (9.8)	2.9 (1.2, 4.5)	<0.001
	UPA 30 mg QD	217	41.1 (11.9)	199	45.4 (10.8)	1.4 (-0.2, 3.1)	0.09
Social Functioning	Placebo	221	40.8 (10.8)	207	43.7 (10.7)		
	UPA 15 mg QD	219	40.6 (9.9)	211	47.1 (8.7)	3.6 (2.0, 5.2)	<0.001
	UPA 30 mg QD	217	40.8 (11.0)	199	46.3 (9.9)	2.4 (0.8, 4.0)	<0.001
Vitality	Placebo	221	41.8 (9.0)	207	44.5 (10.1)		
	UPA 15 mg QD	219	41.5 (9.0)	211	48.2 (10.0)	4.1 (2.5, 5.8)	<0.001
	UPA 30 mg QD	217	41.2 (10.0)	199	49.0 (10.8)	4.7 (3.0, 6.4)	<0.001

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total patients with baseline; QD=once daily; n=number of patients with observed data
[Source: Statistical Reviewer]

Table 61 Individual Components of SF-36, Study M14-465

Components	Treatment Arm	N	Baseline Mean (SD) [1]	n	Week 12 Mean (SD) [1]	Diff (95% CI) [2]	P-value
Physical Component Score							
Bodily Pain	Placebo	647	34.4 (6.7)	632	39.2 (8.1)		
	UPA 15 mg QD	648	34.2 (6.7)	634	44.3 (8.5)	5.2 (4.3, 6.0)	<0.001
	ADA 40 mg EOW	327	33.8 (6.5)	322	42.2 (8.5)	3.2 (2.2, 4.3)	<0.001
General Health	Placebo	647	37.9 (8.3)	632	41.0 (8.8)		
	UPA 15 mg QD	648	37.7 (8.0)	634	44.8 (8.9)	4.0 (3.2, 4.8)	<0.001
	ADA 40 mg EOW	327	37.9 (8.1)	322	43.5 (8.9)	2.4 (1.5, 3.4)	<0.001
Physical Function	Placebo	647	31.4 (8.9)	632	35.6 (9.4)		
	UPA 15 mg QD	648	31.8 (8.8)	634	39.3 (10.1)	3.7 (2.8, 4.6)	<0.001
	ADA 40 mg EOW	327	31.2 (8.7)	322	37.7 (10.1)	2.4 (1.4, 3.5)	<0.001
Role Physical	Placebo	647	33.4 (7.7)	632	37.3 (8.4)		
	UPA 15 mg QD	648	33.5 (7.9)	634	40.6 (8.7)	3.2 (2.4, 4.0)	<0.001
	ADA 40 mg EOW	327	33.0 (7.4)	322	38.5 (8.8)	1.5 (0.5, 2.5)	<0.001
Mental Component Score							
Mental Health	Placebo	647	40.8 (10.7)	632	44.5 (10.7)		
	UPA 15 mg QD	648	40.7 (10.2)	634	47.2 (9.9)	3.0 (2.0, 3.9)	<0.001
	ADA 40 mg EOW	327	40.8 (10.0)	322	46.3 (9.8)	1.9 (0.8, 3.0)	<0.001
Role Emotional	Placebo	647	37.8 (11.5)	632	41.3 (10.9)		
	UPA 15 mg QD	648	38.2 (11.4)	634	44.0 (10.2)	2.6 (1.6, 3.6)	<0.001
	ADA 40 mg EOW	327	37.2 (11.2)	322	42.5 (10.7)	1.5 (0.3, 2.7)	<0.001
Social Functioning	Placebo	647	38.1 (9.8)	632	41.4 (9.8)		
	UPA 15 mg QD	648	38.1 (9.8)	634	45.0 (9.0)	3.7 (2.8, 4.6)	<0.001
	ADA 40 mg EOW	327	37.8 (10.6)	322	43.4 (9.8)	2.2 (1.1, 3.3)	<0.001
Vitality	Placebo	647	41.1 (9.1)	632	45.4 (9.6)		
	UPA 15 mg QD	648	40.9 (8.9)	634	49.1 (9.3)	3.9 (3.0, 4.8)	<0.001
	ADA 40 mg EOW	327	40.8 (8.9)	322	47.7 (9.1)	2.5 (1.4, 3.5)	<0.001

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to placebo, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total patients with baseline; QD=once daily; n=number of patients with observed data
[Source: Statistical Reviewer]

Table 62 Individual Components of SF-36, Study M15-555

Components	Treatment Arm	N	Baseline Mean (SD) [1]	n	Week 14 Mean (SD) [1]	Diff (95% CI) [2]	P-value
Physical Component Score							
Bodily Pain	MTX	216	35.0 (6.7)	195	39.8 (7.6)		
	UPA 15 mg QD	217	35.3 (6.7)	200	44.1 (8.6)	4.3 (2.7, 5.8)	<0.001
	UPA 30 mg QD	214	35.4 (7.5)	202	46.3 (8.7)	6.4 (4.9, 8.0)	<0.001
General Health	MTX	216	38.7 (7.6)	195	41.5 (8.4)		
	UPA 15 mg QD	217	38.1 (8.4)	200	43.9 (9.6)	2.8 (1.4, 4.3)	<0.001
	UPA 30 mg QD	214	39.1 (8.7)	202	46.1 (9.2)	4.3 (2.9, 5.8)	<0.001
Physical Functioning	MTX	216	33.0 (9.0)	195	36.4 (9.5)		
	UPA 15 mg QD	217	32.7 (9.6)	200	40.8 (10.0)	4.6 (3.0, 6.1)	<0.001
	UPA 30 mg QD	214	32.9 (9.2)	202	42.9 (10.0)	6.3 (4.8, 7.9)	<0.001
Role Physical	MTX	216	35.6 (8.2)	195	38.5 (8.2)		
	UPA 15 mg QD	217	34.6 (8.2)	200	41.7 (8.6)	3.6 (2.2, 5.1)	<0.001
	UPA 30 mg QD	214	35.7 (8.6)	202	43.5 (8.7)	4.9 (3.5, 6.3)	<0.001
Mental Component Score							
Mental Health	MTX	216	42.5 (10.6)	195	45.2 (10.2)		
	UPA 15 mg QD	217	41.5 (10.9)	200	47.5 (10.0)	2.7 (1.1, 4.3)	<0.001
	UPA 30 mg QD	214	42.2 (10.9)	202	48.4 (10.0)	3.2 (1.6, 4.8)	<0.001
Role Emotional	MTX	216	41.2 (11.5)	195	43.4 (10.5)		
	UPA 15 mg QD	217	39.8 (11.3)	200	45.3 (10.0)	2.8 (1.2, 4.4)	<0.001
	UPA 30 mg QD	214	39.8 (11.9)	202	46.8 (10.6)	4.0 (2.4, 5.6)	<0.001
Social Functioning	MTX	216	40.0 (10.2)	195	43.7 (9.5)		
	UPA 15 mg QD	217	39.6 (10.3)	200	46.0 (9.2)	2.7 (1.0, 4.3)	0.001
	UPA 30 mg QD	214	40.1 (10.2)	202	47.2 (10.0)	3.5 (1.8, 5.1)	<0.001
Vitality	MTX	216	41.4 (9.0)	195	44.8 (9.6)		
	UPA 15 mg QD	217	41.4 (9.3)	200	49.7 (9.7)	4.9 (3.3, 6.6)	<0.001
	UPA 30 mg QD	214	42.0 (9.0)	202	50.3 (9.6)	5.2 (3.6, 6.9)	<0.001

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to MTX, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total patients with baseline; QD=once daily; n=number of patients with observed data; MTX=methotrexate
[Source: Statistical Reviewer]

Table 63 Individual Components of SF-36, Study M13-545

Components	Treatment Arm	N	Baseline Mean (SD) [1]	n	Week 12 Mean (SD) [1]	Diff (95% CI) [2]	P-value
Physical Component Score							
Bodily Pain	MTX	313	34.0 (7.2)	299	41.0 (8.8)		
	UPA 15 mg QD	315	33.0 (6.8)	306	45.3 (8.8)	4.8 (3.5, 6.1)	<0.001
	UPA 30 mg QD	312	34.3 (6.9)	306	46.0 (8.8)	4.9 (3.6, 6.2)	<0.001
General Health	MTX	313	40.2 (9.1)	299	44.6 (9.6)		
	UPA 15 mg QD	315	39.7 (9.6)	306	46.2 (9.8)	2.1 (0.9, 3.3)	<0.001
	UPA 30 mg QD	312	39.6 (8.7)	306	46.7 (8.8)	2.6 (1.3, 3.8)	<0.001
Physical Functioning	MTX	313	31.8 (9.3)	299	37.4 (10.0)		
	UPA 15 mg QD	315	31.7 (9.3)	306	41.9 (10.6)	4.6 (3.3, 6.0)	<0.001
	UPA 30 mg QD	312	33.0 (9.4)	306	42.7 (10.3)	4.7 (3.3, 6.1)	<0.001
Role Physical	MTX	313	33.9 (8.2)	299	38.3 (8.7)		
	UPA 15 mg QD	315	33.6 (8.6)	306	42.0 (9.7)	4.0 (2.7, 5.3)	<0.001
	UPA 30 mg QD	312	34.4 (8.5)	306	42.7 (9.0)	4.1 (2.8, 5.5)	<0.001
Mental Component Score							
Mental Health	MTX	313	41.1 (10.2)	299	45.3 (10.4)		
	UPA 15 mg QD	315	40.4 (10.5)	306	47.5 (10.2)	2.6 (1.2, 4.0)	<0.001
	UPA 30 mg QD	312	41.5 (11.0)	306	47.8 (10.7)	2.4 (0.9, 3.8)	<0.001
Role Emotional	MTX	313	38.5 (11.5)	299	42.3 (10.8)		
	UPA 15 mg QD	315	37.9 (11.7)	306	44.5 (10.5)	2.5 (1.0, 4.0)	<0.001
	UPA 30 mg QD	312	38.2 (12.4)	306	45.1 (10.6)	3.0 (1.5, 4.4)	<0.001
Social Functioning	MTX	313	38.0 (10.6)	299	42.6 (10.1)		
	UPA 15 mg QD	315	37.9 (10.3)	306	45.7 (10.1)	3.2 (1.7, 4.6)	0.001
	UPA 30 mg QD	312	38.8 (10.8)	306	46.6 (9.8)	3.8 (2.3, 5.2)	<0.001
Vitality	MTX	313	40.9 (9.7)	299	46.5 (10.0)		
	UPA 15 mg QD	315	40.0 (9.2)	306	49.9 (9.4)	4.0 (2.7, 5.4)	<0.001
	UPA 30 mg QD	312	41.6 (9.2)	306	50.0 (9.6)	3.4 (2.0, 4.7)	<0.001

1: Observed mean and standard deviation in parenthesis based on all patients remaining in the study regardless of adherence to randomized treatment at week of interest.
2: Estimated difference in adjusted mean change from baseline compared to MTX, respective 95% CI were reported based on a linear regression fit to the change from baseline in component adjusting for treatment groups, baseline component, and key stratification factor listed in Table 4.
Abbreviations: UPA=upadacitinib; SD=standard deviation; CI=confidence interval; N=total patients with baseline; QD=once daily; n=number of patients with observed data; MTX=methotrexate
[Source: Statistical Reviewer]

Table 64 Applicant's Primary Analysis of the Mean Change from Baseline in mTSS and Components at Week 26 using Linear Extrapolation for Patients Rescued, Study M14-465

Treatment Arm	N	Baseline Mean (SD)	n	Change from Baseline Mean (SD)	Est Diff (95% CI)[1,2]	P-value	Est Diff (95% CI)[1,3]	P-value
mTSS								
Placebo	649	35.9 (51.7)	599	0.9 (3.5)				
UPA 15 mg QD	644	34.0 (50.1)	593	0.2 (1.6)	-0.67 (-0.97, -0.37)	<0.001	0.14 (-0.23, 0.51)	0.45
ADA 40 mg EOW	326	34.5 (47.1)	296	0.1 (2.4)	-0.82 (-1.2, -0.45)	<0.001		
Erosion Score								
Placebo	649	17.0 (27.4)	599	0.5 (2.2)				
UPA 15 mg QD	644	16.5 (26.4)	593	0.1 (0.8)	-0.41 (-0.58, -0.24)	<0.001	0.009 (-0.2, 0.22)	0.93
ADA 40 mg EOW	326	15.4 (23.1)	296	0.0 (0.9)	-0.42 (-0.63, -0.21)	<0.001		
Joint Space Narrowing								
Placebo	649	18.9 (26.1)	599	0.5 (2.4)				
UPA 15 mg QD	644	17.5 (25.1)	593	0.1 (1.1)	-0.36 (-0.56, -0.17)	<0.001	0.08 (-0.16, 0.32)	0.50
ADA 40 mg EOW	326	19.2 (25.8)	296	0.0 (1.0)	-0.45 (-0.69, -0.21)	<0.001		

1: Estimated difference in the adjusted mean change from baseline compared to placebo, respective 95% CI, and p-values were based on a linear regression fit to the change from baseline in mTSS adjusting for treatment, baseline mTSS, key stratification factor.
2: Comparison is made with respect to placebo arm
3: Comparison is made with adalimumab 40 mg EOW
Abbreviations: UPA=upadacitinib; QD=once daily; MTX=methotrexate; ADA=adalimumab; EOW=every other week; SD=standard deviation; CI=confidence interval; mTSS=modified total Sharpe score; N=total number of patients at baseline with measurement; n=total number of patients at Week 26 with observed x-ray
[Source: Statistical Reviewer]

Table 65 EAIR for Treatment Emergent AEs of Special Interest up to 12 Weeks During Placebo-Controlled Period, Study M13-542 and M13-549

AE Category	Study ID	Placebo Events/Exposure (EAIR)	UPA 15 mg QD Events/Exposure (EAIR)	UPA 30 mg QD Events/Exposure (EAIR)
Infections	M13-542	52/29.9 (174)	54/30.3 (178)	56/30.6 (183)
	M13-549	47/43.1 (109)	64/41.5 (154)	70/39.3 (178)
Serious Infections	M13-542	0/36.0 (-)	1/37.2 (2.7)	4/36.9 (10.8)
	M13-549	1/49.4 (2)	1/49.4 (2)	3/47.9 (6.3)
Opportunistic Infection	M13-542	0/36.0 (-)	2/37.0 (5.4)	3/36.6 (8.2)
	M13-549	1/49.4 (2)	0/49.5 (-)	3/48.0 (6.3)
Confirmed Malignancy	M13-542	0/36.0 (-)	1/37.3 (2.7)	2/36.8 (5.4)
	M13-549	0/49.4 (-)	0/49.5 (-)	3/47.9 (6.3)
Malignancy (Excl NMSC)	M13-542	0/36.0 (-)	1/37.3 (2.7)	2/36.8 (5.4)
	M13-549	0/49.4 (-)	0/49.5 (-)	1/47.9 (2.1)
Lymphoma	M13-542	0/36.0 (-)	0/37.3 (-)	0/36.9 (-)
	M13-549	0/49.4 (-)	0/49.5 (-)	1/47.9 (2.1)
Any Hepatic Disorder	M13-542	2/35.9 (5.6)	2/37.1 (5.4)	3/36.5 (8.2)
	M13-549	7/48.7 (14.4)	3/49.3 (6.1)	6/47.3 (12.7)
Any GI perforation	M13-542	0/36.0 (-)	0/37.3 (-)	0/36.9 (-)
	M13-549	0/49.4 (-)	0/49.5 (-)	0/48.1 (-)
Anemia	M13-542	0/36.0 (-)	0/37.3 (-)	1/36.9 (2.7)
	M13-549	2/49.2 (4.1)	1/49.3 (2.0)	3/47.7 (6.3)
Neutropenia	M13-542	0/36.0 (-)	5/36.6 (13.7)	4/36.3 (11)
	M13-549	0/49.4 (-)	5/48.7 (10.3)	8/47.1 (17)
Lymphopenia	M13-542	2/35.7 (5.6)	2/37.1 (5.4)	2/36.7 (5.5)
	M13-549	1/49.2 (2)	1/49.3 (2)	5/47.6 (10.5)
Herpes Zoster	M13-542	1/35.8 (2.8)	1/37.3 (2.7)	4/36.7 (10.9)
	M13-549	1/49.4 (2)	1/49.3 (2)	2/48.0 (4.2)
Active/Latent TB	M13-542	0/36.0 (-)	0/37.3 (-)	0/36.9 (-)
	M13-549	0/49.4 (-)	0/49.5 (-)	0/48.1 (-)
MACE	M13-542	0/36.0 (-)	1/37.3 (2.7)	0/36.9 (-)
	M13-549	0/49.4 (-)	0/49.5 (-)	1/48.1 (2.1)
Any Adjudicated VTE	M13-542	0/36.0 (-)	1/37.3 (2.7)	0/36.9 (-)
	M13-549	0/49.4 (-)	0/49.5 (-)	0/48.1 (-)

Total number of patients with at least one AE/Total exposure and EAIR in parenthesis are reported for each study
Abbreviations: MTX=methotrexate; UPA=upadacitinib; QD=once daily; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; EAIR=exposure-adjusted incidence rates;
[Source: Statistical Reviewer]

Table 66 Exposure Adjusted Incidence Rates for Treatment Emergent Adverse Events of Special Interests During Placebo-Controlled Period 12/14 Weeks

	Placebo + cDMARD (N=390)	UPA 15 mg QD + cDMARD (N=385)	UPA 15 mg Vs Placebo
	E /Exposure (100PY) (IR /100 PY) [1]	E /Exposure (100PY) (IR /100 PY) [1]	Diff [95% CI] [2]
Infections	218/228.2 (95.6)	284/222.0 (128)	32.2 [12.6, 51.7]
Serious Infections	6/256.6 (2.3)	12/258.3 (4.6)	2.3 [-0.9, 5.5]
Opportunistic Infection	3/256.8 (1.2)	5/258.5 (1.9)	0.8 [-1.4, 2.9]
Malignancy	2/256.7 (0.8)	1/259.3 (0.4)	-0.4 [-1.7, 0.9]
Malignancy (excluding NMSC)	1/256.8 (0.4)	1/259.3 (0.4)	-0.0 [-1.1, 1.1]
Lymphoma	0/256.8 (-)	0/259.3 (-)	-
Any Hepatic Disorder	38/252.5 (15)	46/253.4 (18.2)	3.2 [-3.9, 10.3]
Any GI perforation	0/256.8 (-)	2/259.1 (0.8)	0.8 [-0.3, 1.8]
Anemia	21/254.6 (8.2)	11/258.5 (4.3)	-4.0 [-8.3, 0.3]
Neutropenia	2/256.7 (0.8)	23/256.4 (9)	8.2 [4.4, 12.0]
Lymphopenia	11/254.9 (4.3)	14/257.3 (5.4)	1.1 [-2.7, 4.9]
Herpes Zoster	3/256.4 (1.2)	7/258.7 (2.7)	1.5 [-0.9, 3.9]
Active/Latent TB	0/256.8 (-)	1/259.4 (0.4)	0.4 [-0.4, 1.1]
MACE	3/256.8 (1.2)	1/259.3 (0.4)	-0.8 [-2.3, 0.7]
Adjudicated VTEs	1/256.8 (0.4)	2/259.2 (0.8)	0.4 [-0.9, 1.7]

Treatment-emergent adverse event is defined as any adverse event with an onset date on or after the first dose of study drug in period 1 and prior to the week 14 dose date in period 1 or up to 30 days after the last dose of placebo or upadacitinib and 70 days for adalimumab, if subject discontinued study drug prematurely before week 14 dosing in period 1 of the study.

1: EAIR rates were stratified by study using MH weights using total exposure within each study.

2: Difference in EAIR comparing upadacitinib 15 mg with placebo and respective 95% CI accounting for total exposure within each study via MH weights were reported

Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; EOW=every other week; incidence rates; PY=person years; CI=confidence intervals

[Source: Statistical Reviewer]

Table 67 EAIR for Treatment Emergent AEs of Special Interest up to 12/14 Weeks During Active-Controlled Period for the Monotherapy Studies

AE Category	Study ID	MTX Events/Exposure (EAIR)	UPA 15 mg QD Events/Exposure (EAIR)	UPA 30 mg QD Events/Exposure (EAIR)
Infections	M13-545	70/60.6 (116)	63/63.3 (99.5)	74/61.7 (120)
	M15-555	57/45.8 (124)	41/50.1 (81.9)	54/49.5 (109)
Serious Infections	M13-545	1/68.0 (1.5)	2/71.0 (2.8)	8/69.9 (11.4)
	M15-555	1/53.7 (1.9)	1/55.5 (1.8)	0/56.4 (-)
Opportunistic Infection	M13-545	0/68.0 (-)	0/71.1 (-)	1/70.3 (1.4)
	M15-555	1/53.5 (1.9)	0/55.6 (-)	3/55.8 (5.4)
Confirmed Malignancy	M13-545	1/67.9 (1.5)	1/71.2 (1.4)	0/70.3 (-)
	M15-555	1/53.6 (1.9)	2/55.6 (3.6)	0/56.4 (-)
Malignancy (Excl NMSC)	M13-545	1/67.9 (1.5)	1/71.2 (1.4)	0/70.3 (-)
	M15-555	0/53.7 (-)	2/55.6 (3.6)	0/56.4 (-)
Lymphoma	M13-545	0/68.0 (-)	0/71.1 (-)	0/70.3 (-)
	M15-555	0/53.7 (-)	1/55.6 (1.8)	0/56.4 (-)
Any Hepatic Disorder	M13-545	6/67.1 (8.9)	15/68.9 (21.8)	11/69.2 (15.9)
	M15-555	4/53.1 (7.5)	4/54.8 (7.3)	5/55.7 (9)
Any GI perforation	M13-545	0/68.0 (-)	0/71.1 (-)	2/70.3 (2.8)
	M15-555	0/53.7 (-)	0/55.6 (-)	0/56.4 (-)
Anemia	M13-545	4/67.4 (5.9)	5/70.2 (7.1)	5/69.6 (7.2)
	M15-555	2/53.5 (3.7)	0/55.6 (-)	1/56.3 (1.8)
Neutropenia	M13-545	2/67.7 (3.0)	6/70.4 (8.5)	10/68.8 (14.5)
	M15-555	1/53.5 (1.9)	2/55.2 (3.6)	4/55.6 (7.2)
Lymphopenia	M13-545	5/67.3 (7.4)	2/70.7 (2.8)	2/70.1 (2.9)
	M15-555	2/53.3 (3.8)	0/55.6 (-)	2/56.0 (3.6)
Herpes Zoster	M13-545	1/67.9 (1.5)	3/70.8 (4.2)	2/70.3 (2.8)
	M15-555	1/53.5 (1.9)	3/55.2 (5.4)	6/55.9 (10.7)
Active/Latent TB	M13-545	0/68.0 (-)	0/71.1 (-)	0/70.3 (-)
	M15-555	0/53.7 (-)	0/55.6 (-)	0/56.4 (-)
MACE	M13-545	1/68.1 (1.5)	1/71.1 (1.4)	1/70.4 (1.4)
	M15-555	0/53.7 (-)	1/55.6 (1.8)	2/56.3 (3.6)
Any Adjudicated VTE	M13-545	0/68.0 (-)	0/71.1 (-)	1/70.3 (1.4)
	M15-555	0/53.7 (-)	1/55.6 (1.8)	0/56.4 (-)

Total number of patients with at least one AE/Total exposure and EAIR in parenthesis are reported for each study
Abbreviations: MTX=methotrexate; UPA=upadacitinib; QD=once daily; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; EAIR=exposure-adjusted incidence rates;
[Source: Statistical Reviewer]

Table 68 EAIR Treatment Emergent AEs of Special Interest Up to 1-yr Exposure for the Two Doses of Upadacitinib (Part 1/3) by Study

