Official Title: A Phase IIIB, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Clinical Efficacy Study of Baloxavir Marboxil for the Reduction of Direct Transmission of Influenza From Otherwise Healthy Patients to Household Contacts

NCT Number: NCT03969212

Document Date: SAP Version 4: 06-August-2024

STATISTICAL ANALYSIS PLAN

STUDY TITLE: **A PHASE IIIb, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, CLINICAL EFFICACY STUDY OF BALOXAVIR MARBOXIL FOR THE REDUCTION OF DIRECT TRANSMISSION OF INFLUENZA FROM OTHERWISE HEALTHY PATIENTS TO HOUSEHOLD CONTACTS**

STUDY NUMBER: MV40618

VERSION NUMBER: 4

ROCHE COMPOUND Baloxavir Marboxil (RO7191686)

EUDRACT NUMBER: 2018-004056-37

IND NUMBER: 126653

NCT NUMBER: NCT03969212

PLAN PREPARED BY: [redacted], MD, Ph.D.
[redacted], Ph.D.
[redacted], MSc
[redacted], MSc
[redacted], Ph.D.

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

SPONSOR: F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS: Grenzacherstrasse 124
4070 Basel, Switzerland

DATE FINAL: See electronic date stamp on the last page of this document

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP model document V2.0.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
Version 4	see electronic date stamp on the last page of this document	Version 4, 29 March 2022
Version 3	7-June-2024	Version 4, 29 March 2022
Version 2	19-July-2023	Version 4, 29 March 2022
Version 1	28-September-2022	Version 4, 29 March 2022

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP, along with the rationale for each change, are summarized below.

Section	Description of Change	Rationale for Change
4.2	Updated description of inter-current event of death	Clarification on handling of death of an IP and strategy for both HHC and IP deaths
4.2	Removed RR from population-level summary measures	The odds ratio is the main population-level summary and the RR is supportive
4.2.2	Updated CI for RRR to use bootstrap method	To address concerns with the Zhang and Yu (1998) approach for the CI
4.3.1	Updated significance level (and associated confidence interval width) for all secondary endpoints in the testing hierarchy	To ensure strong control of the type I error
4.6.2	Updated the analysis sets for the summaries of treatment group comparability	Clarification on the analysis sets that will be used for demography, vital signs, medical history and concomitant medication summaries
4.7.1	Updated description of testing at interim and final analysis	Clarification that interim analysis was one-sided but final analysis is two-sided

Additional minor changes have been made throughout to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AE	adverse event
BLOQ	below the limit of quantification
BXM	baloxavir marboxil
CI	confidence interval
CMH	cochran-mantel-haenszel
COVID-19	Coronavirus Disease 2019
CSR	Clinical Study Report
CTCAE	common terminology criteria for adverse events
DAP M2	Data Analysis Plan Module 2
eCRF	electronic Case Report Form
EQ-5D-5L	Euroqol 5-dimension Questionnaire
FAS-HC	full household contacts analysis set
FASi-HC	full household contacts infected analysis set
FAS-HH	full household analysis set
FASi-HH	full household infected analysis set
FAS-IP	full index patients analysis set
GCP	Good Clinical Practice
GEE	generalized estimating equations
HCP	healthcare professional
HHC	household contact
I38X	substitution of isoleucine for another amino acid at position 38 (of the polymerase acidic protein)
IA	interim analysis
iDMC	independent Data Monitoring Committee
IMP	investigational medicinal product
IP	index patient
IxRS	interactive voice or web-based response system
MDD	minimum detectable difference
MedDRA	The Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
OR	odds ratio
OwH	otherwise healthy
PA	polymerase acidic (protein)
PAS-HC	primary household contacts analysis set
PAS-HH	primary households analysis set

Abbreviation or Term	Description
PAS-IP	primary index patients analysis set
PCR	polymerase chain reaction
POC	point of care
PT	preferred terms
RAS	resistance associated substitution (PA/I38X or PA/T20K [influenza B only] and other USPI defined substitutions)
RIDT	rapid influenza diagnostic test
RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
RR	relative risk
RRR	relative risk reduction
SAE	serious adverse event
SAP	Statistical Analysis Plan
SARS-CoV-2	syndrome coronavirus 2
SAS	statistical analysis system
SCR-HHC	screened household contacts population
SCR-IP	screened index patients population
SD	standard deviation
SOC	system organ class
T20K	substitution of threonine to lysine at position 20 in the polymerase acidic protein
$TCID_{50}$	50% tissue culture infectious dose
US	United States
USPI	United States Package Insert
VAS	Visual Analogue Scale
WPAI + CIQ	Work Productivity and Activity Impairment Questionnaire plus Classroom Impairment Questions