AE Category	Study ID	UPA 15 mg QD Events/Exposure (EAIR)	UPA 30 mg QD Events/Exposure (EAIR)	UPA 30 mg vs 15 mg Diff [95 % CI]
Infections	M13-542	120/132.6 (90.5)	146/114.2 (128)	37.3 [-20.4, 95.0]
	M13-549	176/185.0 (95.1)	182/173.3 (105)	9.9 [-47.8, 67.6]
	M13-545	165/210.7 (78.3)	179/197.0 (90.9)	12.5 [-45.2, 70.2]
	M15-555	154/205.3 (75)	167/193.1 (86.5)	11.5 [-46.2, 69.2]
Serious Infections	M13-542	8/209.0 (3.8)	13/200.9 (6.5)	2.6 [-55.0, 60.3]
	M13-549	7/289.8 (2.4)	19/273.0 (7)	4.5 [-53.1, 62.2]
	M13-545	11/302.1 (3.6)	15/294.0 (5.1)	1.5 [-56.2, 59.2]
	M15-555	12/280.7 (4.3)	12/281.3 (4.3)	-0.0 [-57.7, 57.7]
Opportunistic Infection	M13-542	2/208.9 (1)	3/201.5 (1.5)	0.5 [-57.2, 58.2]
	M13-549	2/291.7 (0.7)	5/274.0 (1.8)	1.1 [-56.6, 58.8]
	M13-545	1/304.4 (0.3)	2/297.0 (0.7)	0.3 [-57.3, 58.0]
	M15-555	2/283.9 (0.7)	5/280.3 (1.8)	1.1 [-56.6, 58.8]
Herpes Zoster	M13-542	7/208.3 (3.4)	15/197.1 (7.6)	4.2 [-53.4, 61.9]
	M13-549	12/287.2 (4.2)	20/267.1 (7.5)	3.3 [-54.4, 61.0]
	M13-545	13/297.7 (4.4)	12/294.6 (4.1)	-0.3 [-58.0, 57.4]
	M15-555	9/279.8 (3.2)	17/275.1 (6.2)	3.0 [-54.7, 60.7]
Gastrointestinal Perforations	M13-542	0/210.4 (-)	2/203.0 (1)	1.0 [-56.7, 58.7]
	M13-549	0/292.3 (-)	0/276.1 (-)	0.0 [-57.7, 57.7]
	M13-545	0/304.4 (-)	2/298.1 (0.7)	0.7 [-57.0, 58.4]
	M15-555	1/284.3 (0.4)	1/284.3 (0.4)	-0.0 [-57.7, 57.7]
Active/Latent Tuberculosis	M13-542	1/210.3 (0.5)	1/203.0 (0.5)	0.0 [-57.7, 57.7]
	M13-549	10/291.0 (3.4)	3/275.5 (1.1)	-2.3 [-60.0, 55.3]
	M13-545	8/303.9 (2.6)	8/297.2 (2.7)	0.1 [-57.6, 57.8]
	M15-555	8/283.9 (2.8)	7/283.3 (2.5)	-0.3 [-58.0, 57.3]

Estimated difference and 95% CI are based on the difference in EAIR within each study.
Abbreviations: UPA=upadacitinib; QD=once daily; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; CI=confidence intervals; MH=Mantel Haenszel; EAIR=exposure-adjusted incidence rates; E=number of patients with at least 1 event
[Source: Statistical Reviewer]

Table 69 EAIR Treatment Emergent AEs of Special Interest Up to 1-yr Exposure for the Two Doses of Upadacitinib (Part 2/3) by Study

AE Category	Study ID	UPA 15 mg QD Events/Exposure (EAIR)	UPA 30 mg QD Events/Exposure (EAIR)	UPA 30 mg vs 15 mg Diff [95 % CI]
MACE	M13-542	1/209.6 (0.5)	1/203.1 (0.5)	0.0 [-57.7, 57.7]
	M13-549	3/291.9 (1)	3/275.8 (1.1)	0.1 [-57.6, 57.8]
	M13-545	2/304.4 (0.7)	3/297.5 (1)	0.4 [-57.3, 58.0]
	M15-555	3/284.2 (1.1)	5/283.1 (1.8)	0.7 [-57.0, 58.4]
Malignancy Excl NMSC	M13-542	2/210.2 (1)	3/202.9 (1.5)	0.5 [-57.2, 58.2]
	M13-549	4/291.6 (1.4)	6/275.8 (2.2)	0.8 [-56.9, 58.5]
	M13-545	3/303.9 (1)	2/297.8 (0.7)	-0.3 [-58.0, 57.4]
	M15-555	4/284.4 (1.4)	3/284.3 (1.1)	-0.4 [-58.0, 57.3]
Any VTE	M13-542	4/209.5 (1.9)	2/202.3 (1)	-0.9 [-58.6, 56.8]
	M13-549	0/292.3 (-)	1/276.1 (0.4)	0.4 [-57.3, 58.1]
	M13-545	0/304.4 (-)	1/298.1 (0.3)	0.3 [-57.4, 58.0]
	M15-555	1/284.5 (0.4)	0/284.3 (-)	-0.4 [-58.0, 57.3]
Lymphoma	M13-542	0/210.4 (-)	0/203.1 (-)	0.0 [-57.7, 57.7]
	M13-549	0/292.3 (-)	1/275.9 (0.4)	0.4 [-57.3, 58.1]
	M13-545	0/304.4 (-)	0/298.1 (-)	0.0 [-57.7, 57.7]
	M15-555	1/284.4 (0.4)	0/284.3 (-)	-0.4 [-58.0, 57.3]
Hepatic Disorder	M13-542	10/206.2 (4.8)	12/197.0 (6.1)	1.2 [-56.5, 58.9]
	M13-549	34/278.1 (12.2)	23/264.0 (8.7)	-3.5 [-61.2, 54.2]
	M13-545	29/286.2 (10.1)	28/286.3 (9.8)	-0.4 [-58.0, 57.3]
	M15-555	11/278.3 (4)	24/272.6 (8.8)	4.9 [-52.8, 62.5]

Estimated difference and 95% CI are based on the difference in EAIR within each study.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; CI=confidence intervals; MH=Mantel Haenszel ; EAIR=exposure-adjusted incidence rates; E=number of patients with at least 1 event
[Source: Statistical Reviewer]

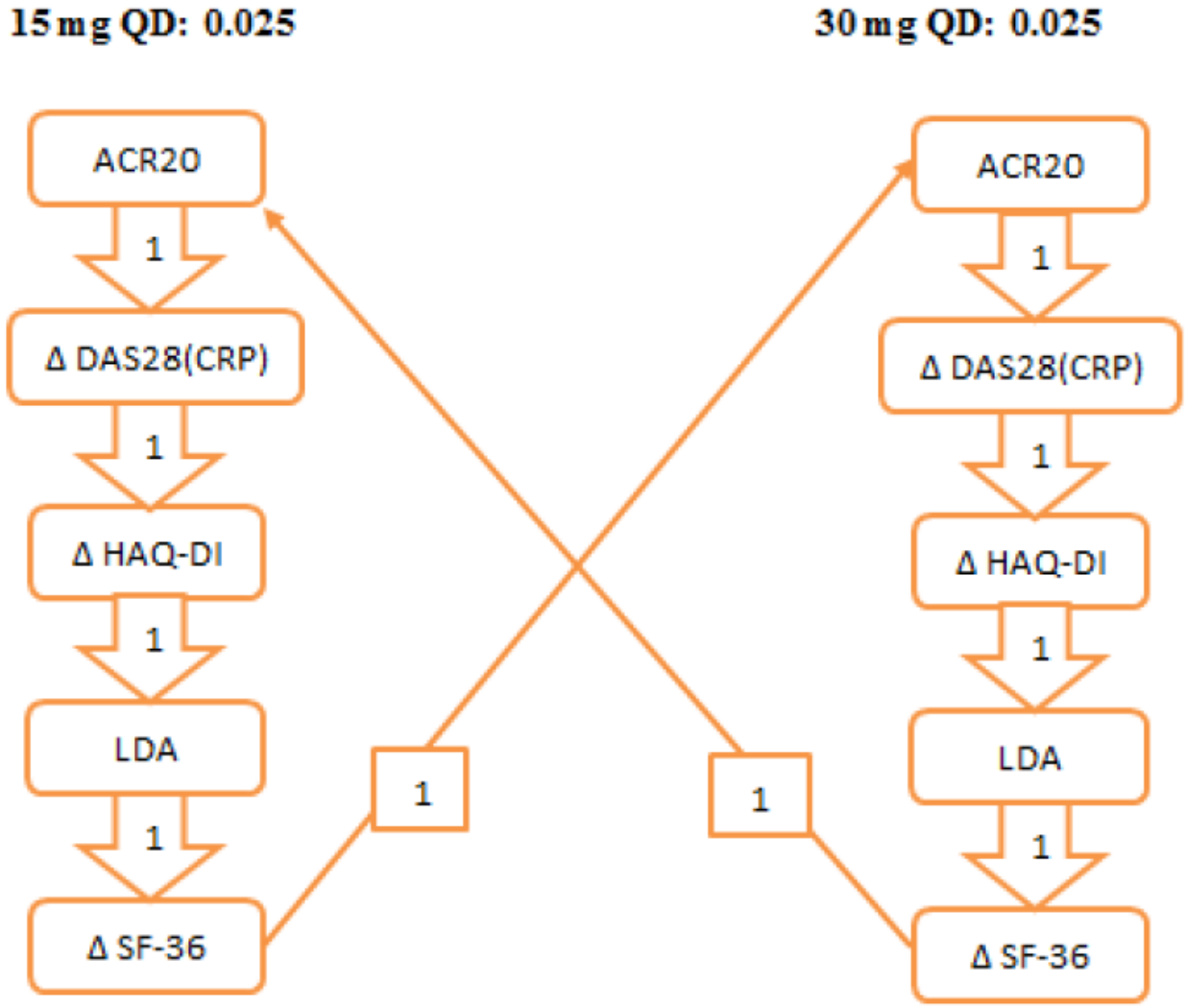
Table 70 EAIR Treatment Emergent AEs of Special Interest Up to 1-yr Exposure for the Two Doses of Upadacitinib (Part 3/3) by Study

AE Category	Study ID	UPA 15 mg QD Events/Exposure (EAIR)	UPA 30 mg QD Events/Exposure (EAIR)	UPA 30 mg vs 15 mg Diff [95 % CI]
Creatine Phosphokinase	M13-542	12/205.0 (5.9)	13/195.6 (6.6)	0.8 [-56.9, 58.5]
	M13-549	20/282.9 (7.1)	21/264.6 (7.9)	0.9 [-56.8, 58.6]
	M13-545	23/291.4 (7.9)	45/268.3 (16.8)	8.9 [-48.8, 66.6]
	M15-555	18/274.7 (6.6)	38/262.3 (14.5)	7.9 [-49.8, 65.6]
Anemia	M13-542	8/207.6 (3.9)	12/195.6 (6.1)	2.3 [-55.4, 60.0]
	M13-549	12/287.5 (4.2)	17/270.1 (6.3)	2.1 [-55.6, 59.8]
	M13-545	15/295.4 (5.1)	19/286.6 (6.6)	1.6 [-56.1, 59.2]
	M15-555	10/279.7 (3.6)	14/279.0 (5)	1.4 [-56.2, 59.1]
Neutropenia	M13-542	10/203.5 (4.9)	7/198.2 (3.5)	-1.4 [-59.1, 56.3]
	M13-549	10/285.7 (3.5)	21/264.9 (7.9)	4.4 [-53.3, 62.1]
	M13-545	13/295.1 (4.4)	23/285.0 (8.1)	3.7 [-54.0, 61.4]
	M15-555	4/281.4 (1.4)	14/274.5 (5.1)	3.7 [-54.0, 61.4]
Lymphopenia	M13-542	4/207.7 (1.9)	6/199.9 (3)	1.1 [-56.6, 58.8]
	M13-549	3/290.3 (1)	10/270.4 (3.7)	2.7 [-55.0, 60.4]
	M13-545	4/302.1 (1.3)	8/293.7 (2.7)	1.4 [-56.3, 59.1]
	M15-555	2/282.9 (0.7)	3/281.6 (1.1)	0.4 [-57.3, 58.1]
Renal Dysfunction	M13-542	1/210.3 (0.5)	4/202.6 (2)	1.5 [-56.2, 59.2]
	M13-549	1/292.3 (0.3)	2/275.2 (0.7)	0.4 [-57.3, 58.1]
	M13-545	1/304.4 (0.3)	4/297.5 (1.3)	1.0 [-56.7, 58.7]
	M15-555	1/284.5 (0.4)	3/284.3 (1.1)	0.7 [-57.0, 58.4]

Estimated difference and 95% CI are based on the difference in EAIR within each study.
Abbreviations: cDMARD=conventional disease modifying anti-rheumatic drugs; UPA=upadacitinib; QD=once daily; CMH=Cochran Mantel Haenszel; NMSC=non-melanoma skin cancer; TB=tuberculosis; MACE=major adverse cardiovascular events; VTE=venous thromboembolic events; GI=gastrointestinal; incidence rates; PY=person years; CI=confidence intervals; MH=Mantel Haenszel ; EAIR=exposure-adjusted incidence rates; E=number of patients with at least 1 event
[Source: Statistical Reviewer]

6.4 Additional Figures

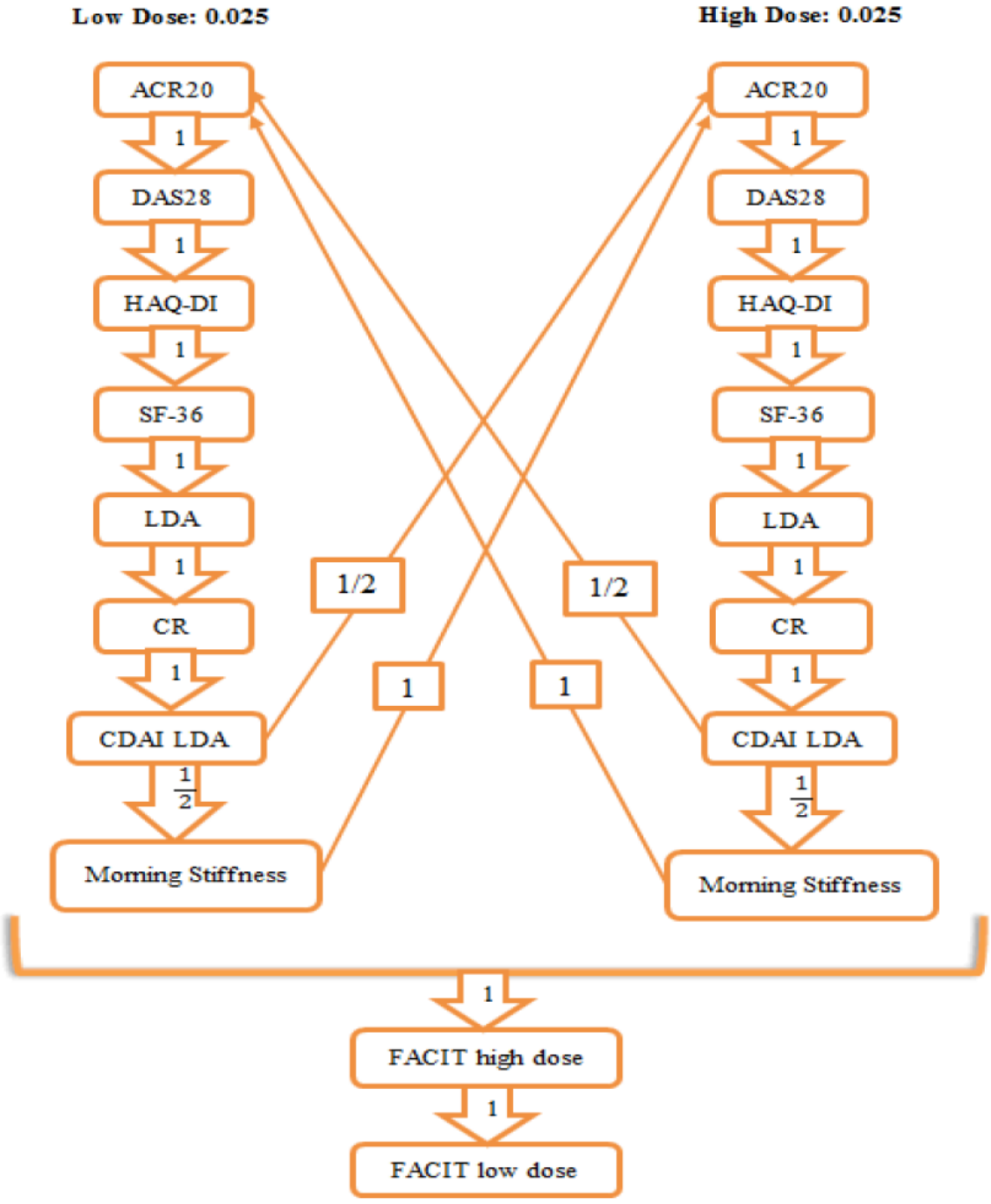
Figure 9 Graphical Testing Procedure, Study M13-542



Abbreviations: ACR=American College of Rheumatology; CRP=C-reactive protein; DAS28(CRP)=Disease Activity Score 28 based on CRP; HAQ-DI=health assessment questionnaire-disability index; LDA=low disease activity defined based on DAS28(CRP)<10; SF-36=short-form 36-item survey; Δ=change from baseline; QD=once daily

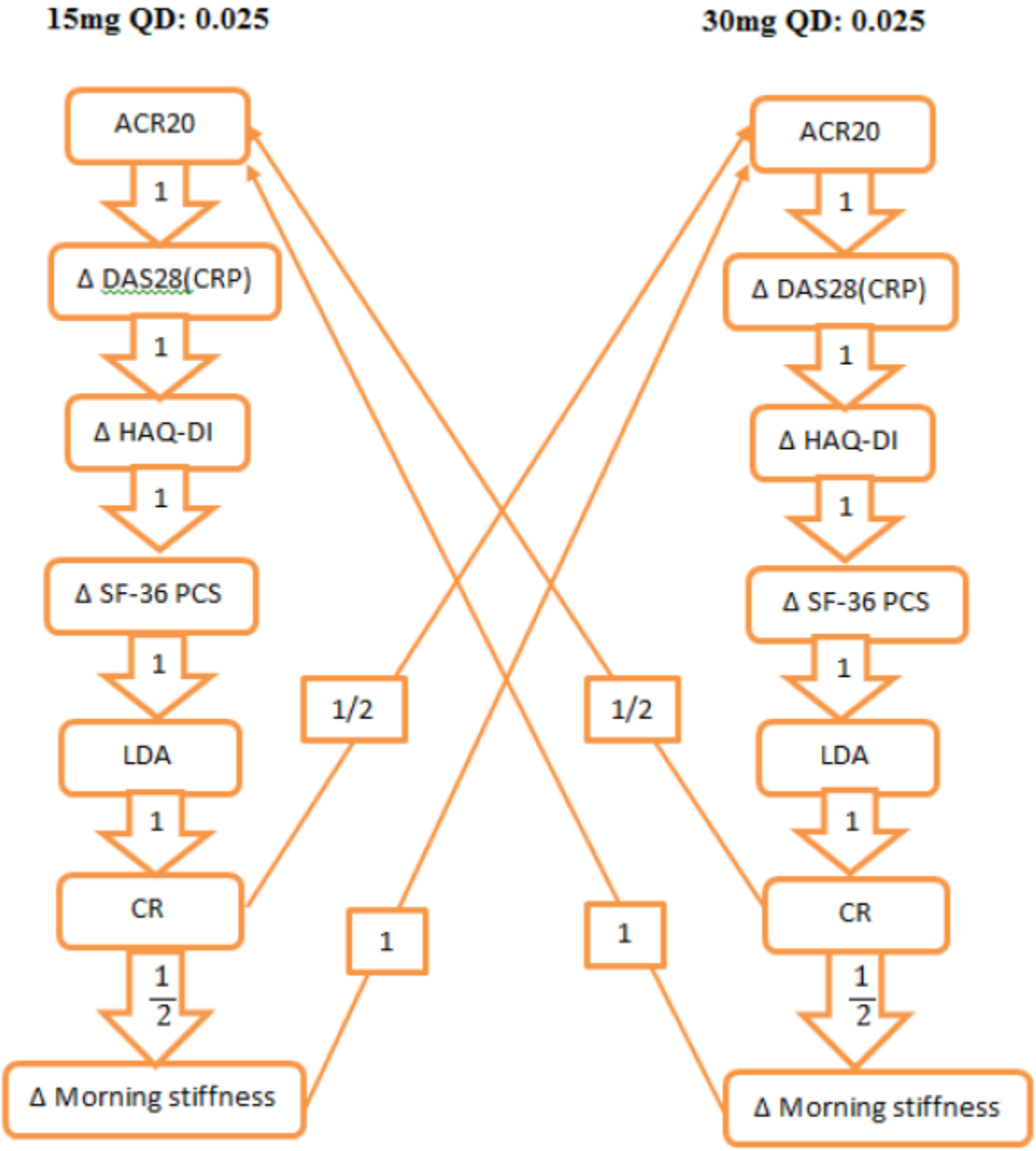
[Source: M13-542 SAP]

Figure 10 Graphical Testing Procedure, Study M13-549



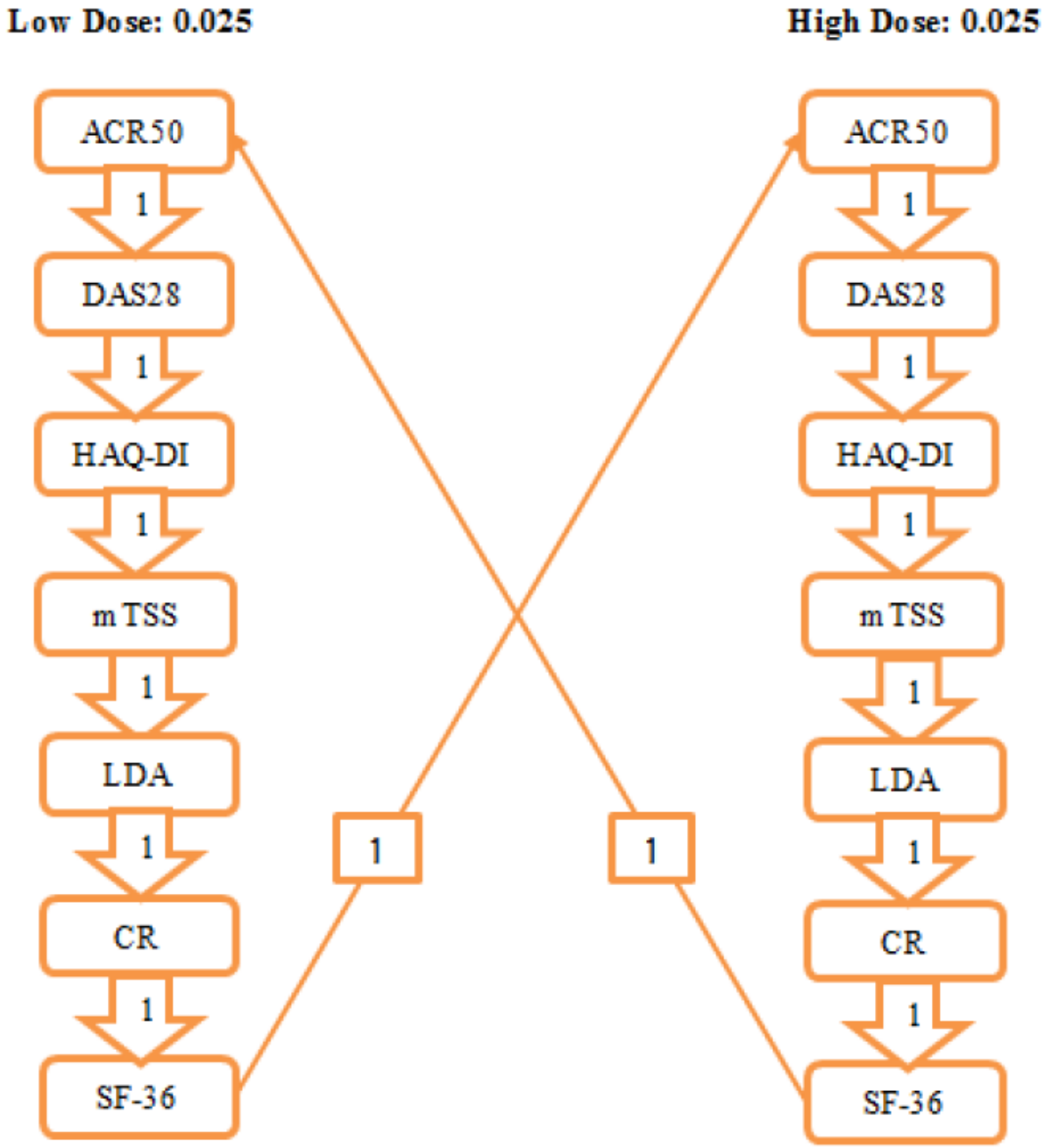
Abbreviations: ACR=American College of Rheumatology; CRP=C-reactive protein; DAS28(CRP)=Disease Activity Score 28 based on CRP; HAQ-DI=health assessment questionnaire-disability index; LDA=low disease activity defined based on DAS28(CRP)<10; SF-36=short-form 36-item survey; Δ=change from baseline; QD=once daily; CDAI LDA=clinical disease activity index LDA; FACIT-F=functional assessment of chronic illness-fatigue [Source: M13-549 SAP]

Figure 11 Graphical Testing Procedure, Study M15-555



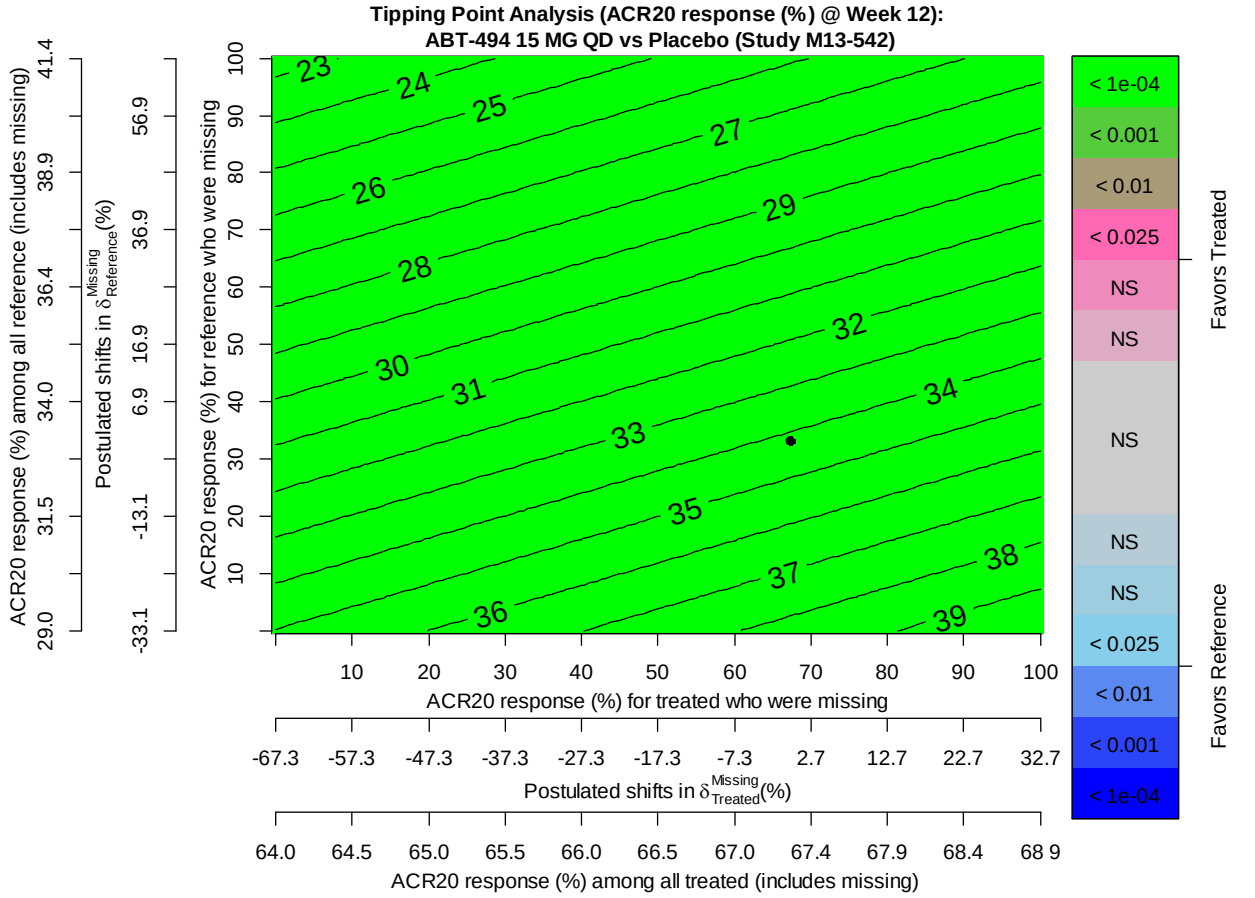
Abbreviations: ACR=American College of Rheumatology; CRP=C-reactive protein; DAS28(CRP)=Disease Activity Score 28 based on CRP; HAQ-DI=health assessment questionnaire-disability index; LDA=low disease activity defined based on DAS28(CRP)<10; SF-36=short-form 36-item survey; Δ=change from baseline; QD=once daily; CDAI LDA=clinical disease activity index LDA; FACIT-F=functional assessment of chronic illness-fatigue [Source: M15-555 SAP]

Figure 12 Graphical Testing Procedure, Study M13-545



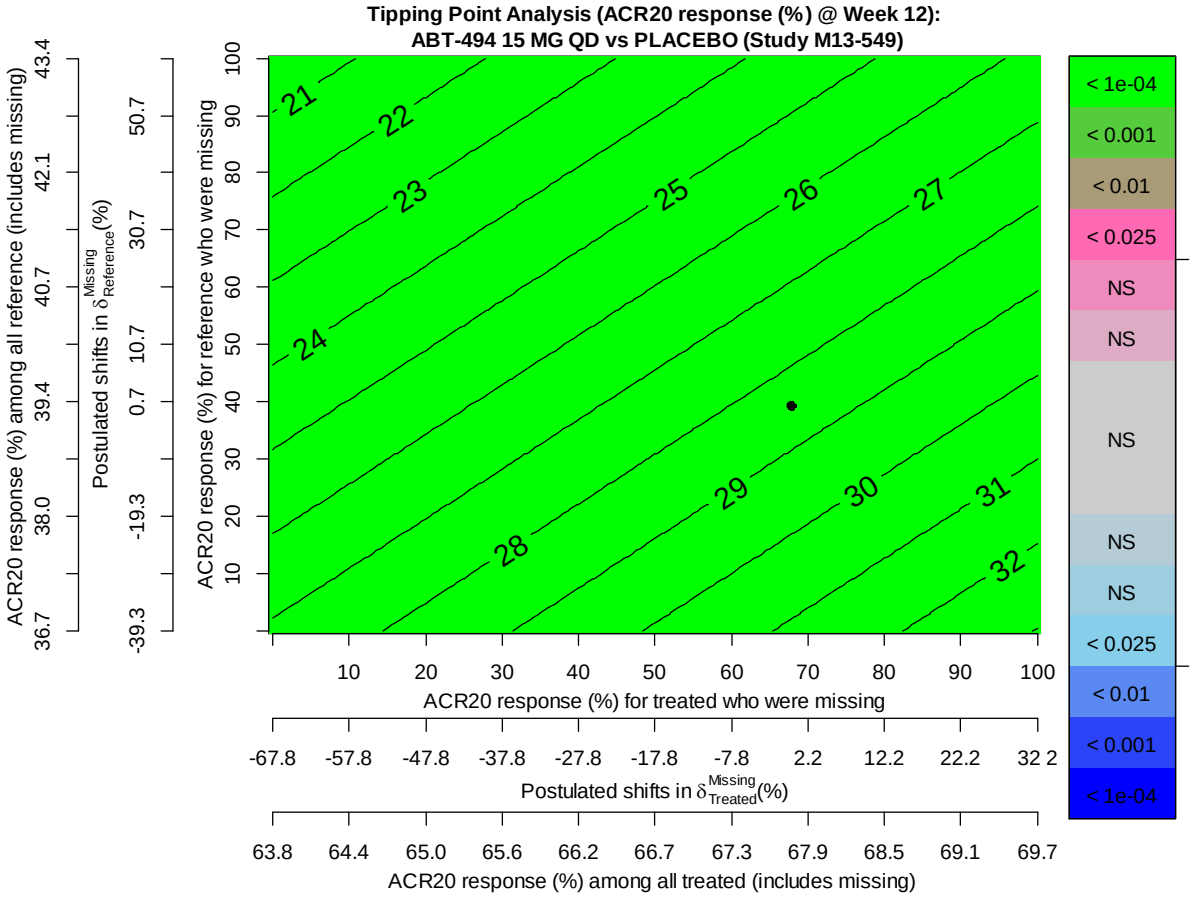
Abbreviations: ACR=American College of Rheumatology; CRP=C-reactive protein; DAS28(CRP)=Disease Activity Score 28 based on CRP; HAQ-DI=health assessment questionnaire-disability index; LDA=low disease activity defined based on DAS28(CRP)<10; SF-36=short-form 36-item survey; Δ=change from baseline; QD=once daily; mTSS=modified Total Sharp Score; CR=clinical remission based on DAS28(CRP)<2.6
[Source: M13-545 SAP]

Figure 13 Tipping Point Sensitivity Analysis for the Probability of ACR20 Response at Week 12 Comparing Upadacitinib (ABT-494) 15 mg QD vs Placebo, Study M13-542



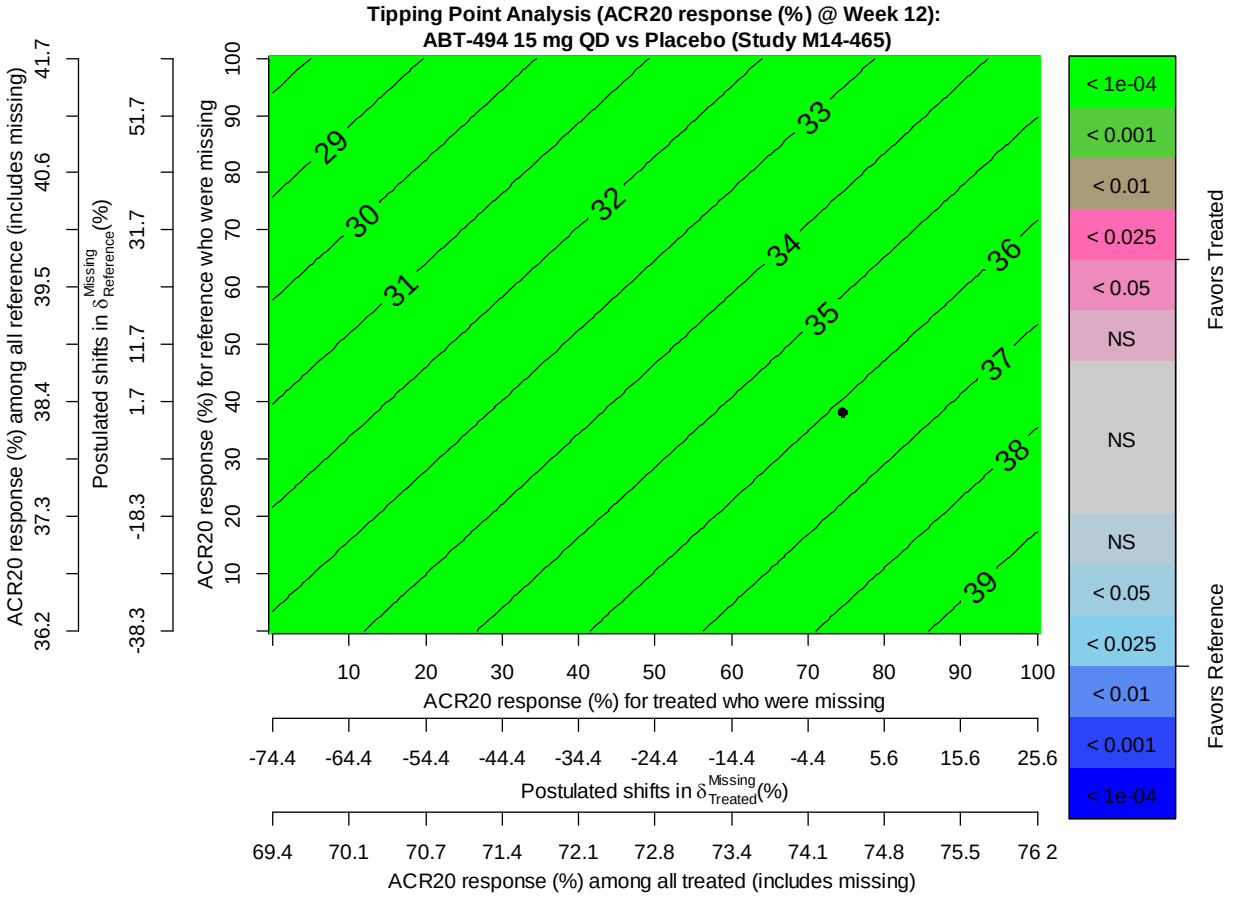
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025
[Source: Statistical Reviewer]

Figure 14 Tipping Point Sensitivity Analysis for the Probability of ACR20 Response at Week 12 Comparing Upadacitinib 15 mg QD vs Placebo, Study M13-549



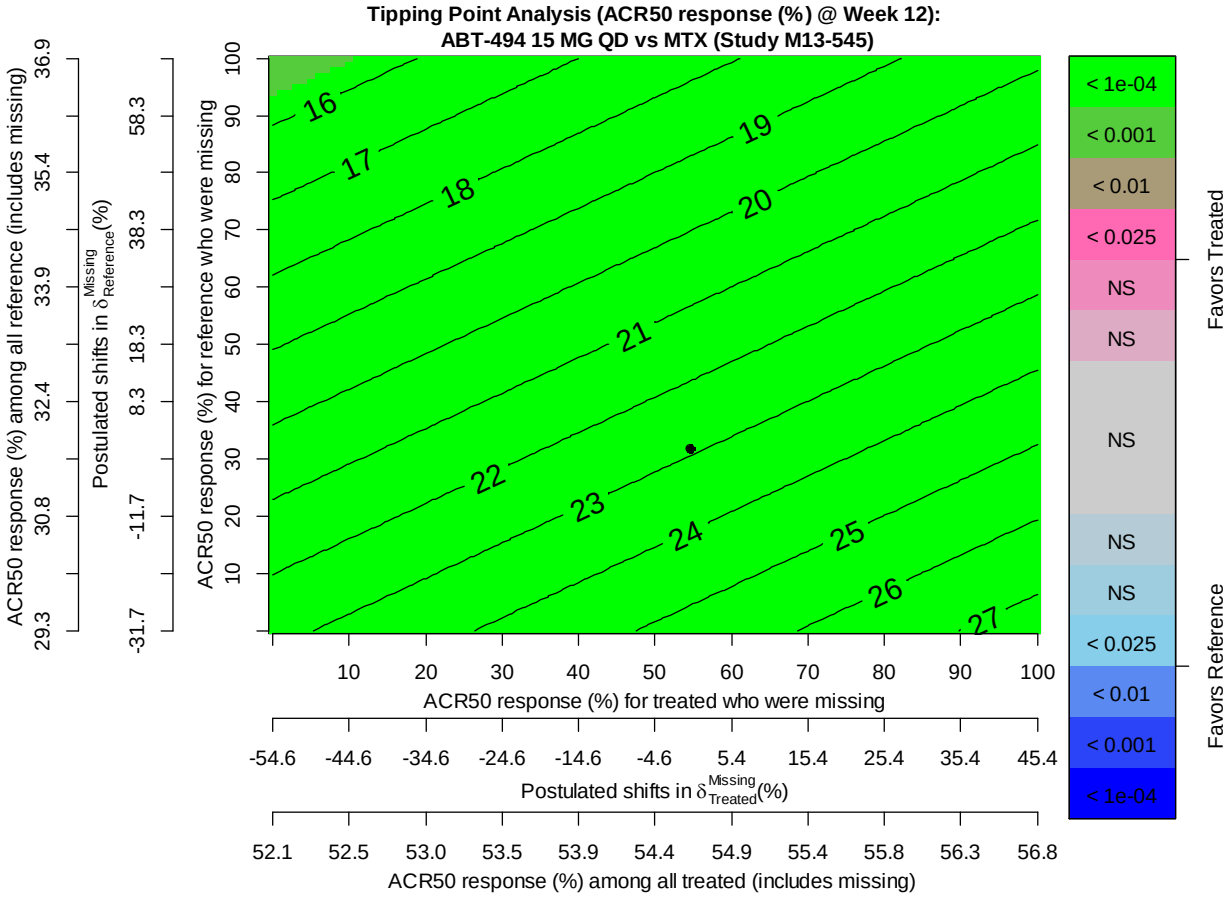
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025
[Source: Statistical Reviewer]

Figure 15 Tipping Point Sensitivity Analysis for the Probability of ACR20 Response at Week 12 Comparing Upadacitinib 15 mg QD vs Placebo, Study M14-465



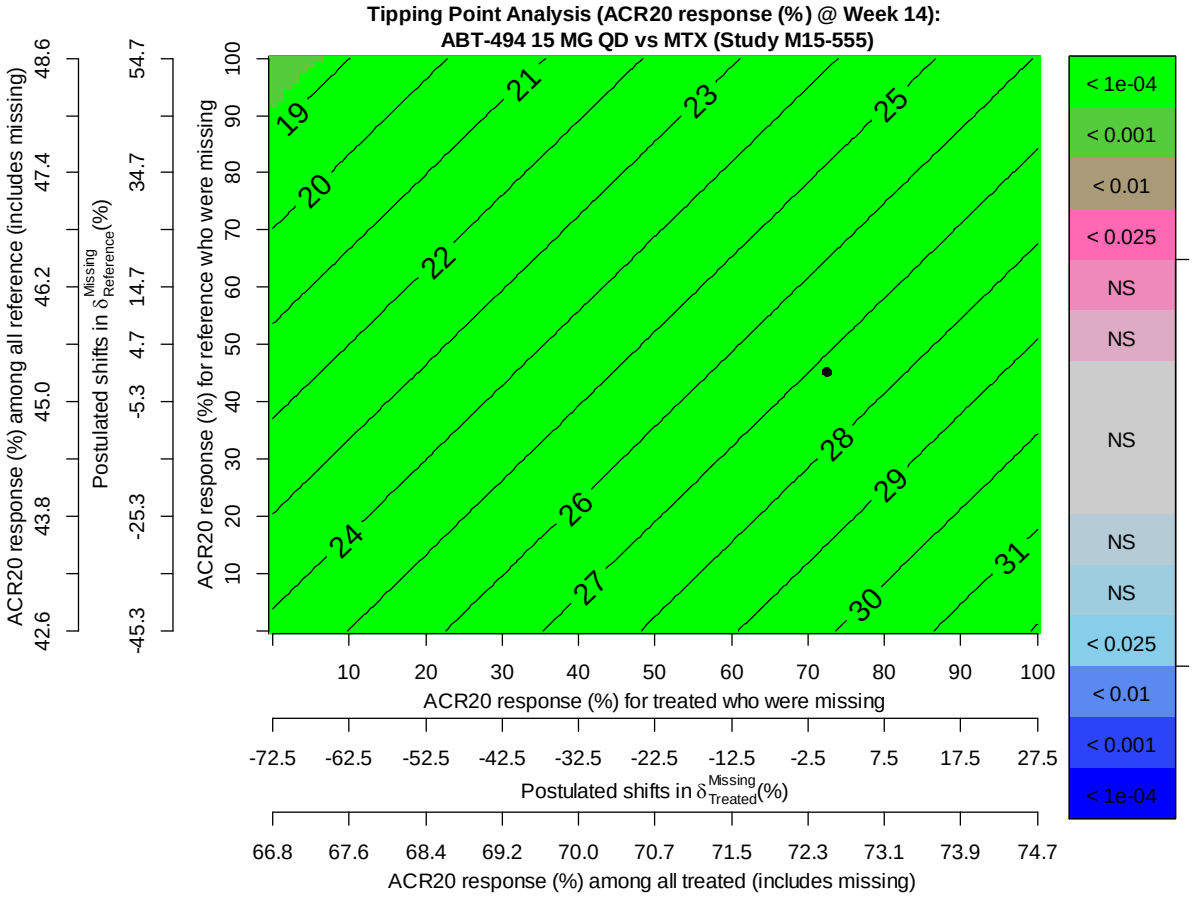
Abbreviations: ACR=American College of Rheumatology; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.05
[Source: Statistical Reviewer]

Figure 16 Tipping Point Sensitivity Analysis for the Probability of ACR50 Response at Week 12 Comparing Upadacitinib 15 mg QD vs MTX Monotherapy, Study M13-545



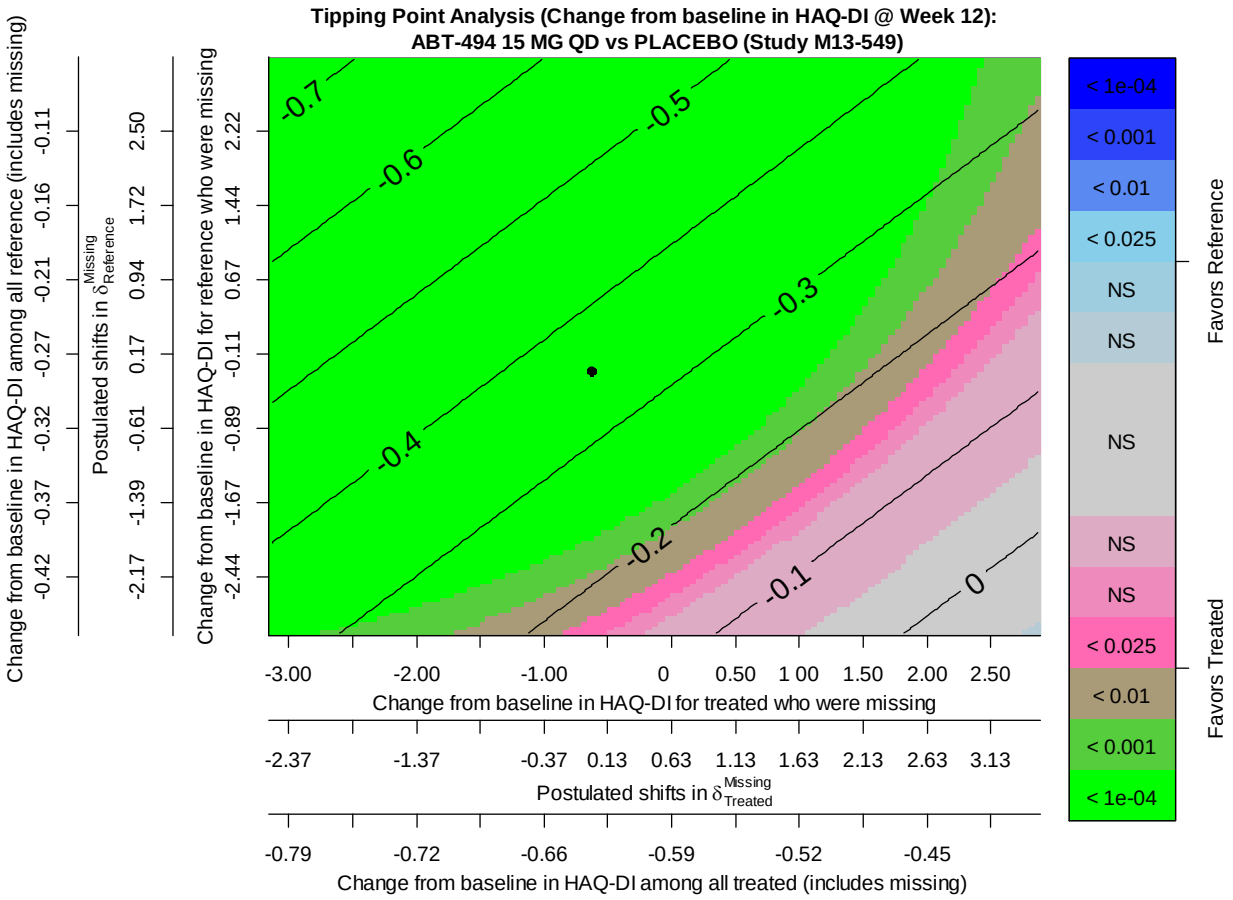
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025; MTX=methotrexate
[Source: Statistical Reviewer]

Figure 17 Tipping Point Sensitivity Analysis for the Probability of ACR20 Response at Week 14 Comparing Upadacitinib 15 mg QD Monotherapy vs MTX Monotherapy, Study M15-555



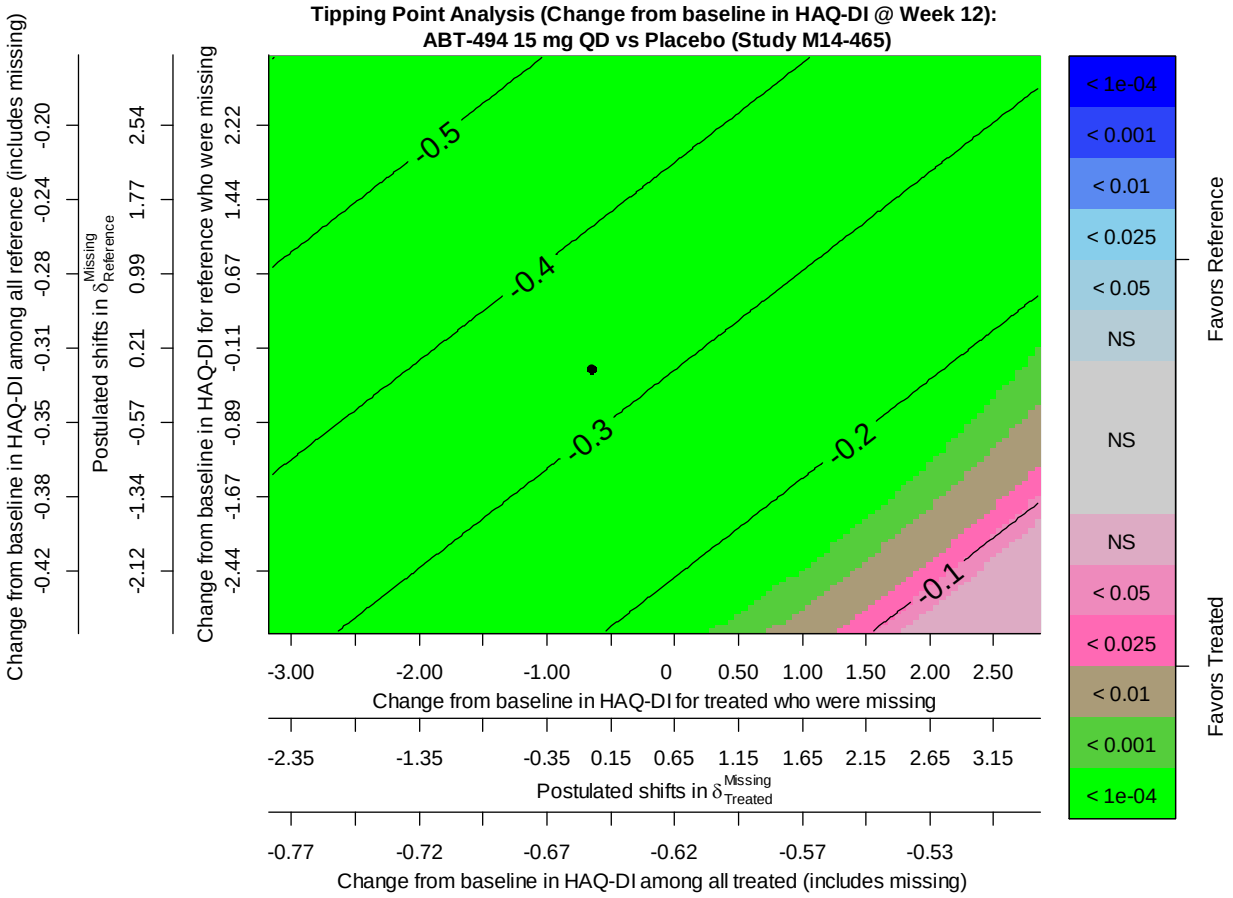
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025; MTX=methotrexate
[Source: Statistical Reviewer]