1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analyses that are planned to be performed for the Clinical Study Report (CSR) of Study MV40618 (Centerstone). Additional detailed algorithms for endpoint derivation and tables, listings and figures shells are provided in the Data Analysis Plan Module 2 (DAP M2).

1.1 OBJECTIVES AND ENDPOINTS

This study will evaluate the efficacy of a single, oral dose of baloxavir marboxil (BXM) compared with placebo for the reduction of the direct transmission rate of influenza A or B from otherwise healthy index patients (IPs) to household contacts. Only household contacts (HHCs) with influenza type A and B will be considered for primary and secondary efficacy endpoints. Specific objectives and corresponding endpoints for the study based on the study protocol are outlined in Table 1 and Table 2.

Table 1 Primary and Selected Secondary Endpoints

Primary Objective(s)	Definition
• To evaluate the efficacy of a single, oral dose of BXM compared with placebo to prevent secondary within-household transmission of influenza A or B	• Virological transmission by Day 5: Proportion of HHCs who become PCR(+) [confirmed at central laboratory] for influenza by Day 5 Visit, with virus subtype consistent with IP
Selected Secondary Objective(s)	**Definition**
• To evaluate the efficacy of a single, oral dose of BXM compared with placebo to prevent transmission of influenza A/ B beyond secondary within-household transmission	• Symptomatic transmission by Day 5: Proportion of HHCs who become PCR (+) [confirmed at central laboratory] for influenza by Day 5 Visit, with virus subtype consistent with IP, AND: 1. For HHC aged ≥ 12 years: – Temperature ≥ 38.0 °C*[1] and one respiratory symptom, OR – One respiratory symptom and one general systemic symptom (with or without fever). 2. For HHC aged ≥ 2 to < 12 years: – Temperature ≥ 38.0 °C* AND signs or symptoms of an upper respiratory tract infection.

[1] Temperature obtained from tympanic thermometers provided to households (or exceptionally from other thermometers/locations in case of tympanic thermometer failure for any reason)

Table 1 Primary and Selected Secondary Endpoints

	3. Symptoms must be either new, or have worsened versus baseline in HHC with baseline symptoms due to a pre-existing comorbidity
	• Virological transmission at the household level by Day 5: Proportion of households with at least one HHC who meets the primary endpoint
	• Symptomatic transmission at the household level by Day 5: Proportion of households with at least one HHC who meets the "Symptomatic transmission by Day 5 endpoint.
	• Virological transmission by Day 9: Proportion of HHCs who become PCR (+) (confirmed at central laboratory) for influenza by Day 9 Visit, with virus subtype consistent with IP, including:
	  • all HHC meeting primary endpoint, AND
	  • all HHC cases detected after Day 5 Visit meeting the following criteria:
	  – Included HHC case is in a household where another HHC has already met the primary endpoint, OR
	  – Included HHC case is PCR(+) [confirmed at central laboratory] for influenza bearing an amino acid substitution of isoleucine for another amino acid at position 38 (I38X) in the polymerase acidic (PA) protein (PA/I38X substitution) or amino acid substitution of threonine to lysine at position 20 in the polymerase acidic (PA) protein for influenza B only (PA/T20K)
	• Symptomatic transmission by Day 9: Proportion of HHCs who meet the "Virological transmission by Day 9" endpoint AND are symptomatic per the definition for symptoms in the "Symptomatic transmission by Day 5" endpoint.
	• Any virological infection by Day 9: Proportion of HHCs who become PCR (+) for influenza (confirmed at central laboratory) by Day 9.
	• Any virological infection at the household level by Day 9: Proportion of households with at least one

Table 1 Primary and Selected Secondary Endpoints

	HHC who meets the “Any virological infection by Day 9” endpoint. • Any symptomatic infection by Day 9: Proportion of HHCs who meet the “Any virological infection by Day 9” endpoint AND are symptomatic per the definition for symptoms in the “Symptomatic transmission by Day 5” endpoint. • Any symptomatic infection at the household level by Day 9: Proportion of households with at least one HHC who meets the “Any symptomatic infection by Day 9” endpoint.

BXM = baloxavir marboxil; HHC = household contacts; IP = index patient; PA/I38X = polymerase acidic (protein)/ substitution of isoleucine for another amino acid at position 38 (of the polymerase acidic protein); PA/T20K = polymerase acidic (protein)/ substitution of threonine to lysine at position 20; PCR = polymerase chain reaction.

Table 2 Other Secondary Objectives and Endpoints

Additional Efficacy Objectives	Corresponding Endpoints
• To further evaluate the efficacy of a single, oral dose of BXM compared with placebo to prevent transmission of influenza A or B.	• The effect of IP baseline viral titer on the rate of direct transmission of influenza A or influenza B to HHCs. • Baloxavir arm only: Proportion of IPs who are PCR(+) [confirmed at central laboratory] for influenza bearing a PA/I38X or substitution of threonine to lysine at position 20 in the polymerase acidic protein (T20K) (influenza B only) substitution (or other identified substitution) post randomization. • Baloxavir arm only: Proportion of HHCs who become PCR(+) [confirmed at central laboratory] for influenza bearing a PA/I38X or T20K (influenza B only) substitution (or other identified substitution) with matched PA substitution with their IP by Day 9 Visit. • Baloxavir arm only: Proportion of symptomatic cases among HHCs who become PCR(+) [confirmed at central laboratory] for influenza bearing a PA/I38X or T20K (influenza B only) substitution (or other identified substitution) with matched PA substitution with their IP by Day 9 Visit (utilizing same symptoms definition as used for secondary symptomatic transmission endpoint). • Proportion of HHC who become PCR(+) [confirmed at central laboratory] for influenza by Day 9 Visit, with virus subtype consistent with IP. • Measurement of viral titer by PCR and 50% tissue culture infectious dose (TCID50) in IPs over the Day 0, 3, 5, and 9 Visits • **Subgroup analysis:** in a post-hoc analysis, endpoints requiring a subtype match will be assessed with a match based on sequencing in a subset of households (sequencing will be performed for samples from selected sites and countries based on feasibility, including availability of adequate number of IP sequences for specific community).

Table 2 Other Secondary Objectives and Endpoints

Safety Objectives	Corresponding Endpoints
• To evaluate the safety and tolerability of a single, oral dose of BXM compared with placebo.	• **Index patient only:** The incidence, severity, and timing of adverse events, and serious adverse events.
Health Status Utility Objectives	**Corresponding Endpoints**
• To evaluate index patients treated with BXM compared with those receiving placebo.	• **Index patient aged 12 and above only:** Change from baseline in health-related quality of life according to EuroQol 5 dimensions 5 (EQ-5D-5L) Questionnaire at Day 3 and Day 9 Visits. • **Index patient age 12 and above only**: Work Productivity and Activity Impairment Questionnaire plus Classroom Impairment Questions: SHP, Version 2 (WPAI + CIQ:SHP, V 2) to measure lost work or school day assessed at Day 9 Visit.
Additional Objective	**Palatability and Acceptability Endpoints**
• To describe the palatability and acceptability of BXM oral suspension	• **Index patient aged 12 and below only**: Distribution of responses to the Palatability and Acceptability Questionnaire of Study Drug

BXM = baloxavir marboxil; HHC = household contacts; IP = index patient; PA/I38X = polymerase acidic (protein)/ substitution of isoleucine for another amino acid at position 38 (of the polymerase acidic protein); PA/T20K = polymerase acidic (protein)/ substitution of threonine to lysine at position 20; PCR = polymerase chain reaction; EQ-5D-5L = EuroQoL 5-Dimension Questionnaire (see Section 5.2); WPAI + CIQ = Work Productivity and Activity Impairment Questionnaire plus Classroom Impairment Questions (see Section 5.3); Palatability and Acceptability Questionnaire of Study Drug (see Section 5.4).