Figure 18 Tipping Point Sensitivity Analysis for the Mean Change from Baseline in HAQ-DI at Week 12 Comparing Upadacitinib 15 mg QD vs Placebo, Study M13-549



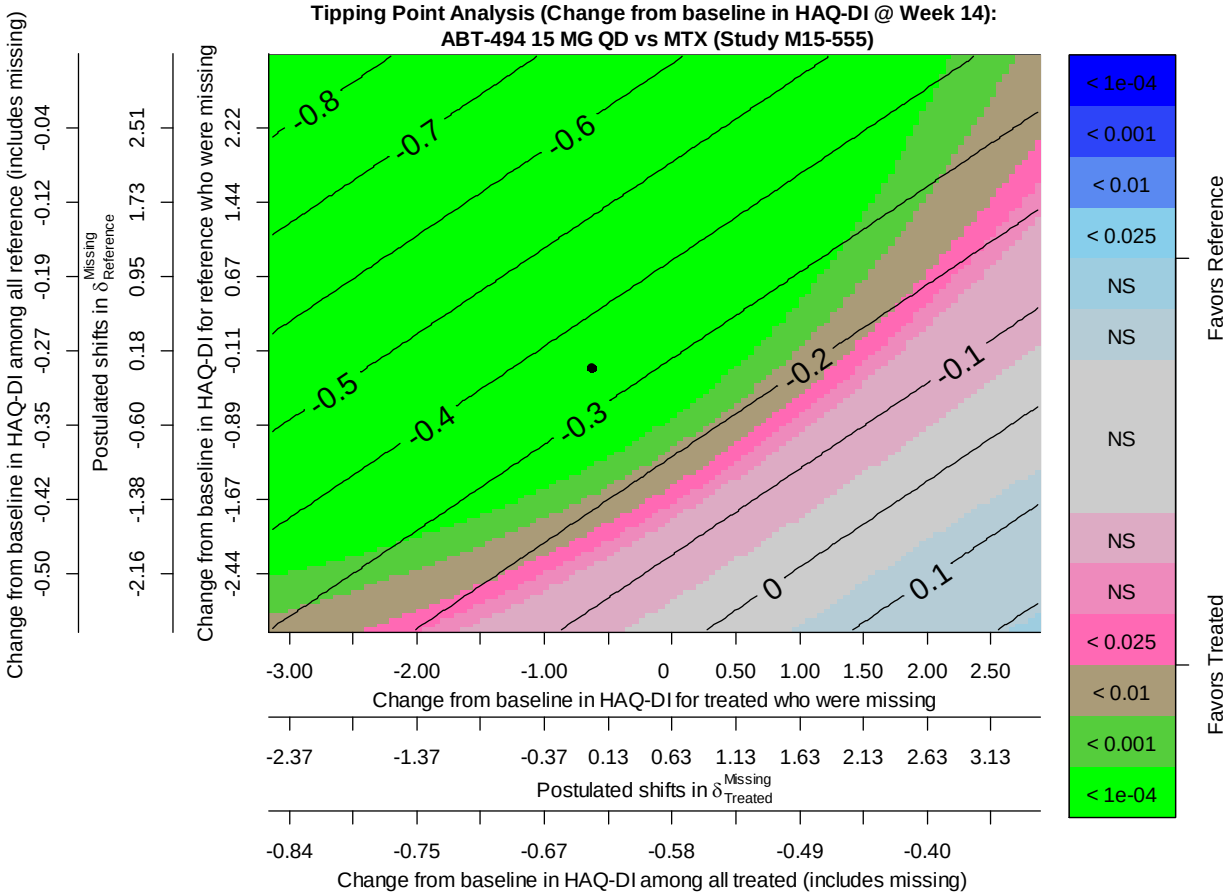
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: HAQ-DI=health assessment questionnaire-disability index; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025
[Source: Statistical Reviewer]

Figure 19 Tipping Point Sensitivity Analysis for the Mean Change from Baseline in HAQ-DI at Week 12 Comparing Upadacitinib 15 mg QD vs Placebo, Study M14-465



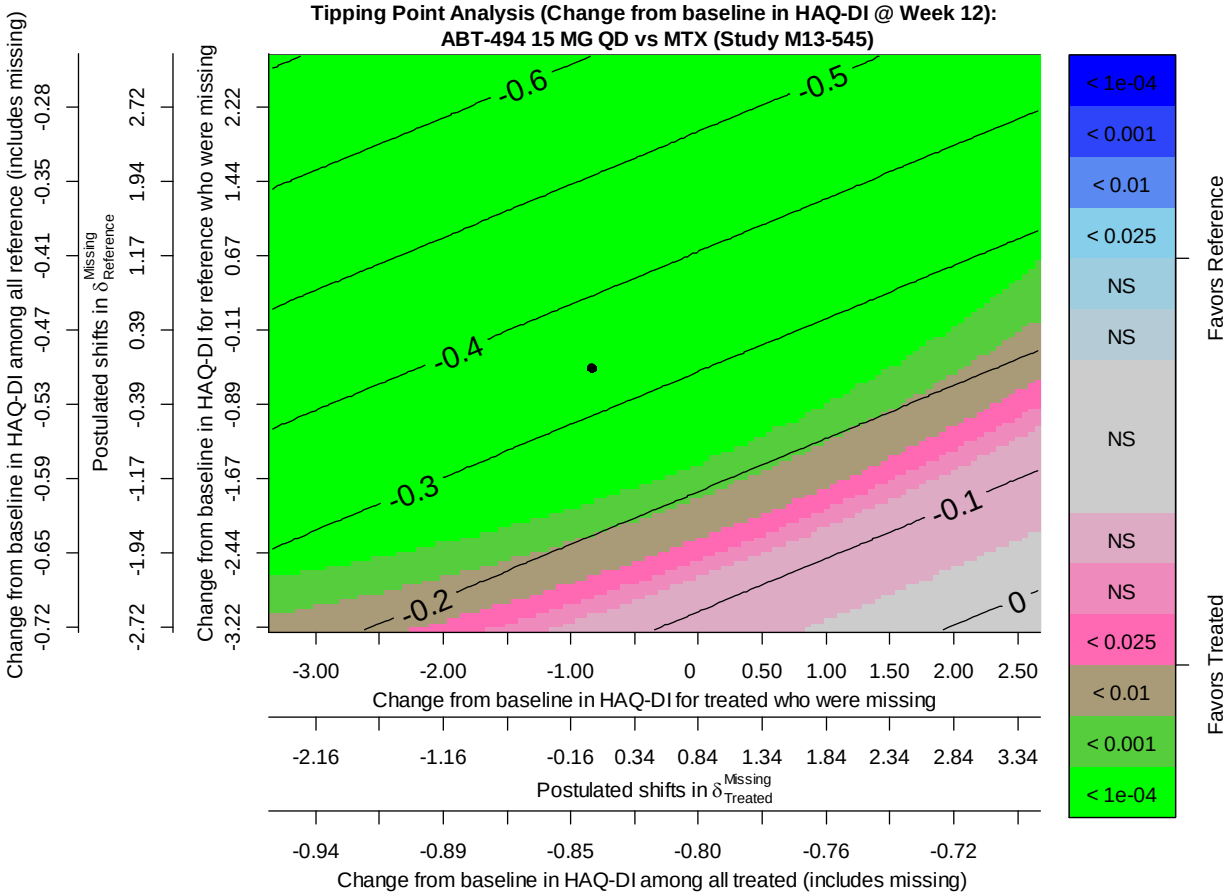
Abbreviations: HAQ-DI=health assessment questionnaire-disability index; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.05
[Source: Statistical Reviewer]

Figure 20 Tipping Point Sensitivity Analysis for the Mean Change from Baseline in HAQ-DI at Week 14 Comparing Upadacitinib 15 mg QD Monotherapy vs MTX Monotherapy, Study M15-555



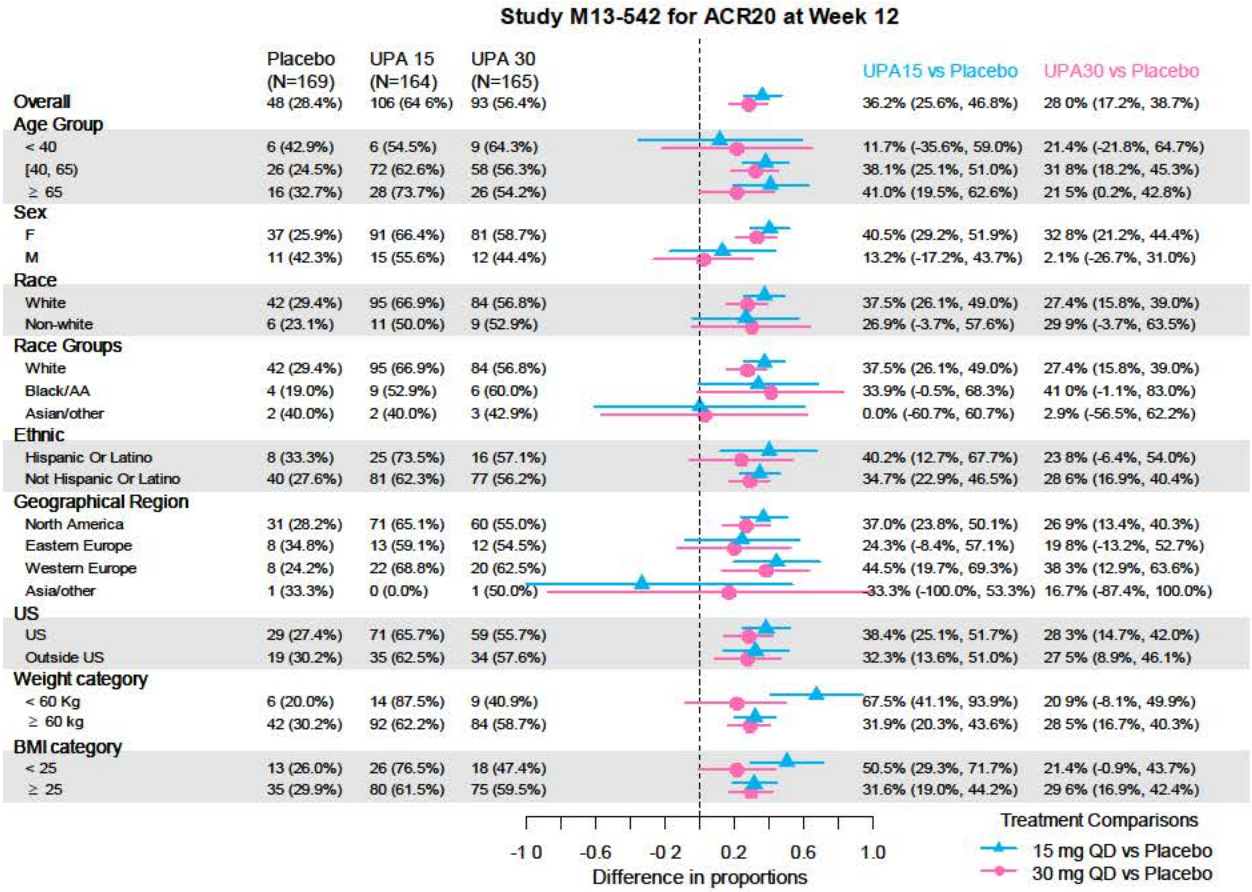
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: HAQ-DI=health assessment questionnaire-disability index; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025
[Source: Statistical Reviewer]

Figure 21 Tipping Point Sensitivity Analysis for the Mean Change from Baseline in HAQ-DI at Week 12 Comparing Upadacitinib 15 mg QD Monotherapy vs MTX Monotherapy, Study M13-545



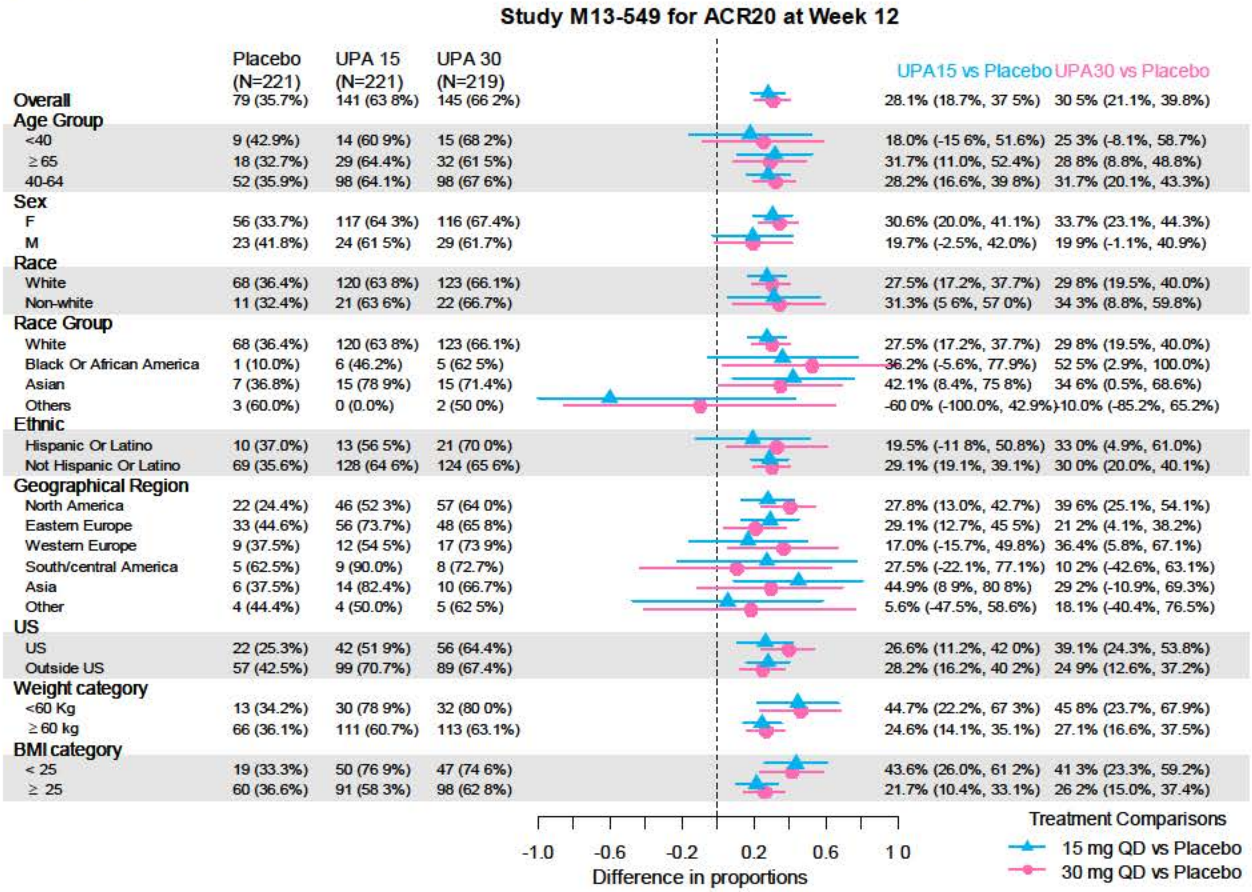
The 2-sided alpha of 0.05 was split between the two doses evaluated.
Abbreviations: HAQ-DI=health assessment questionnaire-disability index; QD=once daily; NS=non-significant based on a conservative 2-sided level alpha of 0.025
[Source: Statistical Reviewer]

Figure 22 Demographic Subgroup Analysis for the Probability of ACR20 Response at Week 12, Study M13-542



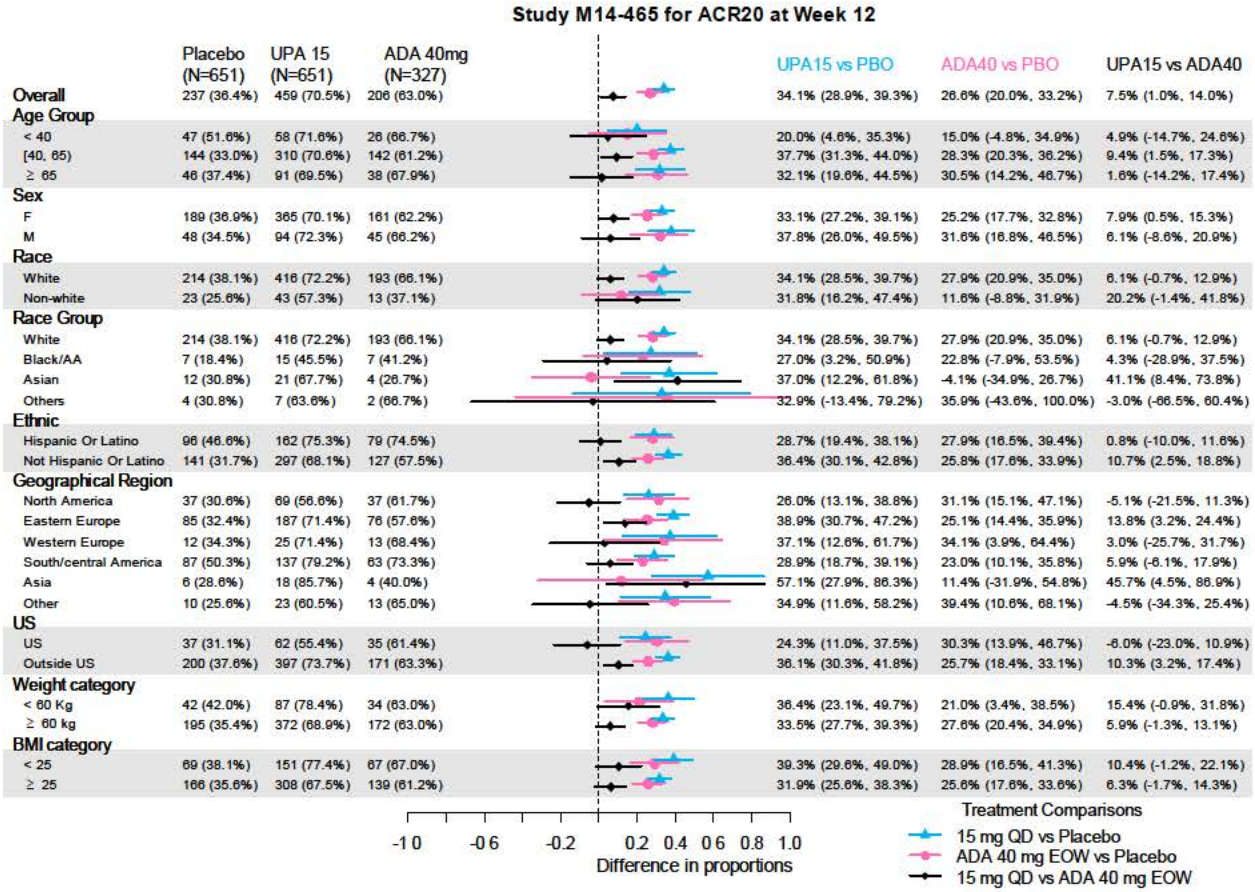
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; BMI=body mass index(kg/m^2)
[Source: Statistical Reviewer]

Figure 23 Demographic Subgroup Analysis for the Probability of ACR20 Response at Week 12, Study M13-549



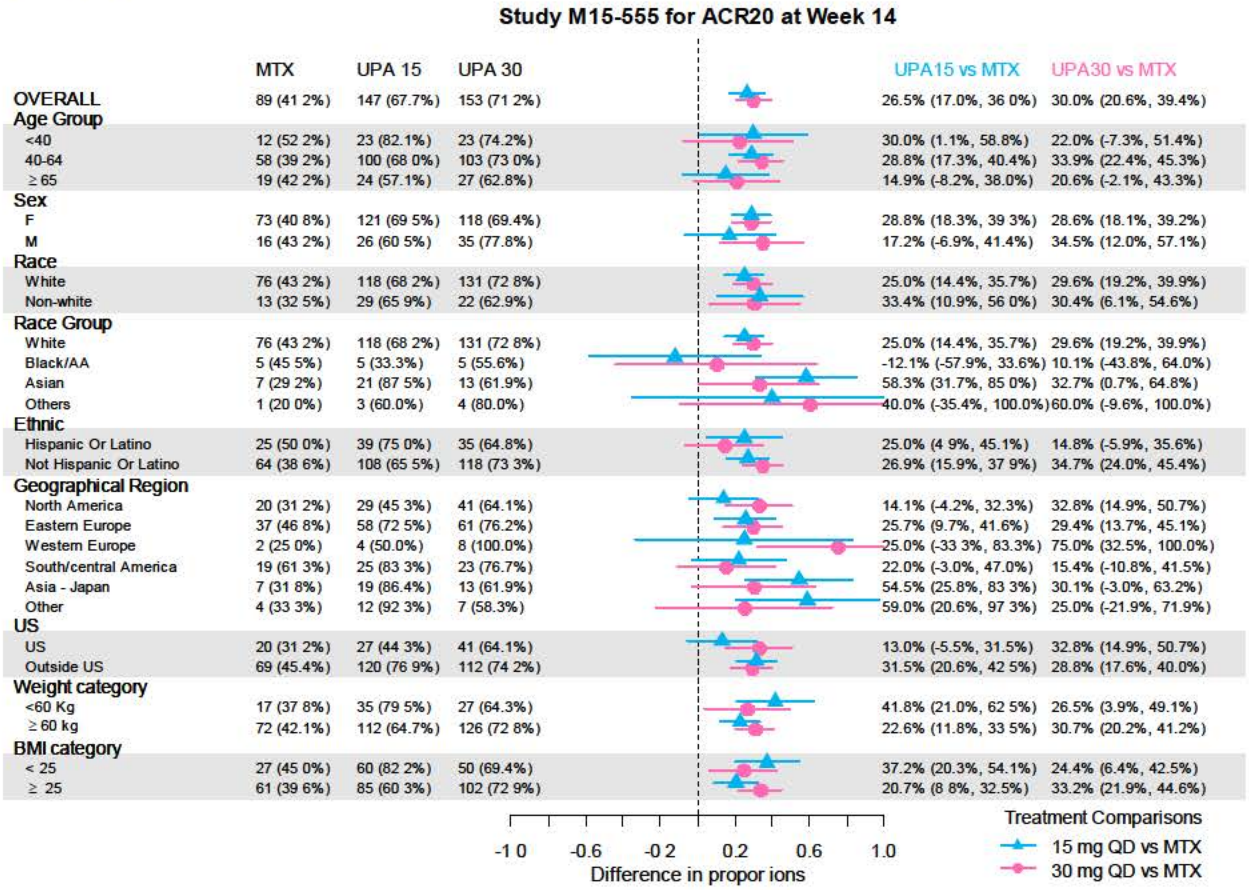
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; BMI=body mass index(kg/m^2)
[Source: Statistical Reviewer]

Figure 24 Demographic Subgroup Analysis for the Probability of ACR20 Response at Week 12, Study M14-465



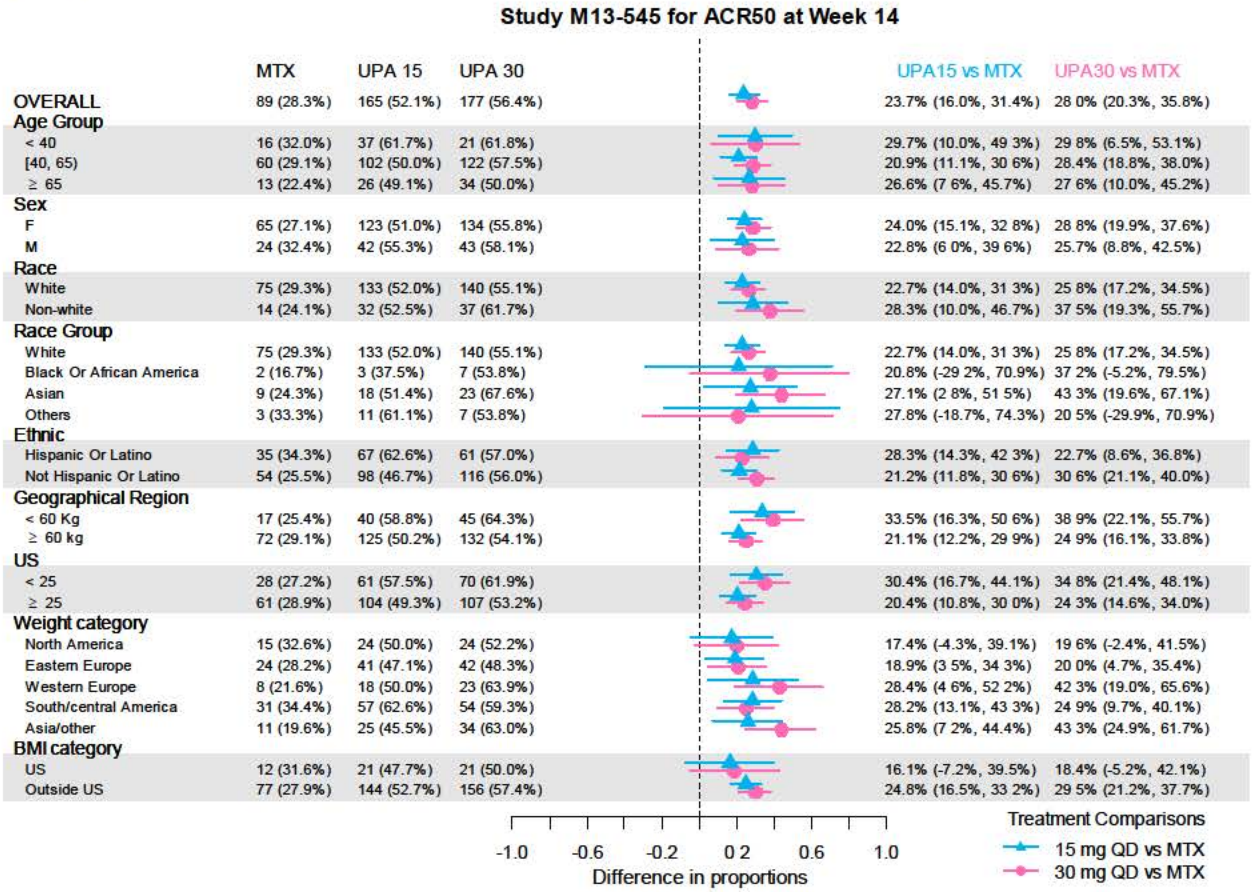
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; ADA=adalimumab; EOW=every other week; BMI=body mass index(kg/m^2)
[Source: Statistical Reviewer]

Figure 25 Demographic Subgroup Analysis for the Probability of ACR20 Response at Week 14, Study M15-555



Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; MTX=methotrexate; BMI=body mass index(kg/m^2)
[Source: Statistical Reviewer]

Figure 26 Demographic Subgroup Analysis for the Probability of ACR50 Response at Week 14, Study M13-545

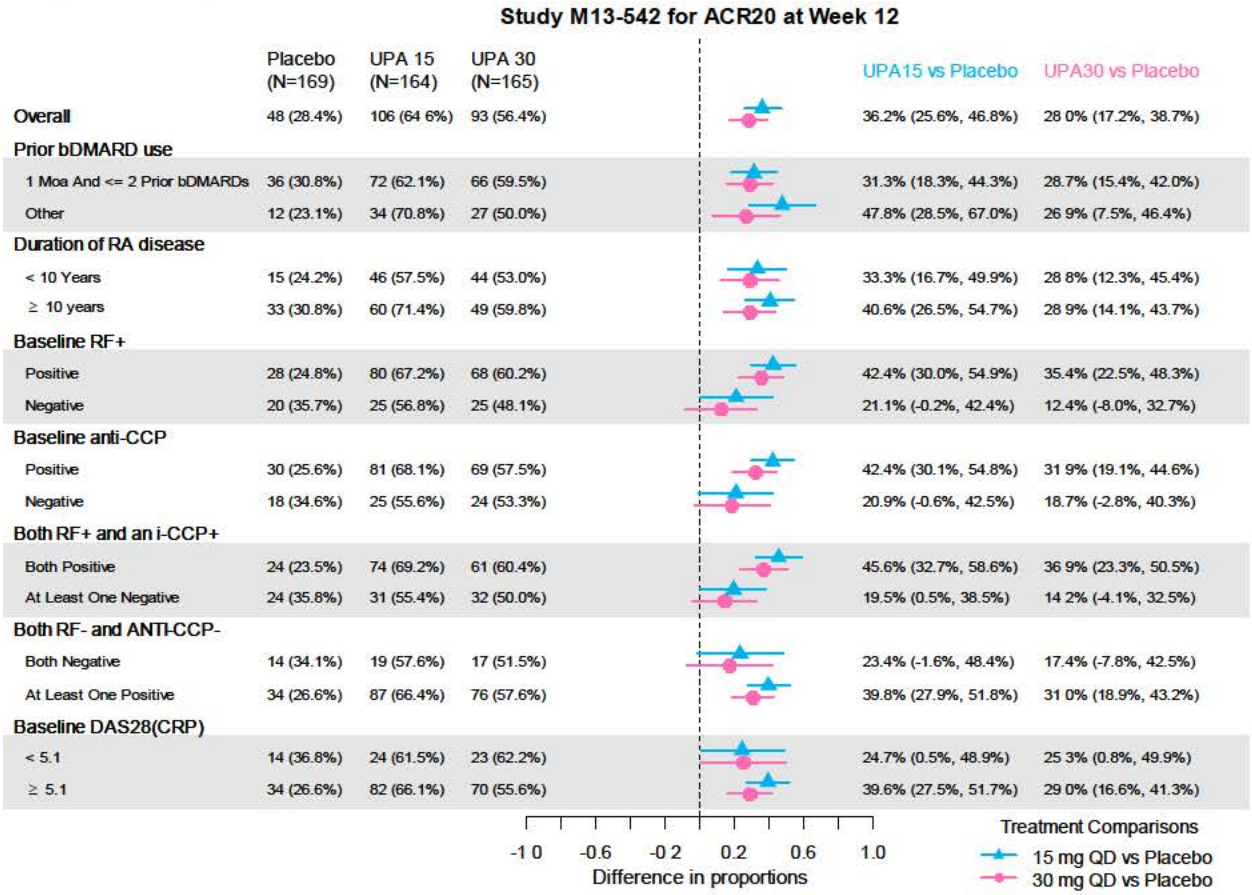


Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest for each dose comparison with MTX.

Abbreviations: ACR=American College of Rheumatology; UPA=upadacitinib; MTX=methotrexate; F=female; M=male; BMI=body mass index(kg/m^2)

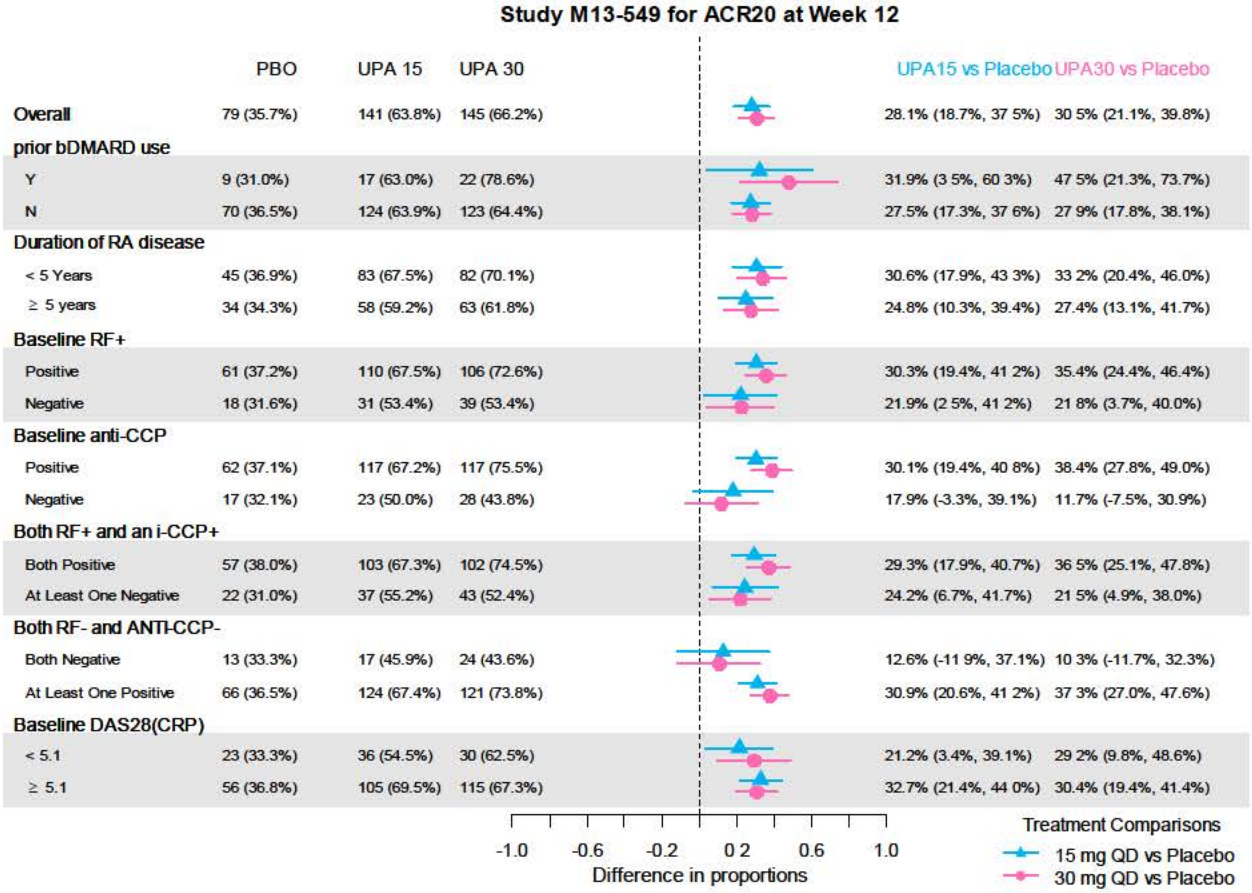
[Source: Statistical Reviewer]

Figure 27 Subgroup Analysis for Disease Characteristics based on the Probability of ACR20 Response at Week 12, Study M13-542



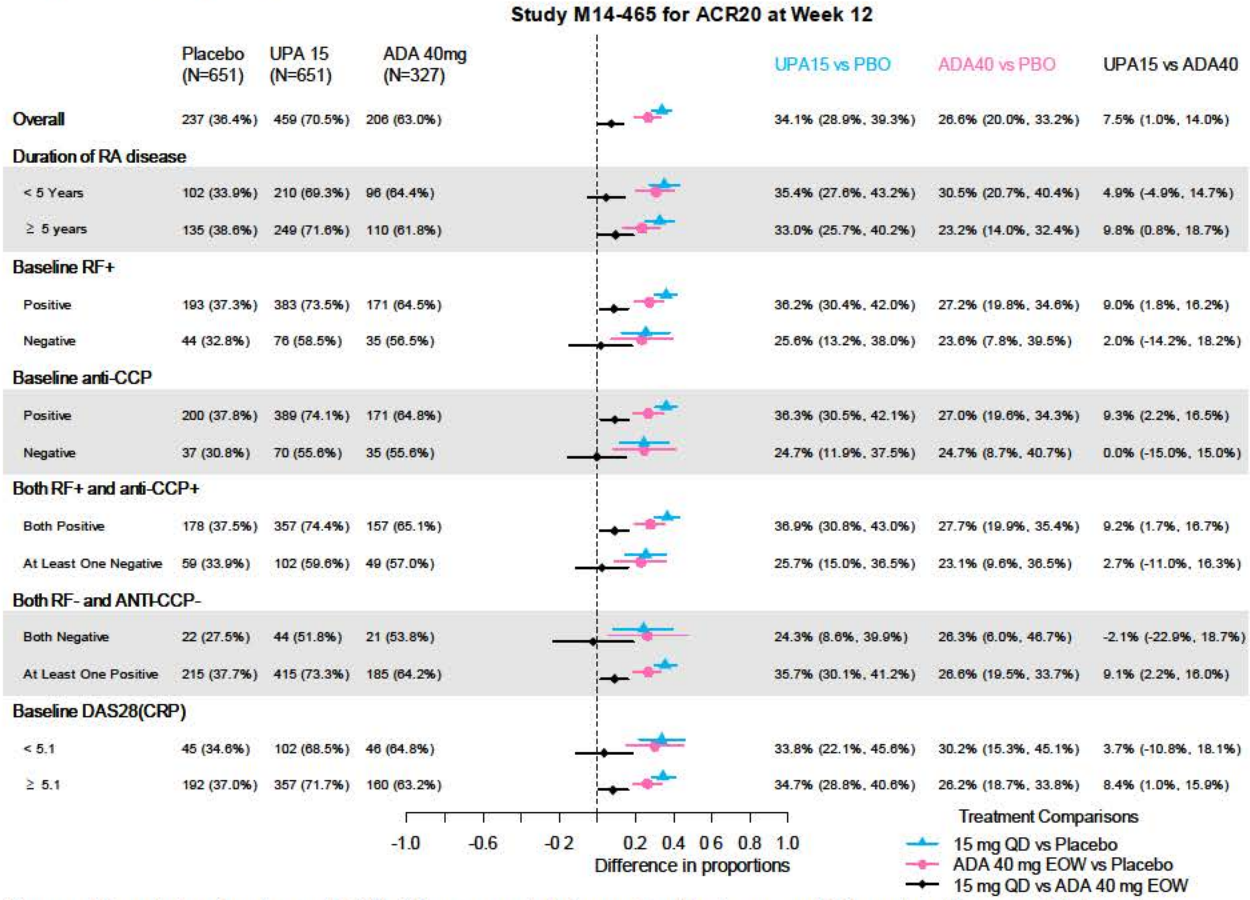
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; RF=rheumatoid factor; anti-CCP= anti-cyclic citrullinated peptide; CRP=C-reactive protein; DAS28=disease activity score based on 28 joints; RA=rheumatoid arthritis; bDMARD=biologic disease modifying anti-rheumatic drug; MOA=mechanism of action
[Source: Statistical Reviewer]

Figure 28 Subgroup Analysis for Disease Characteristics based on the Probability of ACR20 Response at Week 12, Study M13-549



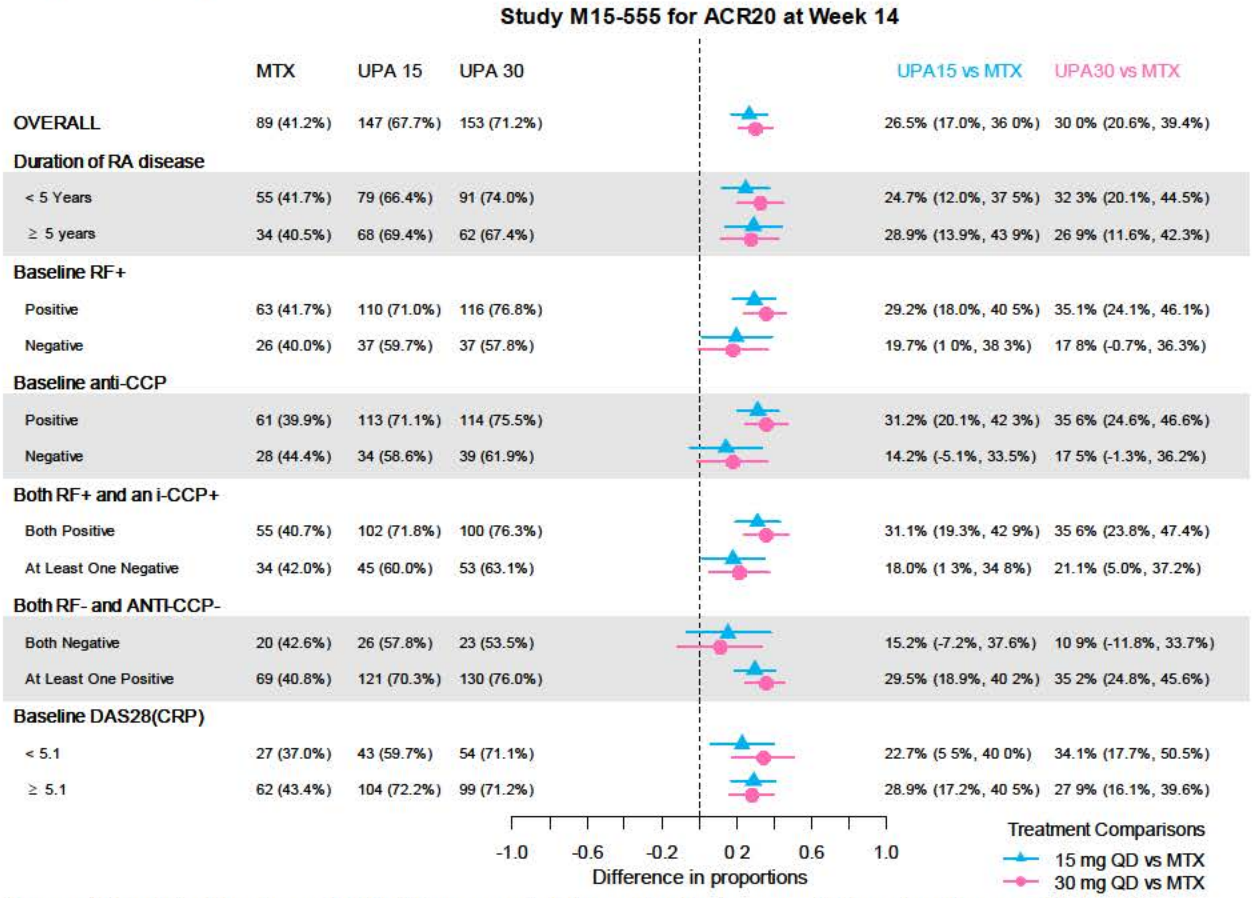
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; RF=rheumatoid factor; anti-CCP= anti-cyclic citrullinated peptide; CRP=C-reactive protein; DAS28=disease activity score based on 28 joints; RA=rheumatoid arthritis; bDMARD=biologic disease modifying anti-rheumatic drug; MOA=mechanism of action; PBO=placebo
[Source: Statistical Reviewer]

Figure 29 Subgroup Analysis for Disease Characteristics based on the Probability of ACR20 Response at Week 12, Study M14-465



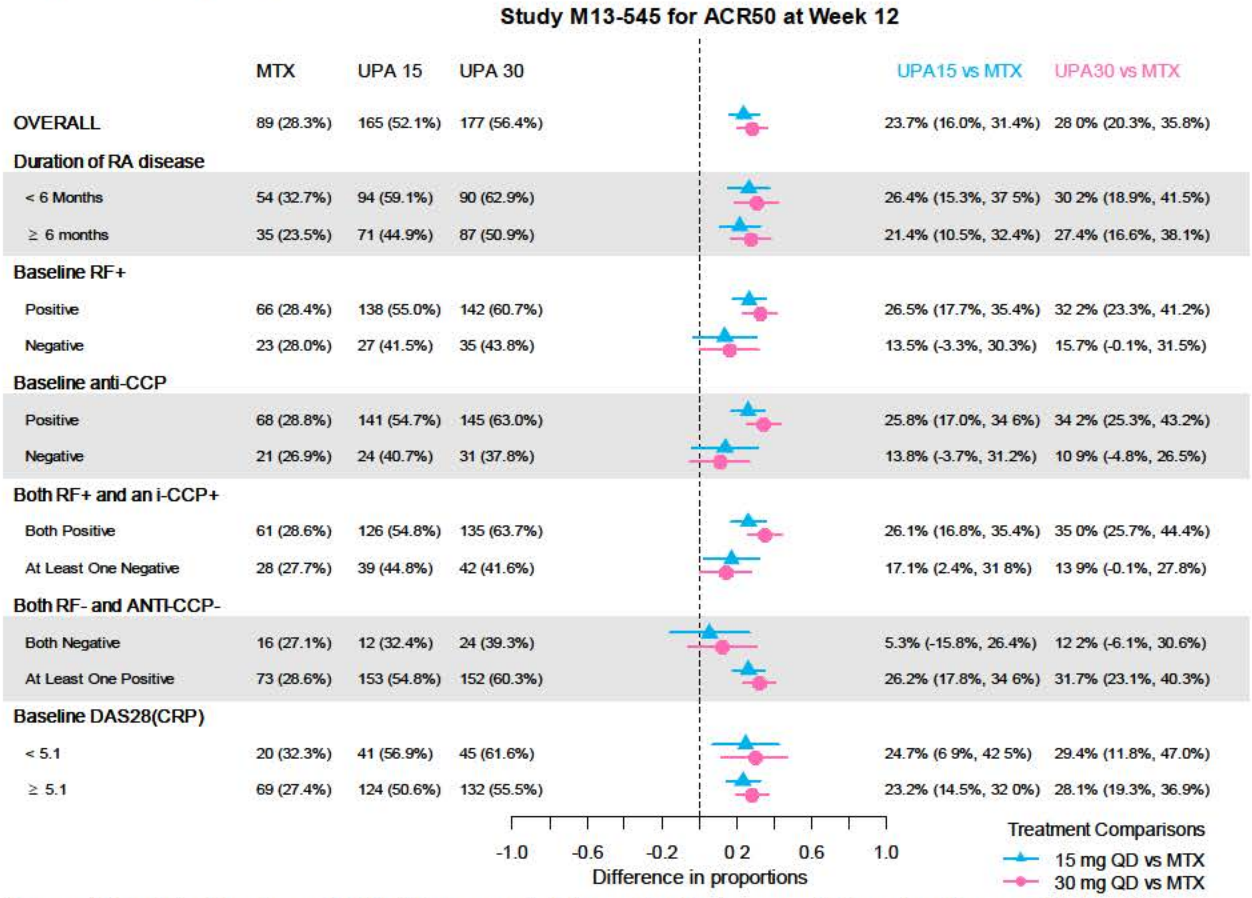
Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; ADA=adalimumab; EOW=every other week; RF=rheumatoid factor; anti-CCP= anti-cyclic citrullinated peptide; CRP=C-reactive protein; DAS28=disease activity score based on 28 joints; RA=rheumatoid arthritis; bDMARD=biologic disease modifying anti-rheumatic drug; MOA=mechanism of action; PBO=placebo
[Source: Statistical Reviewer]

Figure 30 Subgroup Analysis for Disease Characteristics based on the Probability of ACR20 Response at Week 14, Study M15-555



Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; RF=rheumatoid factor; MTX=methotrexate; anti-CCP= anti-cyclic citrullinated peptide; CRP=C-reactive protein; DAS28=disease activity score based on 28 joints; RA=rheumatoid arthritis; bDMARD=biologic disease modifying anti-rheumatic drug; MOA=mechanism of action
[Source: Statistical Reviewer]

Figure 31 Subgroup Analysis for Disease Characteristics based on the Probability of ACR50 Response at Week 12, Study M13-545



Frequentist point estimates and 95% CI are reported for each subcategory within each subgroup of interest.
Abbreviations: ACR=American College of Rheumatology; QD=once daily; UPA=upadacitinib; RF=rheumatoid factor; MTX=methotrexate; anti-CCP= anti-cyclic citrullinated peptide; CRP=C-reactive protein; DAS28=disease activity score based on 28 joints; RA=rheumatoid arthritis; bDMARD=biologic disease modifying anti-rheumatic drug; MOA=mechanism of action
[Source: Statistical Reviewer]

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

WILLIAM J KOH
08/13/2019 01:37:08 PM

PEILING YANG
08/13/2019 01:40:18 PM

MARK D ROTHMANN
08/13/2019 02:56:09 PM
I concur

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CARCINOGENICITY STUDY

IND/NDA Number:	NDA 211675
Drug Name:	Upadacitinib
Indication:	Treatment of Rheumatoid Arthritis
Applicant:	AbbVie Inc. 1 North Waukegan Road North Chicago, Illinois 60064 Test Facility for Rats and mice Studies: MPI Research 54943 North Main Street Mattawan, Michigan 49071
Documents Reviewed:	Study reports (Study TA15-032 and TD16-088)and Electronic data submitted on December 18, 2018 via NDA211675;
Biometrics Division:	Division of Biometrics -6
Statistical Reviewer:	Zhuang Miao, Ph.D.
Concurring Reviewer:	Feng Zhou, MS Karl Lin, Ph.D.
Medical Division:	Division of Pulmonary, Allergy and Rheumatology Products
Reviewing Pharmacologist:	Jones, Brett, Ph.D.
Project Manager:	Ton, Phuong Nina
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Summary

In this submission the sponsor included reports of two animal carcinogenicity studies, one in CD®[Crl:CD®(SD)] rats and one in CByB6F1-Tg(HRAS)2Jic mice. These studies were intended to assess the carcinogenic potential of A-1293543 (also known as A-1293543 hydrate C, A-1293543.0 Hemihydrate, and ABT-494), when administered orally by gavage at appropriate drug levels for 101 weeks in rats and 26 weeks in mice.

Rat Study: . Three hundred and fifty CD®[Crl:CD®(SD)] rats of each sex were randomly assigned to the three treated groups and the two vehicle control groups in equal size of 70 rats per group. The dose levels for treated groups in male rats were 4, 7.5, and 15 mg/kg/day. The dose levels for treated groups in female rats were 3, 7.5 and 20 mg/kg/day. The rats in the vehicle control group received the vehicle(0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study for the rats was designed to continue for up to 99 weeks, however in accordance with study termination criteria, all surviving male rats were sacrificed during Week 101.