1.2 STUDY DESIGN

This is a randomized, double-blind, multicenter, parallel group, placebo-controlled study designed to evaluate the clinical efficacy of BXM for the reduction of direct transmission of influenza A or B from otherwise healthy (OwH) IPs to their HHCs.

For this household study, IPs with influenza will be randomized to receive BXM or placebo, and their HHCs will be repeatedly tested for influenza virus and assessed for influenza symptoms during the next 9 days. The total number of new HHC infections (symptomatic and asymptomatic) and total number of new HHC infections associated with symptoms will be key measures used to evaluate the “virological transmission” and “symptomatic transmission” endpoints (see Section 2.2).

Approximately 1130 IPs with influenza and approximately 2030 evaluable HHCs are planned to participate (assuming an average of 2.5 HHCs per IP, a 15% exclusion rate at the household level, and a 15% drop-out rate) in the study. The total study duration for each household will be up to 9 ($\pm$1) days (IPs aged 5–11 years old will have a safety follow-up visit at 21 [+2] days). This study will be conducted at approximately 200 sites globally in the Northern and Southern hemispheres.

At the end of the 2022/2023 Northern hemisphere influenza season, an interim efficacy analysis will be conducted with early stopping rules for proven efficacy (based on a group sequential design) and futility. See Section 4.7.1 for details.

Figure 1 presents an overview of the study design. A schedule of activities for the IPs and HHCs is provided in study protocol.

Index Patient: Screening and Randomization

Eligible IPs must be aged from $\geq$5 to $\leq$64 years old, have influenza symptom onset within 48 hours, test positive for influenza A/B, and be OwH (i.e., not at high risk for complications of influenza). Screening assessments informing IP eligibility include physical examination, vital signs, height and weight, medical history and concomitant therapies, urine pregnancy testing, and respiratory sampling for influenza testing.

With regard to the household, IPs should be determined at screening to live with $\geq$1 HHCs who have not received an influenza vaccine in the past 6 months ("unvaccinated HHCs"), are likely to fulfill all HHC eligibility criteria, and are expected to participate in the "full study" (i.e., participate in all study assessments).

IPs who meet the eligibility criteria (see Protocol Section 4.1) will be randomized in a 1:1 ratio to receive a single dose of either BXM or placebo within 2 hours of randomization. The dose and formulation of BXM is based on weight and age (see Section 1.2.1).

Index Patient: Post-Randomization

Respiratory samples and adverse events (AEs) will be collected. IPs who are $\geq$12 years old will also complete questionnaires describing their health status and absence from work or school (see Section 5.2 and Section 5.3). IPs who are $<$12 years old will be asked to complete a questionnaire on the palatability and acceptability of the study drug oral suspension.

Household Contacts: Screening and Enrollment

All HHCs present in the home must have their screening visit start within 24 hours of IP randomization (see Protocol Section 4.5.8 for details). If any HHC tests positive for influenza A/B, then all HHCs fail screening. If $\geq$1 unvaccinated HHCs meet all HHC eligibility criteria and agree to participate in the full study, it is allowable for additional HHCs to not participate beyond the screening visit (even if they meet all HHC eligibility

criteria). There is no maximum number of vaccinated and unvaccinated HHCs that can participate in the full study.

HHCs who agree to participate in the full study ("full study HHCs") must meet the full study criteria (HHC inclusion criteria 7 to 14; see Protocol Section 4.1.1, including that they must reside in the household for 7 of the 9 study days, and must not have any influenza symptoms at screening (mild symptoms determined by the investigator to be due to a pre-existing condition are allowed).

Screening assessments informing HHC eligibility include influenza symptoms (see Section 5.1), medical history, concomitant medications, and respiratory sampling for influenza testing.

Full Study Household Contacts: Post-Screening

Only full study HHCs participate in visits and assessments post-screening. Full study HHCs are monitored for new or worsening influenza symptoms (see Section 5.1) and will also maintain a daily temperature diary (see Protocol Section 4.5.6). HHCs will be instructed to telephone the site if they develop influenza symptoms or fever so that a scheduled or unscheduled visit occurring within 24 hours can be arranged (see Protocol Section 4.5.6). Respiratory samples, adverse events due to study procedures, and concomitant medications will be collected.

SARS-CoV-2 Testing

Given overlapping influenza and Coronavirus disease 2019 (COVID-19) symptoms, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) testing will be conducted at screening and as needed during study conduct (see Protocol Section 4.5.5.2). If any subject (IP or HHC) tests positive for SARS-CoV-2, then all subjects in the household should be discontinued (see Protocol Section 4.6.2).

Independent Data Monitoring Committee

An external independent Data Monitoring Committee (iDMC) will evaluate safety according to policies and procedures detailed in an iDMC Charter. The involvement of the iDMC in interim efficacy evaluations is described in Section 4.7.

1.2.1 Treatment Assignment and Blinding

This is a single, oral dose study. IPs will receive the initial and only dose of study drug (BXM or matching placebo) to be taken in tablet or oral suspension form at the study center. See Table 3 and Table 4 for study drug (BXM or matching placebo) administration.

Table 3 Study Drug Administration and Dose for Index Patients < 12 Years Old

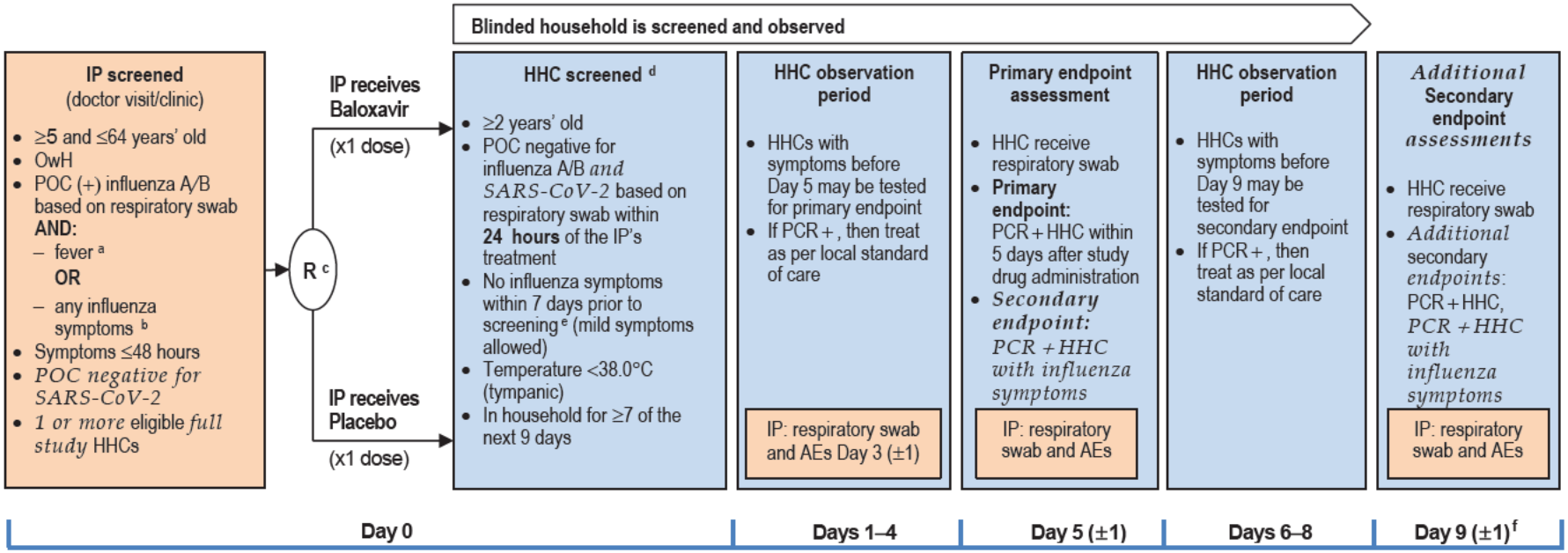
Weight	Dose, Formulation
< 20 kg	2 mg/kg, oral suspension
≥ 20 kg	40 mg, oral suspension

Table 4 Study Drug Administration and Dose for Index Patients ≥ 12 Years Old

Weight	Dose, Formulation
< 80 kg	40 mg, two 20 mg tablets
≥ 80 kg	80 mg, four 20 mg tablets

See details in study Protocol Section 4.2 “Method of Treatment Assignment and Blinding”.