Survival analysis: for female rats, the pairwise comparisons showed a statistically significant decrease in mortality in the 7.5 mg/kg/day group when compared to the combined vehicle controls.

Tumor analysis: for female rats, the pairwise comparisons showed a statistically significant increase in incidence in Ganglioneuroma in Thyroid Gland (P-value=0.0413<0.05) in 3 mg/kg/day group.

Mouse Study: One hundred CByB6F1-Tg(HRAS)2Jic mice of each sex were randomly assigned to the three treated groups and the vehicle control group in equal size of 25 mice per group. There are 15 mice of each sex in the positive control group. The dose levels for treated groups were 5, 10, and 20 mg/kg/day. The mice in the vehicle control group received the vehicle (0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study was designed to continue for up to 26 weeks for both sexes, however in accordance with study termination criteria, all surviving mice were sacrificed during Week 27. The mice in the positive control group received N-nitroso-N-methylurea via intraperitoneal injection once on Day 1 at a dose of 75 mg/kg.

Survival analysis: the pairwise comparisons showed a statistically significant increase in mortality in the positive control group when compared to the vehicle control for the female mice.

Tumor analysis: the pairwise comparisons between the vehicle control and the positive control showed statistically significant increases in incidence of Adenoma, Bronchiolar Alveolar in Lung (P-value=0.0108), Lymphoma in Lung (P-value=0.0136), Combined tumor of Adenoma and Carcinoma, Bronchiolar Alveolar in Lung (P-value=0.0423), Lmyphoma in Multicentric Neoplasm (P-value=0.0136), Papilloma, Squamous Cell in Skin (P-value=0.0431), Lymphoma in Spleen (P-value=0.0431), Papilloma, Squamous Cell in Stomach (P-value<0.001) and Lymphoma in Thymus (P-value=0.0136) in males; Lymphoma in Lung (P-value=0.0035), Lmyphoma in Multicentric Neoplasm (P-value<0.001), Papilloma, Squamous Cell in Skin (P-value<0.001), Lymphoma in Spleen (P-value=0.0122), Papilloma, Squamous Cell in Stomach (P-value<0.001), Lymphoma in Thymus (P-value<0.001) and Polyp, Glandular in Uterus with Cervis (P-value=0.0056) in females.

2. Background

In this submission the sponsor included reports of two animal carcinogenicity studies, one in CD®[Crl:CD®(SD)] rats and one in CByB6F1-Tg(HRAS)2Jic mice. These studies were intended to assess the carcinogenic potential of A-1293543 (also known as A-1293543 hydrate C, A-1293543.0 Hemihydrate, and ABT-494), when administered orally by gavage at appropriate drug levels for 101 weeks in rats and 26 weeks in mice. Results of this review have been discussed with the reviewing pharmacologist Dr. Jones. This review analyzed the SAS data sets of these studies received from the sponsor on December 18, 2018 via NDA211675.

In this review the phrase "dose response relationship" refers to the linear component of the effect of treatment, and not necessarily to a strictly increasing or decreasing mortality or tumor incidence rate as the dose increases.

3. Rat Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups and two vehicle control groups. Three hundred and fifty CD®[Crl:CD®(SD)] rats of each sex were randomly assigned to the treated and vehicle control groups in equal size of 70 rats per group. The dose levels for treated groups in male rats were 4, 7.5, and 15 mg/kg/day. The dose levels for treated groups in female rats were 3, 7.5 and 20 mg/kg/day. The rats in the two vehicle control groups received the vehicle(0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study for the rats was designed to continue for up to 99 weeks, however in accordance with study termination criteria, all surviving male rats were sacrificed during Week 101.

Table 1: Study Design in Rat Study

Protocol Group No.	Dose Levels (mg/kg/day)	Identification	Number of Animals Enrolled	
			Males	Females
1	0	Vehicle	70	
2	0	Vehicle	70	
3	4	E2006	70	
4	7.5	E2006	70	
5	15	E2006	70	
6	3	E2006		70
7	7.5	E2006		70
8	20	E2006		70

3.1. Sponsor's analyses

3.1.1. Survival analysis

Survival data were analyzed using life table techniques: log-rank test for pairwise group comparisons (control vs. treated groups) and Tarone's test for a dose-response trend. The Kaplan-Meier survival curves were calculated for each sex/group and displayed graphically over time.

Results of all pair-wise comparisons were reported at the 0.05 and 0.01 significance levels. All endpoints were analyzed using two-tailed tests.

Sponsor's findings: Sponsor's analysis showed the numbers (percents) of death were 117 (84%), 55 (79%), 54 (77%), and 55 (79%) in combined vehicle controls, 4 mg/kg/day, 7.5 mg/kg/day and 15 mg/kg/day dose groups, respectively in males and 114 (81%), 55 (79%), 50 (71%), and 53 (76%) in combined vehicle controls, 3 mg/kg/day, 7.5 mg/kg/day and 20 mg/kg/day dose groups, respectively in females.

The sponsor concluded that there was no test A-1293543 Hydrate C-related effect on survival in either sex.

3.1.2. Tumor data analysis

Tumor incidence data were analyzed using the Peto method. This method requires that each tumor, if not found through palpation, be classified as fatal or incidental. For palpable tumors, the onset rate method of analysis was used. For fatal tumors the death-rate method of analysis were used. For incidental tumors, including those found at terminal sacrifice, the prevalence analysis were used. Results from these analyses were combined to obtain a final assessment of the tumorigenicity of the test article. For sparse data, exact analysis was performed to obtain the p-values. For each tumor type, the incidence rate in the vehicle control groups and the high dose group were compared. For tumor types for which tumor data are available from all dose groups, a test for a positive dose-response trend was performed. These tests were one-sided and unadjusted p-values were reported. Evaluation criteria (p-values of significance) were applied differently for rare tumors (background rate of 1% or less) and common tumors (background rate greater than 1%). For positive trends tests, common and rare tumors were tested at 0.005 and 0.025 significance levels, respectively. For control-high pair-wise comparisons, common and rare tumors were tested at 0.01 and 0.05 significance levels, respectively.

Sponsor's findings: The sponsor's analyses did not show any A-1293543 Hydrate C-related increases in tumor incidence in either sex.

3.2. Reviewer's analyses

To verify sponsor's analyses and to perform additional analyses suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically on December 18, 2018 via NDA211675.

3.2.1. Survival analysis

The survival distributions of animals in all four groups were estimated by the Kaplan-Meier product limit method. The dose response relationship and homogeneity of survival distributions were tested for combined vehicle controls, low, medium and high dose groups using the Likelihood Ratio test and the Log-Rank test. The intercurrent mortality data are given in Tables 6 and 7 in the appendix for males and females, respectively. The Kaplan-Meier curves for survival rate are given in Figures 1 and 2 in the appendix for males and females, respectively. Results of the tests for dose response relationship and homogeneity of survivals, are given in Tables 8 and 9 in the appendix for males and females, respectively.

Reviewer's findings: This reviewer's analysis showed the numbers (percents) of death were 60 (86%), 57 (82%) 55 (79%), 54 (77%), and 55 (79%) in male rats in the vehicle control 1, vehicle control 2, 4 mg/kg/day, 7.5 mg/kg/day and 15 mg/kg/day groups respectively; and 56 (80%), 58 (83%), 55 (79%), 50 (71%), and 53 (76%) in female rats in the vehicle control 1, vehicle control 2, 3 mg/kg/day, 7.5 mg/kg/day and 20 mg/kg/day groups, respectively.

The survival analyses did not show a statistically significant dose responserelationship in mortality across combined vehicle control group and treated groups for either males or females. The pairwise comparisons did

not show any statistically significant differences in mortality between the combined vehicle control group and each of the treated groups for males.

For female rats, the pairwise comparisons showed a statistically significant decrease in mortality in the 7.5 mg/kg/day group when compared to the combined vehicle control group. The p-values for Likelihood Ratio test is 0.0256 (<0.05) and the p-values for Log-Rank test is 0.0254 (<0.05).

3.2.2. Tumor data analysis

The tumor data were analyzed for the positive dose response relationships and the positive pairwise comparison increases between each of the treated groups with control group. Both the dose response relationship tests and pairwise comparisons were performed using the Poly-K method described in the paper of Bailer and Portier (1988) and Bieler and Williams (1993). In this method an animal that lives the full study period ($w_{\max}$) or dies before the terminal sacrifice but develops the tumor type being tested gets a score of s_h =1. An animal that dies at week w_h without a tumor before the end of the study gets a score of $s_h = \left(\frac{w_h}{w_{\max}} \right)^k < 1$. The adjusted group size is defined as Σs_h . As an interpretation, an animal with score s_h =1 can be considered as a whole animal while an animal with score $s_h < 1$ can be considered as a partial animal. The adjusted group size Σs_h is equal to N (the original group size) if all animals live up to the end of the study or if each animal that dies before the terminal sacrifice develops at least one tumor, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise) tests using the Cochran-Armitage test. One critical point for Poly-k test is the choice of the appropriate value of k, which depends on the tumor incidence pattern with the increased dose. For long term 104 week standard rat and mouse studies, a value of k=3 is suggested in the literature. Hence, this reviewer used k=3 for the analysis of this data. For the calculation of p-values the exact permutation method was used. The tumor rates and the p-values for the positive dose response relationship tests and pairwise comparisons are listed in Tables 10 and 11 in the appendix for male and female rats, respectively.

Adjustment for multiple testing: For the chronic study in rats, the adjustment of multiple testing of the dose response relationship for a submission with one chronic rat study and one transgenic mouse study, the more recently revised draft (January, 2013) FDA guidance for the carcinogenicity studies suggests the use of test levels α =0.005 for common tumors and α=0.025 for rare tumors for the chronic rat study. For pairwise comparisonsfor the chronic rat study in the above type of submission with one chronic rat study and one transgenic mouse study, the same guidance document suggests the use of test levels α =0.01 for common tumors and α =0.05 for rare tumors for the chronic rat study.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 either for dose response relationship or pairwise comparisons between the vehicle control group and each of the treated

groups.

Table 2: Tumor Types with P-Values ≤ 0.05 for Dose Response Relationship or Pairwise Comparisons between the Combined Vehicle Controls and the Treated Groups-Female Rats

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-Value Trend	3 mg/kg/day Low (N=70) P-Value – Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-Value – Combined Vehicles vs. Medium	20 mg/kg/day High (N=70) P-Value – Combined Vehicles vs. High
THYROID GLAND	GANGLIONEUROMA	0/140 (70) 0.7295	3/70 (38) 0.0413	0/70 (43) NC	0/70 (38) NC

& X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;
NC = Not calculable.

Reviewer's findings: Based on the above criterion for multiple testing adjustment, we make the folloing conclusion,

1. For female rats, the pairwise comparisons between the combined vehicle controls and the treated groups showed a statistically significant increase in incidence in Ganglioneuroma in Thyroid Gland (P-value=0.0413<0.05) in 3 mg/kg/day group when compared with the combined vehicle controls group.

Reviewer's additional findings: We also performed our analyses using the vehicle control 1 and the treated groups, and using the vehicle control 2 and the treated groups. We make the following conclusion:

1. For male rats, the trend test using the vehicle control 2 and the treated groups showed a statistically significant increase in incidence in Leukemia, Granulocytic in Multicentric Neoplasm(P-value=0.0153<0.025).

4. Mouse Study

Two separate experiments were conducted, one in males and one in females. In each of these two experiments there were three treated groups, one vehicle control group, and one positive control group. One hundred CByB6F1-Tg(HRAS)2Jic mice of each sex were randomly assigned to the treated and vehicle control group in equal size of 25 mice per group. There are 15 mice of each sex in the positive control group. The dose levels for treated groups were 5, 10, and 20 mg/kg/day. The mice in the vehicle control group received the vehicle (0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study was designed to continue for up to 26 weeks for both sexes, however in accordance with study termination criteria, all surviving mice were sacrificed during Week 27. The mice in the positive control group received N-nitroso-N-methylurea via intraperitoneal injection once on Day 1 at a dose of 75 mg/kg.

Table 3: Study Design in Mouse Study

Protocol Group No.	Dose Levels (mg/kg/day)	Identification	Number of Animals Enrolled	
			Males	Females
1	0.2% hydroxypropyl methylcellulose	Vehicle	25	25
2	5	Low Dose	25	25
3	10	Middle Dose	25	25
4	20	High Dose	25	25
5	75 mg/kg MNU	Positive	15	15

4.1. Sponsor's analyses

4.1.1. Survival analysis

The sponsor used the same survival analysis methods used for the rats study in this mouse study.

Sponsor's findings: The sponsor's analysis showed 2 (8%), 0 (0%), 1 (4%), 1 (8%), and 4 (27%) mortalities in male mice, and 0 (0%), 1 (4%), 0 (0%), 1 (4%), and 6 (40%) mortalities in female mice in vehicol control, low, medium, high dose groups and positive control group, respectively.

There were no statistically significant differences in survival rates in either males or females at any dose groups.

4.1.2. Tumor data analysis

The sponsor used the same tumor data analysis methods used for the rat study in this mouse study

Sponsor's findings: The sponsor's analyses did not show any A-1293543 Hydrate C-related increases in tumor incidence in either sex.

4.2. Reviewer's analyses

To verify sponsor's analyses and to perform additional analyses suggested by the reviewing pharmacologist, this reviewer independently performed survival and tumor data analyses. Data used in this reviewer's analyses were provided by the sponsor electronically on December 18, 2018 via NDA211675. The significance level for all statistical tests was set at 0.05.

4.2.1. Survival analysis

The survival distributions of three treated groups, one vehical control group, one water control group and one positive control group were estimated using the Kaplan-Meier product limit method. The dose response relationship in survival was tested using the likelihood ratio test and the homogeneity of survival distributions was tested using the log-rank test. The Kaplan-Meier curves for survival rates are given in Figures 3 and 4 in the appendix for male and female mice, respectively. The intercurrent mortality data are given in Tables 12 and 13 in the appendix for male and female mice, respectively. Results of the tests for dose response relationship and homogeneity of survivals among the vehicle control and three treated groups are given in Tables 14 and 15 in the appendix for male and female mice, respectively.

Reviewer's findings: This reviewer's analysis showed the numbers (percents) of death 2 (8%), 0 (0%), 1 (4%), 1 (8%), and 4 (27%) mortalities in male mice, and 0 (0%), 1 (4%), 0 (0%), 1 (4%), and 6 (40%) in female rats in the vehicle control group, 5 mg/kg/day, 10 mg/kg/day, 20 mg/kg/day groups, and positive control group, respectively.

The survival analyses did not show a statistically significant dose responserelationship in mortality across vehicle control and treated groups for either males or females. The pairwise comparisons did not show statistically significant differences in mortality between the vehicle control and each of the treated groups for either males or females.

The pairwise comparisons showed a statistically significant increase in mortality in the positive control group when compared to the vehicle control for the female mice. The p-value for Likelihood Ratio test is <0.0001 and the p-value for Log-Rank test are 0.0001.

4.2.2. Tumor data analysis

The reviewer used the same tumor data analysis methods for the rat study in this mouse study.

The tumor rates and the p-values for the positive dose response relationship tests and pairwise comparisons between vehicle control and three treated groups, vehicle control and positive control are listed in Tables 16, 17, 18 and 19 in the appendix for male and female mice, respectively.

Adjustment for multiple testing: For the adjustment of multiple testing of dose response relationship for the transgenic mouse study in a submission with one chronic rat study and one transgenic mouse study, the more recently revised draft (January, 2013) FDA guidance for the carcinogenicity studies suggests the use of test levels $\alpha = 0.05$ for both common tumors and rare tumors for the mouse study. For pairwise, the same guidance document suggests the use of test levels $\alpha = 0.05$ for both common tumors and rare tumors for the mouse study.

It should be noted that the FDA guidance for multiple testing for dose response relationship is based on a publication by Lin and Rahman (1998). In this work the authors investigated the use of this rule for Peto analysis. However, in a later work Rahman and Lin (2008) showed that this rule for multiple testing for dose response relationship is also suitable for Poly-K tests.

Reviewer's findings: Following tumor types showed p-values less than or equal to 0.05 for pairwise comparisons between the vehicle control goup and the positive control group.

Table 4: Tumor Types with P-Values ≤ 0.05 for Comparisons between Vehicle Control and Positive Control-Male Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25) P-value - Trend	MNU: 75 Positive (N=15) P-value - Vehicle vs. Positive
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	0/25 (24)	4/15 (13) 0.0108
	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136
Lung	C_bronchiolar alveolar Adeno+Carcin	1/25 (24)	4/15 (13) 0.0423
MULTICENTRIC NEOPL	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136
SKIN	PAPILLOMA, SQUAMOUS CELL	0/25 (24)	3/15 (14) 0.0431
SPLEEN	LYMPHOMA	0/25 (24)	3/15 (14) 0.0431
STOMACH, NONGLANDU	PAPILLOMA, SQUAMOUS CELL	0/25 (24)	12/15 (14) <0.001
THYMUS	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136

Table 5: Tumor Types with P-Values ≤ 0.05 for Comparisons between Vehicle Control and Positive Control-Female Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25) P-value	MNU: 75 Positive (N=15) P-value - Vehicle vs. Positive
LUNG	LYMPHOMA	0/25 (25)	5/15 (14) 0.0035
MULTICENTRIC NEOPL	LYMPHOMA	0/25 (25)	6/15 (14) <0.001
SKIN	PAPILLOMA, SQUAMOUS CELL	0/25 (25)	7/15 (14) <0.001
SPLEEN	LYMPHOMA	0/25 (25)	4/15 (14) 0.0122
STOMACH, NONGLANDU	PAPILLOMA, SQUAMOUS CELL	0/25 (25)	11/15 (14) <0.001

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25) P-value	MNU: 75 Positive (N=15) P-value - Vehicle vs. Positive
THYMUS	LYMPHOMA	0/25 (25)	6/15 (14) <0.001
UTERUS WITH CERVIX	POLYP, GLANDULAR	0/25 (25)	4/15 (11) 0.0056

Reviewer's findings: Based on the criteria of adjustment for multiple testing discussed in the mouse data analysis section, we make the following conclusions

1. The pairwise comparisons between the vehicle control and the positive control showed statistically significant increases in incidence of Adenoma, Bronchiolar Alveolar in Lung (P-value=0.0108), Lymphoma in Lung (P-value=0.0136), Combined tumor of Adenoma and Carcinoma, Bronchiolar Alveolar in Lung (P-value=0.0423), Lmyphoma in Multicentric Neoplasm (P-value=0.0136), Papilloma, Squamous Cell in Skin (P-value=0.0431), Lymphoma in Spleen (P-value=0.0431), Papilloma, Squamous Cell in Stomach (P-value<0.001) and Lymphoma in Thymus (P-value=0.0136) in males.
2. The pairwise comparisons between the vehicle control and the positive control showed statistically significant increases in incidence of Lymphoma in Lung (P-value=0.0035), Lmyphoma in Multicentric Neoplasm (P-value<0.001), Papilloma, Squamous Cell in Skin (P-value<0.001), Lymphoma in Spleen (P-value=0.0122), Papilloma, Squamous Cell in Stomach (P-value<0.001), Lymphoma in Thymus (P-value<0.001) and Polyp, Glandular in Uterus with Cervis (P-value=0.0056) in females.

5. Conclusion

In this submission the sponsor included reports of two animal carcinogenicity studies, one in CD®[Crl:CD®(SD)] rats and one in CByB6F1-Tg(HRAS)2Jic mice. These studies were intended to assess the carcinogenic potential of A-1293543 (also known as A-1293543 hydrate C, A-1293543.0 Hemihydrate, and ABT-494), when administered orally by gavage at appropriate drug levels for 101 weeks in rats and 26 weeks in mice.

Rat Study: . Three hundred and fifty CD®[Crl:CD®(SD)] rats of each sex were randomly assigned to the three treated groups and the two vehicle control groups in equal size of 70 rats per group. The dose levels for treated groups in male rats were 4, 7.5, and 15 mg/kg/day. The dose levels for treated groups in female rats were 3, 7.5 and 20 mg/kg/day. The rats in the vehicle control group received the vehicle(0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study for the rats was designed to continue for up to 99 weeks, however in accordance with study termination criteria, all surviving male rats were sacrificed during Week 101.

Survival analysis: for female rats, the pairwise comparisons showed a statistically significant increase in mortality in the 7.5 mg/kg/day group when compared to the vehicle control.

Tumor analysis: for female rats, the pairwise comparisons showed a statistically significant increase in incidence in Ganglioneuroma in Thyroid Gland (P-value=0.0413<0.05) in 3 mg/kg/day group.

Mouse Study: One hundred CByB6F1-Tg(HRAS)2Jic mice of each sex were randomly assigned to the three treated groups and the vehicle control group in equal size of 25 mice per group. There are 15 mice of each sex in the positive control group. The dose levels for treated groups were 5, 10, and 20 mg/kg/day. The mice in the vehicle control group received the vehicle (0.2% hydroxypropyl methylcellulose (HPMC) in deionized water). The study was designed to continue for up to 26 weeks for both sexes, however in accordance with study termination criteria, all surviving mice were sacrificed during Week 27. The mice in the positive control group received N-nitroso-N-methylurea via intraperitoneal injection once on Day 1 at a dose of 75 mg/kg.

Survival analysis: the pairwise comparisons showed a statistically significant increase in mortality in the positive control group when compared to the vehicle control for the female mice.

Tumor analysis: the pairwise comparisons between the vehicle control and the positive control showed statistically significant increases in incidence of Adenoma, Bronchiolar Alveolar in Lung (P-value=0.0108), Lymphoma in Lung (P-value=0.0136), Combined tumor of Adenoma and Carcinoma, Bronchiolar Alveolar in Lung (P-value=0.0423), Lmyphoma in Multicentric Neoplasm (P-value=0.0136), Papilloma, Squamous Cell in Skin (P-value=0.0431), Lymphoma in Spleen (P-value=0.0431), Papilloma, Squamous Cell in Stomach (P-value<0.001) and Lymphoma in Thymus (P-value=0.0136) in males; and Lymphoma in Lung (P-value=0.0035), Lmyphoma in Multicentric Neoplasm (P-value<0.001), Papilloma, Squamous Cell in Skin (P-value<0.001), Lymphoma in Spleen (P-value=0.0122), Papilloma, Squamous Cell in Stomach (P-value<0.001), Lymphoma in Thymus (P-value<0.001) and Polyp, Glandular in Uterus with Cervis (P-value=0.0056) in females.

Zhuang Miao, Ph.D.
Mathematical Statistician

Concur:
Feng Zhou, MS

Mathematical Statistician, Biometrics-6

Karl Lin, Ph.D.
Mathematical Statistician, Team Leader, Biometrics-6

cc:
Archival NDA (b) (4)
Yi Tsong, Ph.D.
Jones, Brett, Ph.D.
Goodwin, Andrew, Ph.D.

6. Appendix

Table 6: Intercurrent Mortality Rate -Male Rats

	Vehicle Control 1 0 mg|kg|day		Vehicle Control 2 0 mg|kg|day		4 mg|kg|day		7.5 mg|kg|day		15 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 52	6	8.57	11	15.71	6	8.57	5	7.14	9	12.86
53 - 78	28	48.57	23	48.57	26	45.71	28	47.14	21	42.86
79 - 91	15	70.00	13	67.14	13	64.29	4	52.86	15	64.29
92 - 100	11	85.71	10	81.43	10	78.57	17	77.14	10	78.57
Ter. Sac.	10	14.29	13	18.57	15	21.43	16	22.86	15	21.43

Cum. %: Cumulative percentage except for Ter. Sac.

Table 7: Intercurrent Mortality Rate -Female Rats

	Vehicle Control 1 0 mg|kg|day		Vehicle Control 2 0 mg|kg|day		3 mg|kg|day		7.5 mg|kg|day		20 mg|kg|day	
Week	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**	**No. of Death**	**Cum. %**
0 - 52	7	10.00	5	7.14	5	7.14	3	4.29	5	7.14
53 - 78	32	55.71	30	50.00	28	47.14	20	32.86	23	40.00
79 - 91	11	71.43	15	71.43	11	62.86	23	65.71	20	68.57
92 - 100	6	80.00	8	82.86	11	78.57	4	71.43	5	75.71
Ter. Sac.	14	20.00	12	17.14	15	21.43	20	28.57	17	24.29

Cum. %: Cumulative percentage except for Ter. Sac.