Figure 1 Study Schema



AEs = adverse events; HHCs = household contacts; IP = index patient; OwH = otherwise healthy; PCR = polymerase chain reaction; POC = point of care; SARS-CoV-2 = syndrome coronavirus 2.

[a] ≥38.0°C per tympanic or rectal thermometer; ≥37.5°C per axillary, oral or forehead/temporal thermometer.

[b] Cough, sore throat, nasal congestion, headache, feverishness or chills, muscle or joint pain, fatigue.

[c] Stratification factors: age; household size; region; time since symptom onset.

[d] HHC screening must start within 24 hours of IP randomization and may occur on IP study Day 0 or 1.

[e] Symptoms for HHC ≥12 years old: cough, sore throat, nasal congestion, headache, feverishness or chills, muscle or joint pain, fatigue; Symptoms for HHC ≥2 to <12 years old: cough, nasal congestion, or rhinorrhea.

[f] IPs < 12 years old will have a safety follow-up visit on Day 21 (+2 days).

1.2.2 Randomization

IPs will be allocated to one of the two treatment arms in a 1:1 proportion with predefined stratification factors: region (US or Europe, Asia, Rest of the World), number of household contacts living in the household during the study (≤ 2, ≥ 3), age (≥ 5 to < 12 years; ≥ 12 to ≤ 30 years, > 30 years) and duration of symptoms (≤ 24, > 24 to ≤ 48 hours). A placebo for BXM will be used in the control arm. Patients, investigators, and the Sponsor will be blinded to treatment assignment.

At study completion, stratification factors recorded in the Interactive voice or web-based response system (IxRS) will be compared to the stratification factors derived from data recorded on the electronic Case Report Form (eCRF), which will be used in the analysis, to identify any mis-stratifications. These will be listed for all patients.

1.2.3 Independent Review Facility

Not Applicable.

1.2.4 Data Monitoring

Safety interim analyses will be performed regularly and reviewed by an iDMC. Efficacy data will be provided for interim analysis purposes (see Section 4.7) or if requested by the iDMC. Further details on the function and logistics (e.g., frequency of the meetings) of the iDMC are provided in the iDMC Charter.

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATIONS

2.1 STATISTICAL HYPOTHESES

The hypothesis for the primary analysis is as follows:

- H_0: No difference in virological influenza transmission rate between study treatments (odds ratio $= 1$).
- H_A: The virological influenza transmission rate in Baloxavir arm is different from Placebo (odds ratio $\neq 1$)

2.2 SAMPLE SIZE DETERMINATION

The proposed sample size for this study is 2,030 evaluable household contacts.

Assuming a 15% drop-out for inclusion of household contacts into the primary analysis population, an average of 2.5 household contacts per IP and a 15% exclusion rate on the household level, this equates to an approximate number of IPs of 1,130.

A transmission rate in the placebo arm of approximately 20% is assumed based on data from the literature, including work by Welliver et al. 2001. For a 30% relative risk reduction with BXM, which is considered clinically meaningful, the transmission rate in the active treatment group is expected to be 14%. Under these assumptions and for a two-sided test for binary proportions at the 5% significance level, the power for the

between-arm comparison based on the primary endpoint is expected to be > 90%. For the secondary "Symptomatic transmission endpoint" (symptomatic influenza) the power will be approximately 80% for an assumed 65% proportion of symptomatic cases among infected household contacts, and not considering the hierarchical testing procedure. Computations were performed using a normal approximation to the binomial distributions (R function "power.prop.test"). The literature reports a variety of values for the proportion of influenza cases that are symptomatic. The assumption of 65% was authorized by the study steering committee.

The effect of clustering was examined in simulations and using analysis methods accounting for the household effect. The power was largely maintained in simulations that assumed the existence of a random variation of the background transmission rate between households with a standard deviation up to 5% (i.e., with about 95% of the values for the BXM and placebo groups in the [4.2, 23.8] and [10.2, 29.8] range, respectively). This is compatible with the inclusion of a large number of small clusters and a small expected intra-cluster correlation, resulting in a negligible design effect and related reduction in effective sample size.

2.2.1 Sample Size for China Population

This study will enroll participants across all sites including participants in China. No cap on sample size is currently planned in China subpopulation.