Table 8: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Male Rats

Test	Statistic	P_Value Dose Response	P_Value Combined Vehicles vs. Low	P_Value Combined Vehicles vs. Medium	P_Value Combined Vehicles vs. High
Dose-Response	Likelihood Ratio	0.4398	0.5253	0.2487	0.4881
Homogeneity	Log-Rank	0.6268	0.5211	0.2464	0.4848

Table 9: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control -Female Rats

Test	Statistic	P_Value Dose Response	P_Value Combined Vehicles vs. Low	P_Value Combined Vehicles vs. Medium	P_Value Combined Vehicles vs. High
Dose-Response	Likelihood Ratio	0.2368	0.5202	0.0256	0.2777
Homogeneity	Log-Rank	0.1682	0.5164	0.0254	0.2751

Table 10: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons between the Combined Vehicle Controls and the Treated Groups-Male Rats

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	4 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	15 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
ADRENAL GLANDS	ADENOMA, CORTICAL	3/140 (74) 0.7759	1/70 (39) 0.8213	0/70 (40) 1.0000	1/70 (39) 0.8213
	PHEOCHROMOCYTOMA	17/140 (78) 0.9540	6/70 (40) 0.8714	9/70 (42) 0.6047	3/70 (38) 0.9874
BONE, MANDIBLE	OSTEOSARCOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
BRAIN	CARCINOMA, PARS DISTALIS	1/140 (73) 0.6951	0/70 (39) 1.0000	1/70 (40) 0.5847	0/70 (38) 1.0000
	GLIOMA	1/140 (74) 0.6150	3/70 (40) 0.1233	2/70 (40) 0.2811	0/70 (38) 1.0000
	GRANULAR CELL TUMOR	2/140 (74) 0.2411	0/70 (39) 1.0000	1/70 (40) 0.7304	2/70 (39) 0.4287
	MENINGIOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
	OLIGODENDROGLIOMA	0/140 (73) 0.2042	0/70 (39) NC	0/70 (40) NC	1/70 (39) 0.3482
	SCHWANNOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
EYES	CARCINOMA, SQUAMOUS CELL	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
HEART	MESOTHELIOMA, ATRIOCAVAL	1/140 (73) 0.6951	0/70 (39) 1.0000	1/70 (40) 0.5847	0/70 (38) 1.0000
	SCHWANNOMA	1/140 (73) 0.4105	0/70 (39) 1.0000	0/70 (40) 1.0000	1/70 (39) 0.5772
KIDNEYS	ADENOMA, RENAL TUBULE, AMPHOPHIL	0/140 (73) 0.0392	0/70 (39) NC	0/70 (40) NC	2/70 (38) 0.1152
	LIPOSARCOMA	2/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
	PAPILLOMA, TRANSITIONAL CELL	0/140 (73) 0.2042	0/70 (39) NC	0/70 (40) NC	1/70 (39) 0.3482
LIVER	ADENOMA, HEPATOCELLULAR	1/140 (74) 0.8527	1/70 (40) 0.5807	0/70 (40) 1.0000	0/70 (38) 1.0000
	CHOLANGIOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	4 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	15 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	HEMANGIOSARCOMA	0/140 (73) 0.2042	0/70 (39) NC	0/70 (40) NC	1/70 (39) 0.3482
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	2/140 (74) 0.9433	1/70 (39) 0.7231	0/70 (40) 1.0000	0/70 (38) 1.0000
	CARCINOMA, C-CELL	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
	HEMANGIOSARCOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
LYMPH NODE, MANDIB	CARCINOMA, C-CELL	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
LYMPH NODE, MESENT	HEMANGIOMA	0/140 (73) 0.2000	0/70 (39) NC	0/70 (40) NC	1/70 (38) 0.3423
	HEMANGIOSARCOMA	2/140 (75) 0.5481	1/70 (39) 0.7192	1/70 (40) 0.7265	1/70 (39) 0.7192
MAMMARY GLAND	ADENOCARCINOMA	1/140 (74) 0.6921	0/70 (39) 1.0000	1/70 (40) 0.5807	0/70 (38) 1.0000
	ADENOMA	0/140 (73) 0.2042	0/70 (39) NC	0/70 (40) NC	1/70 (39) 0.3482
	FIBROADENOMA	0/140 (73) 0.6158	1/70 (39) 0.3482	0/70 (40) NC	0/70 (38) NC
MULTICENTRIC NEOPL	LEUKEMIA, GRANULOCYTIC	1/140 (74) 0.0375	0/70 (39) 1.0000	0/70 (40) 1.0000	3/70 (39) 0.1178
	LEUKEMIA, LARGE GRANULAR LYMP	0/140 (73) 0.1238	0/70 (39) NC	1/70 (40) 0.3540	1/70 (38) 0.3423
	LYMPHOMA	4/140 (76) 0.7884	0/70 (39) 1.0000	1/70 (40) 0.8848	1/70 (39) 0.8796
	SARCOMA, HISTIOCYTIC	2/140 (74) 0.1176	1/70 (39) 0.7231	2/70 (41) 0.4491	3/70 (39) 0.2233
PANCREAS	ADENOMA, ACINAR CELL	0/140 (73) 0.2000	0/70 (39) NC	0/70 (40) NC	1/70 (38) 0.3423
	ADENOMA, ISLET CELL	9/140 (75) 0.5026	8/70 (41) 0.2048	5/70 (41) 0.5965	5/70 (39) 0.5592
	CARCINOMA, ACINAR CELL	1/140 (74) 0.6921	0/70 (39) 1.0000	1/70 (40) 0.5807	0/70 (38) 1.0000
	CARCINOMA, ISLET CELL	4/140 (74) 0.8102	1/70 (39) 0.8852	1/70 (40) 0.8903	1/70 (39) 0.8852
	C ACINAR CELL ADENO+CARCINOMA	4/140 (74) 0.5744	1/70 (39) 0.8852	1/70 (40) 0.8903	2/70 (39) 0.6780

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	4 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	15 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	C_Adenoma	9/140 (75) 0.3670	8/70 (41) 0.2048	5/70 (41) 0.5965	6/70 (39) 0.4066
	C_CARCINOMA	5/140 (75) 0.8340	1/70 (39) 0.9245	2/70 (40) 0.7709	1/70 (39) 0.9245
	C_ISLET CELL ADENO+CARCINOMA	13/140 (76) 0.6624	9/70 (41) 0.3429	6/70 (41) 0.7241	6/70 (40) 0.7055
PARATHYROID GLANDS	ADENOMA	1/140 (73) 0.7602	2/70 (39) 0.2774	0/70 (40) 1.0000	0/70 (38) 1.0000
PITUITARY GLAND	ADENOMA, PARS DISTALIS	94/140 (114) 0.7520	50/70 (60) 0.5318	53/70 (61) 0.2969	44/70 (57) 0.8479
	ADENOMA, PARS INTERMEDIA	2/140 (75) 0.2802	0/70 (39) 1.0000	0/70 (40) 1.0000	2/70 (40) 0.4335
	CARCINOMA, PARS DISTALIS	0/140 (73) 0.4105	0/70 (39) NC	1/70 (40) 0.3540	0/70 (38) NC
	C PARS DISTALIS ADENO+CARCINOMA	94/140 (114) 0.7349	50/70 (60) 0.5318	54/70 (61) 0.2022	44/70 (57) 0.8479
SKIN	ADENOMA, BASAL CELL	2/140 (74) 0.8833	2/70 (39) 0.4287	0/70 (40) 1.0000	0/70 (38) 1.0000
	ADENOMA, SEBACEOUS CELL	0/140 (73) 0.4043	0/70 (39) NC	2/70 (40) 0.1233	0/70 (38) NC
	CARCINOMA, SEBACEOUS CELL	0/140 (73) 0.6178	1/70 (40) 0.3540	0/70 (40) NC	0/70 (38) NC
	CARCINOMA, SQUAMOUS CELL	1/140 (74) 0.6921	0/70 (39) 1.0000	1/70 (40) 0.5807	0/70 (38) 1.0000
	C_KERATOACA+CARCINO+PAPILLO	8/140 (76) 0.9006	5/70 (41) 0.5037	5/70 (40) 0.4855	1/70 (39) 0.9797
	C SEBACEOUS CELL ADENO+CARCINOMA	0/140 (73) 0.4366	1/70 (40) 0.3540	2/70 (40) 0.1233	0/70 (38) NC
	HAIR FOLLICLE TUMOR	4/140 (75) 0.9148	2/70 (39) 0.6721	2/70 (41) 0.6937	0/70 (38) 1.0000
	KERATOACANTHOMA	5/140 (75) 0.8345	4/70 (40) 0.3831	2/70 (40) 0.7709	1/70 (39) 0.9245
	PAPILLOMA, FIBROUS	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
	PAPILLOMA, SQUAMOUS CELL	3/140 (75) 0.9208	1/70 (39) 0.8179	1/70 (40) 0.8242	0/70 (38) 1.0000
SKIN SUBCUTIS	C_LIPOMA+LIPOSAR	5/140 (75) 0.4376	2/70 (39) 0.7609	2/70 (41) 0.7804	3/70 (38) 0.5430

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	4 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	15 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	C_SARCO+FIBROSAR+LIPOSAR	7/140 (77) 0.7795	2/70 (40) 0.8786	3/70 (42) 0.7557	2/70 (39) 0.8712
SKIN, SUBCUTIS	CARCINOMA, C-CELL	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
	FIBROMA	9/140 (77) 0.7223	4/70 (40) 0.7140	5/70 (41) 0.5768	3/70 (39) 0.8386
	FIBROSARCOMA	3/140 (75) 0.6387	2/70 (40) 0.5708	2/70 (41) 0.5817	1/70 (39) 0.8179
	HEMANGIOSARCOMA	3/140 (74) 0.9330	0/70 (39) 1.0000	1/70 (41) 0.8336	0/70 (38) 1.0000
	LIPOMA	2/140 (74) 0.2724	2/70 (39) 0.4287	2/70 (41) 0.4491	2/70 (38) 0.4181
	LIPOSARCOMA	3/140 (74) 0.7550	0/70 (39) 1.0000	0/70 (40) 1.0000	1/70 (38) 0.8147
	OSTEOSARCOMA	0/140 (73) 0.6158	1/70 (39) 0.3482	0/70 (40) NC	0/70 (38) NC
	SARCOMA, UNDIFFERENTIATED	1/140 (74) 0.6953	0/70 (39) 1.0000	1/70 (41) 0.5879	0/70 (38) 1.0000
	SCHWANNOMA	1/140 (74) 0.7566	2/70 (39) 0.2732	0/70 (40) 1.0000	0/70 (38) 1.0000
SPLEEN	HEMANGIOSARCOMA	1/140 (74) 0.1317	0/70 (39) 1.0000	0/70 (40) 1.0000	2/70 (39) 0.2732
	LEIOMYOSARCOMA	1/140 (74) 0.3169	0/70 (39) 1.0000	1/70 (40) 0.5807	1/70 (39) 0.5732
	LIPOSARCOMA	1/140 (74) 1.0000	0/70 (39) 1.0000	0/70 (40) 1.0000	0/70 (38) 1.0000
TESTES	ADENOMA, LEYDIG CELL	1/140 (74) 0.1382	1/70 (39) 0.5732	2/70 (41) 0.2889	2/70 (38) 0.2653
THYROID GLAND	ADENOMA, C-CELL	12/140 (76) 0.2473	6/70 (41) 0.6615	8/70 (43) 0.4386	8/70 (39) 0.3491
	ADENOMA, FOLLICULAR CELL	3/140 (74) 0.0784	1/70 (39) 0.8213	4/70 (41) 0.2041	4/70 (39) 0.1850
	CARCINOMA, C-CELL	3/140 (75) 0.7480	2/70 (40) 0.5708	0/70 (40) 1.0000	1/70 (38) 0.8112
	CARCINOMA, FOLLICULAR CELL	2/140 (74) 0.2915	1/70 (39) 0.7231	1/70 (41) 0.7375	2/70 (39) 0.4287
	GANGLIONEUROMA	2/140 (74) 0.6276	1/70 (39) 0.7231	0/70 (40) 1.0000	1/70 (38) 0.7156

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	4 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	15 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
Whold Body	C_hemangiosar+heman	6/140 (76) 0.3406	1/70 (39) 0.9501	2/70 (41) 0.8409	4/70 (39) 0.4572
ZYMBAL`S GLAND	CARCINOMA, ZYMBALS GLAND	0/140 (73) 0.2065	2/70 (40) 0.1233	0/70 (40) NC	1/70 (38) 0.3423

Table 11: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons between the Combined Vehicle Controls and the Treated Groups-Female Rats

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	3 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	20 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
ADRENAL GLANDS	ADENOMA, CORTICAL	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	PHEOCHROMOCYTOMA	5/140 (71) 0.8641	1/70 (38) 0.9293	0/70 (43) 1.0000	1/70 (38) 0.9293
BRAIN	CARCINOMA, PARS DISTALIS	6/140 (73) 0.9279	5/70 (39) 0.3205	8/70 (46) 0.1122	0/70 (38) 1.0000
	GLIOMA	1/140 (71) 0.6748	0/70 (37) 1.0000	1/70 (43) 0.6142	0/70 (38) 1.0000
	GRANULAR CELL TUMOR	0/140 (70) 0.4309	0/70 (37) NC	1/70 (43) 0.3805	0/70 (38) NC
	MIXED GLIOMA	2/140 (70) 0.8250	0/70 (37) 1.0000	1/70 (43) 0.7662	0/70 (38) 1.0000
	OLIGODENDROGLIOMA	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
CAVITY, ABDOMINAL	LIPOMA	0/140 (70) 0.2021	0/70 (37) NC	0/70 (43) NC	1/70 (38) 0.3519
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	SCHWANNOMA	0/140 (70) 0.2021	0/70 (37) NC	0/70 (43) NC	1/70 (38) 0.3519
CAVITY, THORACIC	ADENOCARCINOMA	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
HEART	ADENOCARCINOMA	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
KIDNEYS	ADENOMA, RENAL TUBULE, AMPHOPHIL	1/140 (71) 0.3625	0/70 (37) 1.0000	0/70 (43) 1.0000	1/70 (38) 0.5778
	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	3 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	20 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	LIPOSARCOMA	1/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
LIVER	ADENOMA, HEPATOCELLULAR	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
LUNG	ADENOCARCINOMA	2/140 (72) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	CARCINOMA, BRONCHIOLAR ALVEOL	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	CARCINOMA, PARS DISTALIS	0/140 (70) 0.4309	0/70 (37) NC	1/70 (43) 0.3805	0/70 (38) NC
	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
LYMPH NODE, MESENT	HEMANGIOSARCOMA	1/140 (71) 0.8616	1/70 (38) 0.5778	0/70 (43) 1.0000	0/70 (38) 1.0000
MAMMARY GLAND	ADENOCARCINOMA	38/140 (86) 0.9726	30/70 (50) 0.0545	14/70 (47) 0.9661	14/70 (43) 0.9287
	ADENOLIPOMA	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	ADENOMA	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	FIBROADENOMA	70/140 (102) 1.0000	23/70 (45) 0.9861	15/70 (47) 1.0000	13/70 (43) 1.0000
MESENTERY/PERITONE	MESOTHELIOMA	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
MULTICENTRIC NEOPL	LYMPHOMA	4/140 (74) 0.7109	0/70 (37) 1.0000	1/70 (43) 0.9039	1/70 (39) 0.8852
	SARCOMA, HISTIOCYTIC	1/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
OVARIES	SEX-CORD/STROMAL TUMOR	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
PANCREAS	ADENOMA, ISLET CELL	3/140 (70) 0.9238	1/70 (38) 0.8289	1/70 (43) 0.8576	0/70 (38) 1.0000

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	3 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	20 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	CARCINOMA, ISLET CELL	2/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	C_Adenoma	39/140 (87) 0.9762	30/70 (50) 0.0625	14/70 (47) 0.9712	14/70 (43) 0.9378
	C_ISLET CELL ADENO+CARCINOMA	5/140 (71) 0.9809	1/70 (38) 0.9293	1/70 (43) 0.9463	0/70 (38) 1.0000
PARATHYROID GLANDS	ADENOMA	2/140 (71) 0.9492	1/70 (38) 0.7277	0/70 (43) 1.0000	0/70 (38) 1.0000
PITUITARY GLAND	ADENOMA, PARS DISTALIS	111/140 (122) 0.0765	56/70 (63) 0.7665	56/70 (64) 0.8415	62/70 (64) 0.1130
	ADENOMA, PARS INTERMEDIA	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	CARCINOMA, PARS DISTALIS	6/140 (73) 0.9170	5/70 (39) 0.3205	9/70 (47) 0.0703	0/70 (38) 1.0000
	C PARS DISTALIS ADENO+CARCINOMA	117/140 (124) 0.2130	61/70 (65) 0.6881	65/70 (68) 0.5012	62/70 (64) 0.3552
SKIN	ADENOMA, BASAL CELL	1/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	CARCINOMA, BASAL CELL	0/140 (70) 0.4309	0/70 (37) NC	1/70 (43) 0.3805	0/70 (38) NC
	CARCINOMA, SEBACEOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	C_BASAL CELL ADENO+CARCINOMA	1/140 (70) 0.6774	0/70 (37) 1.0000	1/70 (43) 0.6184	0/70 (38) 1.0000
	C_KERATOA+CARCINO	0/140 (70) 0.1361	0/70 (37) NC	1/70 (43) 0.3805	1/70 (39) 0.3578
	HAIR FOLLICLE TUMOR	1/140 (71) 0.8653	2/70 (38) 0.2779	0/70 (43) 1.0000	0/70 (38) 1.0000
	KERATOACANTHOMA	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
SKIN SUBCUTIS	C_SARCO+FIBROSAR	3/140 (72) 0.7803	3/70 (38) 0.3412	0/70 (43) 1.0000	1/70 (39) 0.8282
SKIN, SUBCUTIS	FIBROMA	2/140 (71) 0.2526	1/70 (38) 0.7277	0/70 (43) 1.0000	2/70 (38) 0.4349
	FIBROSARCOMA	3/140 (72) 0.7425	2/70 (38) 0.5663	0/70 (43) 1.0000	1/70 (39) 0.8282
	LIPOMA	1/140 (71) 0.1019	0/70 (37) 1.0000	1/70 (43) 0.6142	2/70 (38) 0.2779

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	3 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	20 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	PAPILLOMA, FIBROUS	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
SMALL INTESTINE, I	LEIOMYOMA	0/140 (70) 0.4309	0/70 (37) NC	1/70 (43) 0.3805	0/70 (38) NC
SMALL INTESTINE, J	ADENOCARCINOMA	1/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
SPINAL CORD, THORA	GLIOMA	1/140 (71) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
SPLEEN	HEMANGIOSARCOMA	1/140 (71) 0.8616	1/70 (38) 0.5778	0/70 (43) 1.0000	0/70 (38) 1.0000
	LEIOMYOSARCOMA	1/140 (71) 0.3625	0/70 (37) 1.0000	0/70 (43) 1.0000	1/70 (38) 0.5778
THYMUS	THYMOMA	1/140 (71) 0.3648	1/70 (38) 0.5778	1/70 (43) 0.6142	1/70 (38) 0.5778
THYROID GLAND	ADENOMA, C-CELL	13/140 (74) 0.7488	5/70 (39) 0.8218	13/70 (46) 0.1246	4/70 (39) 0.9082
	ADENOMA, FOLLICULAR CELL	1/140 (71) 0.8616	1/70 (38) 0.5778	0/70 (43) 1.0000	0/70 (38) 1.0000
	CARCINOMA, C-CELL	2/140 (70) 0.6393	2/70 (38) 0.4408	0/70 (43) 1.0000	1/70 (39) 0.7393
	CARCINOMA, FOLLICULAR CELL	2/140 (70) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	C_C CELL ADENO+CARCINO	14/140 (74) 0.7331	7/70 (39) 0.6423	13/70 (46) 0.1667	5/70 (39) 0.8627
	C FOLLICULAR CELL ADENO+CARCINO	3/140 (71) 0.9815	1/70 (38) 0.8253	0/70 (43) 1.0000	0/70 (38) 1.0000
	GANGLIONEUROMA	0/140 (70) 0.7295	3/70 (38) 0.0413	0/70 (43) NC	0/70 (38) NC
UTERUS WITH CERVIS	C_LEIOMYOMA+LEIOMYOSARCOMA	0/140 (70) 0.2128	1/70 (38) 0.3519	0/70 (43) NC	1/70 (38) 0.3519
	C_POLYP+SARCOMA	6/140 (72) 0.6580	5/70 (39) 0.3290	1/70 (43) 0.9664	3/70 (39) 0.6748
UTERUS WITH CERVIX	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.1361	0/70 (37) NC	1/70 (43) 0.3805	1/70 (39) 0.3578
	GRANULAR CELL TUMOR	3/140 (72) 0.9061	0/70 (37) 1.0000	1/70 (43) 0.8512	0/70 (38) 1.0000

Organ Name	Tumor Name	0 mg/kg/day Combined Vehicles (N=140) P-value - Trend	3 mg/kg/day Low (N=70) P-value - Combined Vehicles vs. Low	7.5 mg/kg/day Med (N=70) P-value - Combined Vehicles vs. Med	20 mg/kg/day High (N=70) P-value - Combined Vehicles vs. High
	HEMANGIOSARCOMA	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	LEIOMYOMA	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	LEIOMYOSARCOMA	0/140 (70) 0.2021	0/70 (37) NC	0/70 (43) NC	1/70 (38) 0.3519
	POLYP, ENDOMETRIAL STROMAL	6/140 (72) 0.6149	4/70 (39) 0.4919	1/70 (43) 0.9664	3/70 (39) 0.6748
	SARCOMA, UNDIFFERENTIATED	0/140 (70) 0.6296	1/70 (38) 0.3519	0/70 (43) NC	0/70 (38) NC
	SCHWANNOMA	1/140 (70) 0.1074	0/70 (37) 1.0000	1/70 (43) 0.6184	2/70 (39) 0.2906
VAGINA	CARCINOMA, SQUAMOUS CELL	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
	GRANULAR CELL TUMOR	5/140 (73) 1.0000	0/70 (37) 1.0000	0/70 (43) 1.0000	0/70 (38) 1.0000
	SCHWANNOMA	0/140 (70) 0.2063	0/70 (37) NC	0/70 (43) NC	1/70 (39) 0.3578
Whold Body	C_hemangiosar+heman	2/140 (72) 0.9309	3/70 (39) 0.2330	0/70 (43) 1.0000	0/70 (38) 1.0000

Table 12: Intercurrent Mortality Rate -Male Mice

	Vehicle 0 mg|kg|day		Low 5 mg|kg|day		Middle 10 mg|kg|day		High 20 mg|kg|day		Positive MNU: 75	
Week	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 13	.	.	.	.	1	4.00	.	.	.	.
14 - 26	2	8.00	.	.	.	.	1	4.00	4	26.67
Ter. Sac.	23	92.00	25	100.00	24	96.00	24	96.00	11	73.33

Cum. %: Cumulative percentage except for Ter. Sac.

Table 13: Intercurrent Mortality Rate -Female Mice

	Vehicle 0 mg|kg|day		Low 50 mg|kg|day		Middle 150 mg|kg|day		High 500 mg|kg|day		Positive MNU: 75	
Week	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %	No. of Death	Cum. %
0 - 13	.	.	1	4.00	.	.	.	.	.	.
14 - 26	.	.	.	.	.	.	1	4.00	6	40.00
Ter. Sac.	25	100.00	24	96.00	25	100.00	24	96.00	9	60.00

Cum. %: Cumulative percentage except for Ter. Sac.

Table 14: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control, Positive Control and Vehicle Control -Male Mice

Test	Statistic	P_Value Vehicle vs Treated Groups Dose Response	P_Value Vehicle vs. Low	P_Value Vehicle vs. Med	P_Value Vehicle vs. High	P_Value Vehicle vs. Positive
Dose-Response	Likelihood Ratio	0.5768	0.5362	0.5678	0.5362	0.1281
Homogeneity	Log-Rank	0.8845	0.5396	0.5717	0.5396	0.1148

Table 15: Intercurrent Mortality Comparison between Treated Groups and Vehicle Control, Positive Control and Vehicle Control --Female Mice

Test	Statistic	P_Value Vehicle vs Treated Groups Dose Response	P_Value Vehicle vs. Low	P_Value Vehicle vs. Med	P_Value Vehicle vs. High	P_Value Vehicle vs. Positive
Dose-Response	Likelihood Ratio	0.4853	0.2390	.	0.2390	<0.0001
Homogeneity	Log-Rank	0.5681	0.3173	.	0.3173	0.0001

Table 16: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons between Vehicle Control and the Treated Groups-Male Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25) P-value - Trend	5 mg/kg/day Low (N=25) P-value - Vehicle vs. Low	10 mg/kg/day Med (N=25) P-value - Vehicle vs. Med	20 mg/kg/day High (N=25) P-value - Vehicle vs. High
ADIPOSE TISSUE	HEMANGIOSARCOMA	0/25 (24) 0.5000	0/25 (25) NC	1/25 (24) 0.5000	0/25 (25) NC
GALLBLADDER	CARCINOMA, SQUAMOUS CELL	1/25 (25) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
HARDERIAN GLANDS	ADENOMA	0/25 (24) 0.8157	2/25 (25) 0.2551	0/25 (24) NC	0/25 (25) NC
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	0/25 (24) 0.7551	1/25 (25) 0.5102	0/25 (24) NC	0/25 (25) NC
	CARCINOMA, BRONCHIOLAR ALVEOL	1/25 (24) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
	CARCINOMA, SQUAMOUS CELL	1/25 (25) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
Lung	C bronchiolar alveolar Adeno+Carcin	1/25 (24) 0.9419	1/25 (25) 0.7653	0/25 (24) 1.0000	0/25 (25) 1.0000
LYMPH NODE	CARCINOMA, SQUAMOUS CELL	1/25 (25) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
MESENTERY/PERITONE	CARCINOMA, SQUAMOUS CELL	1/25 (25) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
MULTICENTRIC NEOPL	HEMANGIOSARCOMA	0/25 (24) 0.1894	0/25 (25) NC	1/25 (24) 0.5000	1/25 (25) 0.5102
SPLEEN	HEMANGIOSARCOMA	0/25 (24) 0.2551	0/25 (25) NC	0/25 (24) NC	1/25 (25) 0.5102
STOMACH, NONGLANDU	CARCINOMA, SQUAMOUS CELL	1/25 (25) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000
THYMUS	THYMOMA	0/25 (24) 0.2551	0/25 (25) NC	0/25 (24) NC	1/25 (25) 0.5102
Whole Body	C_Hemangiosarcoma	0/25 (24) 0.1894	0/25 (25) NC	1/25 (24) 0.5000	1/25 (25) 0.5102

Table 17: Tumor Rates and P-Values for Comparisons between Vehicle Control and Positive Control-Male Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25)	Positive (N=15) P-value - Vehicle vs. Positive
ADIPOSE TISSUE	HEMANGIOSARCOMA	0/25 (24)	0/15 (13) NC
GALLBLADDER	CARCINOMA, SQUAMOUS CELL	1/25 (25)	0/15 (13) 1.0000
HARDERIAN GLANDS	ADENOMA	0/25 (24)	0/15 (13) NC
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	0/25 (24)	4/15 (13) 0.0108
	CARCINOMA, BRONCHIOLAR ALVEOL	1/25 (24)	0/15 (13) 1.0000
	CARCINOMA, SQUAMOUS CELL	1/25 (25)	0/15 (13) 1.0000
	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136
LYMPH NODE	CARCINOMA, SQUAMOUS CELL	1/25 (25)	0/15 (13) 1.0000
LYMPH NODE, MANDIB	LYMPHOMA	0/25 (24)	1/15 (14) 0.3684
LYMPH NODE, MEDIAS	LYMPHOMA	0/25 (24)	2/15 (14) 0.1294
Lung	C_bronchiolar alveolar Adeno+Carcin	1/25 (24)	4/15 (13) 0.0423
MESENTERY/PERITONE	CARCINOMA, SQUAMOUS CELL	1/25 (25)	0/15 (13) 1.0000
MULTICENTRIC NEOPL	HEMANGIOSARCOMA	0/25 (24)	0/15 (13) NC
	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136
SKIN	PAPILLOMA, SQUAMOUS CELL	0/25 (24)	3/15 (14) 0.0431
SPLEEN	LYMPHOMA	0/25 (24)	3/15 (14) 0.0431
STOMACH, NONGLANDU	CARCINOMA, SQUAMOUS CELL	1/25 (25)	2/15 (14) 0.2888
	PAPILLOMA, SQUAMOUS CELL	0/25 (24)	12/15 (14) <0.001

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25)	Positive (N=15) P-value - Vehicle vs. Positive
THYMUS	LYMPHOMA	0/25 (24)	4/15 (14) 0.0136
Whole Body	C_Hemangiosarcoma	0/25 (24)	0/15 (13) NC

Table 18: Tumor Rates and P-Values for Dose Response Relationship and Pairwise Comparisons between Vehicle Control and the Treated Groups-Female Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25) P-value - Trend	5 mg/kg/day Low (N=25) P-value - Vehicle vs. Low	10 mg/kg/day Med (N=25) P-value - Vehicle vs. Med	20 mg/kg/day High (N=25) P-value - Vehicle vs. High
HARDERIAN GLANDS	ADENOMA	0/25 (25) 0.1022	1/25 (24) 0.4898	0/25 (25) NC	2/25 (24) 0.2347
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	0/25 (25) 0.3055	1/25 (24) 0.4898	0/25 (25) NC	1/25 (24) 0.4898
	CARCINOMA, BRONCHIOLAR ALVEOL	1/25 (25) 1.0000	0/25 (24) 1.0000	0/25 (25) 1.0000	0/25 (24) 1.0000
Lung	C bronchiolar alveolar Adeno+Carcin	1/25 (25) 0.5889	1/25 (24) 0.7449	0/25 (25) 1.0000	1/25 (24) 0.7449
MULTICENTRIC NEOPL	HEMANGIOSARCOMA	1/25 (25) 0.7309	0/25 (24) 1.0000	2/25 (25) 0.5000	0/25 (24) 1.0000
SPLEEN	HEMANGIOSARCOMA	1/25 (25) 0.7309	0/25 (24) 1.0000	2/25 (25) 0.5000	0/25 (24) 1.0000
THYMUS	THYMOMA	0/25 (25) 0.7449	1/25 (24) 0.4898	0/25 (25) NC	0/25 (24) NC
Whole Body	C_Hemangiosarcoma	1/25 (25) 0.7309	0/25 (24) 1.0000	2/25 (25) 0.5000	0/25 (24) 1.0000

Table 19: Tumor Rates and P-Values for Comparisons between Vehicle Control and Positive Control -Female Mice

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25)	Positive (N=15) P-value - Vehicle vs. Positive
HARDERIAN GLANDS	ADENOMA	0/25 (25)	0/15 (11) NC
LUNG	ADENOMA, BRONCHIOLAR ALVEOLAR	0/25 (25)	2/15 (12) 0.0991
	CARCINOMA, BRONCHIOLAR ALVEOL	1/25 (25)	0/15 (11) 1.0000
	LYMPHOMA	0/25 (25)	5/15 (14) 0.0035
LYMPH NODE, MANDIB	LYMPHOMA	0/25 (25)	1/15 (12) 0.3243
LYMPH NODE, MESENT	LYMPHOMA	0/25 (25)	1/15 (12) 0.3243
Lung	C bronchiolar alveolar Adeno+Carcin	1/25 (25)	2/15 (12) 0.2407
MEDIASTINUM/PLEURA	LYMPHOMA	0/25 (25)	1/15 (12) 0.3243
MULTICENTRIC NEOPL	HEMANGIOSARCOMA	1/25 (25)	1/15 (11) 0.5238
	LYMPHOMA	0/25 (25)	6/15 (14) <0.001
SKIN	PAPILLOMA, SQUAMOUS CELL	0/25 (25)	7/15 (14) <0.001
SMALL INTESTINE, D	ADENOMA	0/25 (25)	1/15 (11) 0.3056
SPLEEN	HEMANGIOSARCOMA	1/25 (25)	1/15 (11) 0.5238
	LYMPHOMA	0/25 (25)	4/15 (14) 0.0122
STOMACH, NONGLANDU	PAPILLOMA, SQUAMOUS CELL	0/25 (25)	11/15 (14) <0.001
THYMUS	LYMPHOMA	0/25 (25)	6/15 (14) <0.001
	THYMOMA	0/25 (25)	0/15 (11) NC
UTERUS WITH CERVIX	PAPILLOMA, SQUAMOUS CELL	0/25 (25)	1/15 (11) 0.3056

Organ Name	Tumor Name	0 mg/kg/day Vehicle (N=25)	Positive (N=15) P-value - Vehicle vs. Positive
	POLYP, GLANDULAR	0/25 (25)	4/15 (11) 0.0056
Whole Body	C_Hemangiosarcoma	1/25 (25)	1/15 (11) 0.5238

Figure 1: Kaplan-Meier Survival Functions for Male Rats

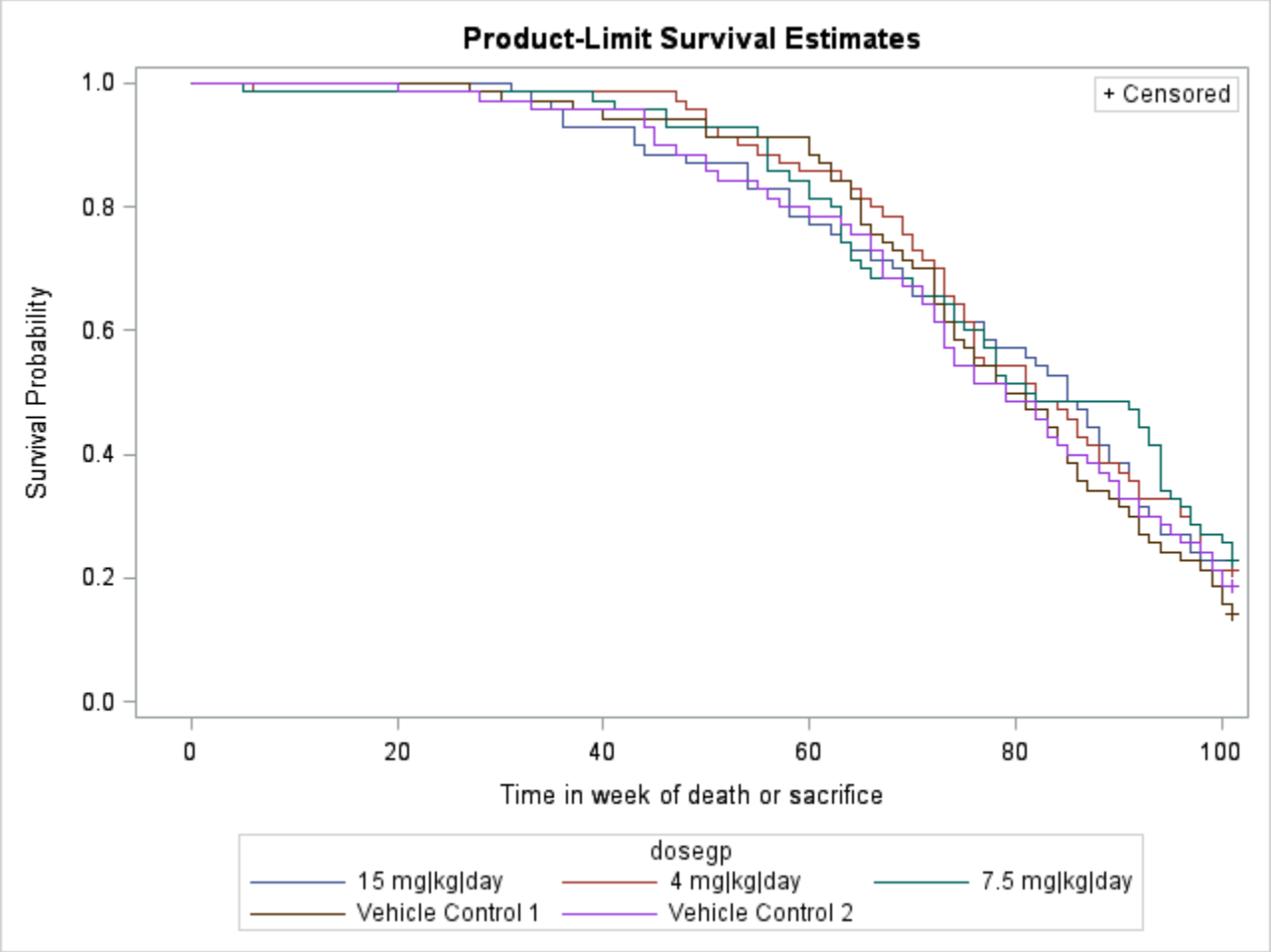


Figure 2: Kaplan-Meier Survival Functions for Female Rats

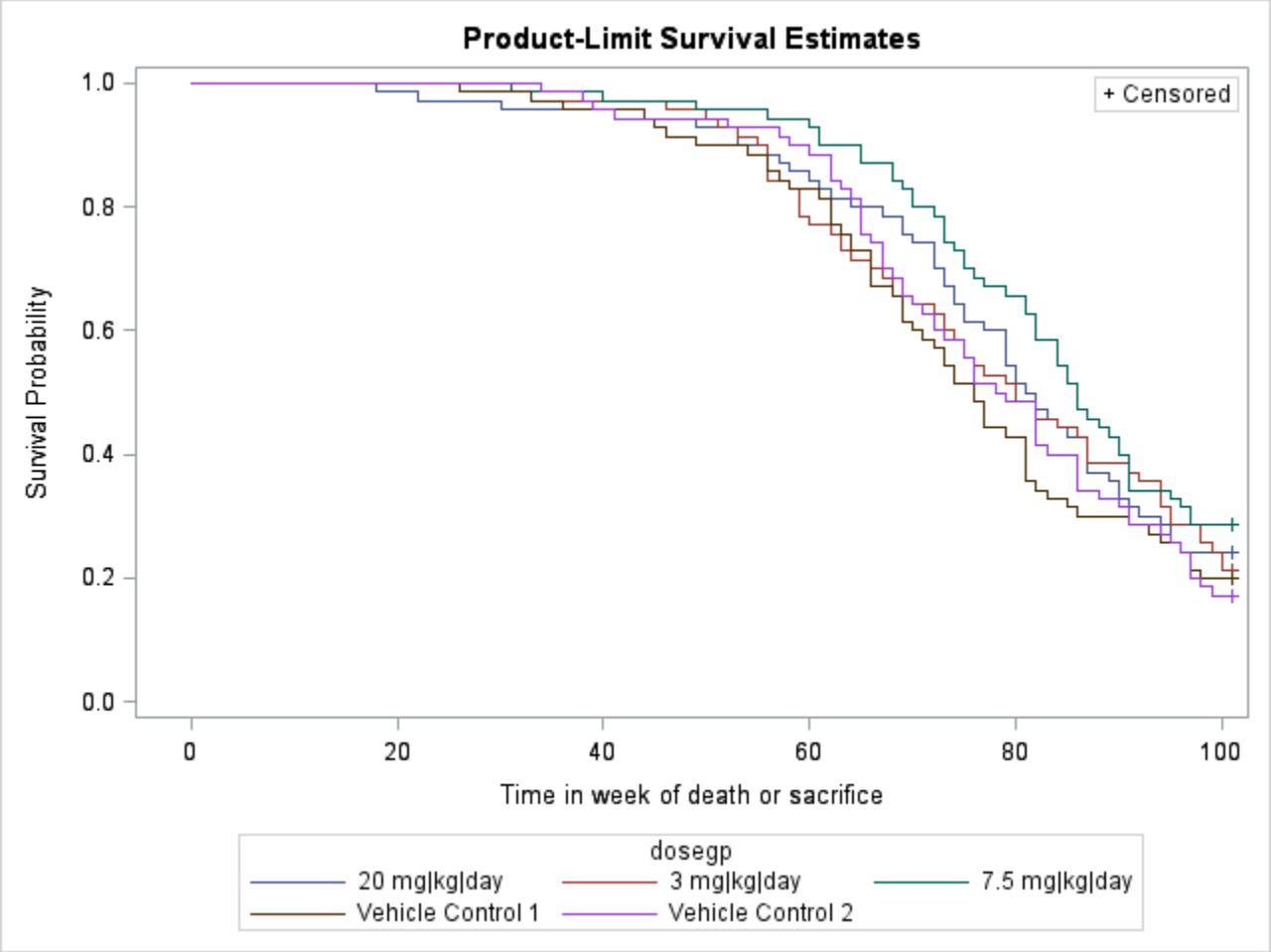


Figure 3: Kaplan-Meier Survival Functions for Male Mice

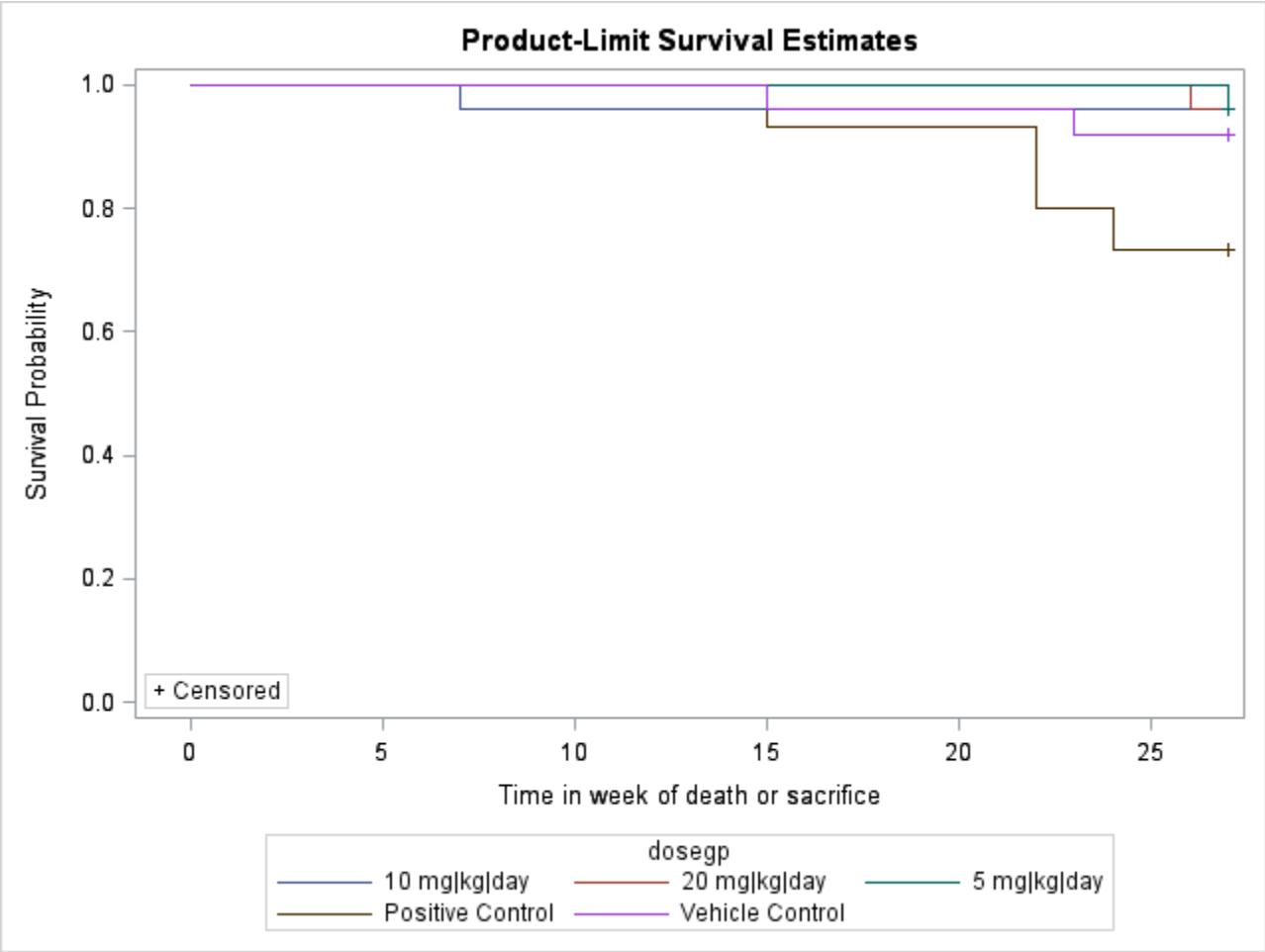
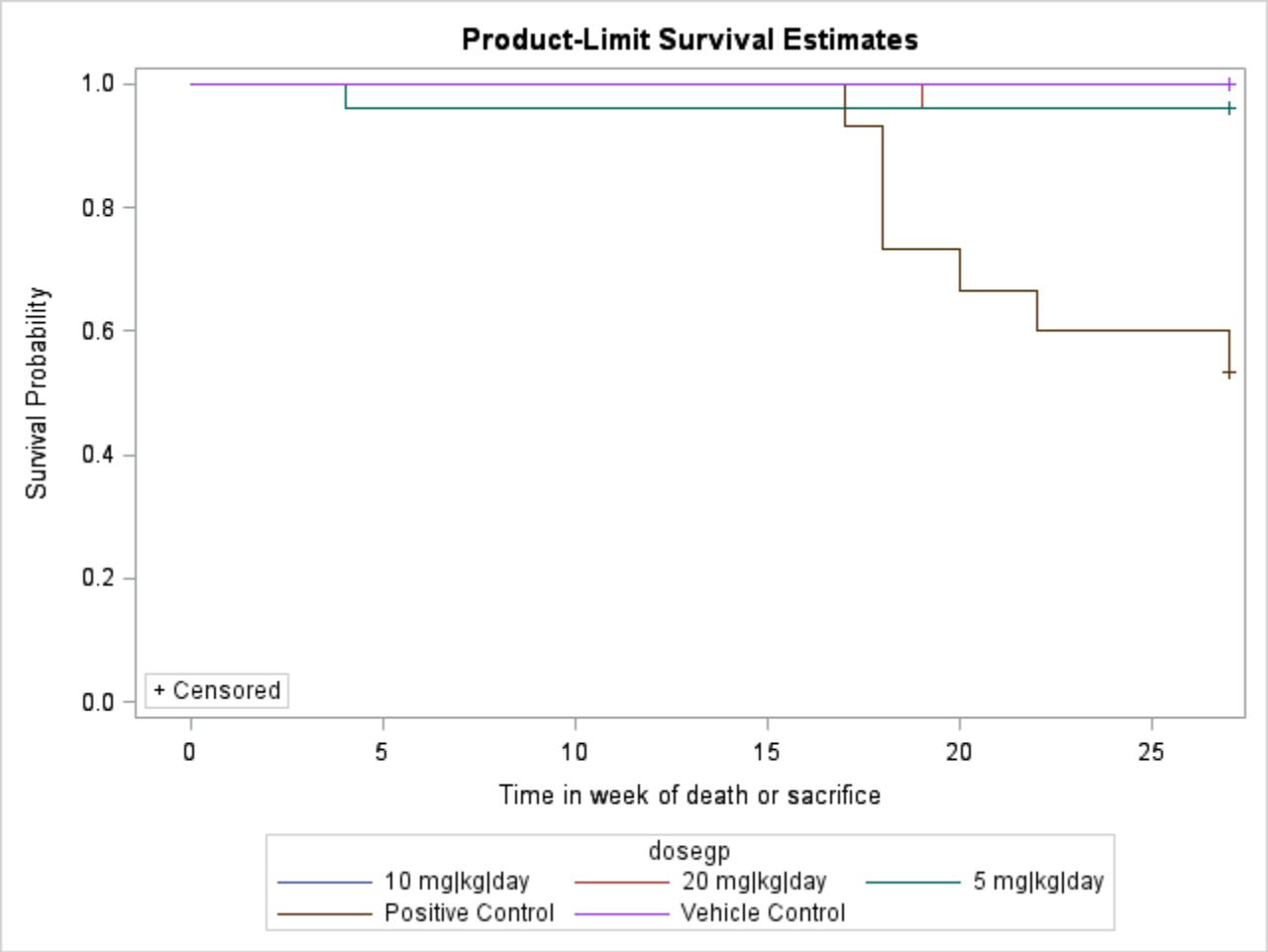


Figure 4: Kaplan-Meier Survival Functions for Female Mice



7. References

- Kaplan EL and Meier P (1958) Nonparametric estimation from incomplete observations. J. Am. Statist. Assoc., 53, 457-481.
- Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports, 50, 163-170.
- Peto R (1974) Guidelines on the analysis of tumour rates and death rates in experimental animals. British J. Cancer, 29, 101-105.
- Lin KK (2000) Carcinogenicity Studies of Pharmaceuticals. In: Encyclopedia of Biopharmaceutical Statistics, ed. Shein-Chung Chow, Marcel Dekker, New York.
- Peto R et al. (1980) Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: Long term and Short term Screen Assays for Carcinogens: A Critical Appraisal. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Supplement 2, pp.311-426. WHO International Agency for Research on Cancer, Lyon.
- SAS Institute (2002) SAS OnlineDoc® Version Nine. SAS Institute Inc., Cary, NC, USA.
- Peto, R., M.C. Pike, N.E. Day, R.G. Gray, P.N. Lee, S. Parish, J. Peto, Richards, and J.Wahrendorf, "Guidelines for sample sensitive significance test for carcinogenic effects in long-term animal experiments", Long term and short term screening assays for carcinogens: A critical appraisal, International agency for research against cancer monographs, *Annex to supplement, World Health Organization, Geneva,* 311-426, 1980.
- Bailer AJ, Portier CJ (1988). "Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples." *Biometrics*, 44, 417-431.
- Bieler, G. S. and Williams, R. L. (1993). "Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity". *Biometrics* 49, 793-801.
- Tarone RE, "Test for trend in life table analysis", *Biometrika* 1975, 62: 679-82
- Lin K.K. and Rahman M.A.," Overall false positive rates in tests for linear trend in tumor incidence in animal carcinogenicity studies of new drugs", *Journal of Biopharmaceutical Statistics,* 8(1), 1-15, 1998.
- Rahman, A.M., and K.K. Lin (2008), "A Comparison of False Positive Rates of Peto and Poly-3 methods for Long-Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", Journal of Biopharmaceutical Statistics, 18:5, 849-858.
- Haseman, J, "A re-examination of false-positive rates for carcinogenesis studies", *Fundamental and Applied Toxicology*, 3: 334-339, 1983.
- Guidance for Industry. Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Statues of Pharmaceuticals (Draft Guidance). U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), May 2001.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

ZHUANG MIAO
03/26/2019 11:44:11 AM

FENG ZHOU
03/27/2019 10:12:58 AM

KARL K LIN
03/27/2019 09:34:22 PM
Concur with review.

YI TSONG
03/28/2019 07:26:42 AM