3. ANALYSIS SETS

The primary study analysis will occur when the last subject has either withdrawn or completed their last visit and will be based on cleaned data for all subjects up to and including this point. The end of the study for all patients (except IPs < 12 years old) is expected to occur 9 days after the last IP is enrolled. For IPs < 12 years old, the last visit will be Day 21. Interim analyses are described in Section 4.7.

3.1 SCREENED POPULATIONS

Screened Index Patients Population (SCR-IP): will include all IPs screened at investigational sites with signed Informed Consent.

Screened Household Contacts Population (SCR-HHC): will include all household contacts of the SCR-IP set, screened at investigational sites or at nurse house visits, with signed Informed Consent.

3.2 RANDOMIZED POPULATIONS

Household Contacts Analysis sets:

- Full Household Contacts Analysis set (FAS-HC) will include all HHCs who are enrolled for the full study and linked to a randomized IP.
  Note: household contacts with Baseline collected later than 2 days after IP

randomization cannot be part of the FAS-HC because they would not be eligible to be enrolled for the full study.

- Full Household Contacts infected Analysis set (FASi-HC): subset of all FAS-HC whose IPs are polymerase chain reaction (PCR)+ for influenza A or B at Baseline (see Section 4.1).
- Primary Household Contacts Analysis Set (PAS-HC) is a subset of FASi-HC including all HHCs:
  - From households where the IP is PCR+ for influenza A or B at Baseline (see Section 4.1), and received study drug
  - who are unvaccinated (i.e., that have not received the influenza vaccine within 6 months prior to screening), and
  - from households where all contacts are PCR negative at Baseline (see Section 4.1). Additional criteria for late arrivals HHCs and screened failed HHCs are reported in Section 3.2.1.

Index Patients Analysis sets:

- Full Index Patients Analysis Set (FAS-IP): will include all randomized IPs.
- Primary Index Patients Analysis Set (PAS-IP): will include all index patients from FAS-IP with at least one HHC in the PAS-HC.
- Full Index Patients Analysis Set for BXM Resistance Analysis (RES-IP): will include all IPs randomized to BXM and exposed to BXM, having at least one PA sequencing result

Households Analysis sets:

- Full Household Analysis Set (FAS-HH) will include all household (HH) of randomized IPs.
- Full Households infected Analysis Set (FASi-HH): All households in FAS-HH with the IP PCR positive at screening and with at least one HHC enrolled for the full study. Of note, this corresponds to the HHs of all subjects in FASi-HHC.
- Primary Households Analysis Set (PAS-HH): subset of FASi-HH where IPs are in PAS-IP.

For analysis set derivation purposes, where a central laboratory sample result is missing for any reason, the local laboratory sample would be tested at the central laboratory in an attempt to avoid missing data, this is applicable to any index patient and household contact (see Section 5.5, Local laboratory sample testing for influenza at the central laboratory, scenario 1). If a central laboratory result remains missing following an attempt to retest the local samples centrally (Section 5.5 scenario 1), the local laboratory result will be utilized for the analysis set derivation if the conducted assay was a PCR method (see Section 4.1.1). For index patients where baseline central laboratory result is missing and the local laboratory test result was conducted by a non PCR

method e.g., RIDT, the central laboratory result on protocol Day 3 will be utilized for analysis set derivation (Section 5.5 scenario 1).

In the situation of specific cases where there is a mismatch between the results from the local and central laboratory sample for the IP baseline influenza testing the selection of the value for analysis set derivation is described in Section 5.5 scenario 2

Treatment groups will be defined for all analysis sets above based on the arm to which the respective IP was randomized.

3.2.1 Handling of Late Arrival HHCs and Screen Failed HHCs for PAS-HC.

For HHCs with Baseline PCR values collected after IP randomization + 2 day (see Section 4.1), the full household will be excluded from PAS-HC if:

- HHC has informed consent form (ICF) date < IP randomization + 7 days (prior to protocol study Day 7) and at least one of the following holds true:
  - HHC has no PCR assessment date < IP randomization + 7 days, or
  - HHC has at least one positive PCR assessment, with date < IP randomization + 7 days, having type/subtype matching with IP or with type/subtype missing

The full household would also be excluded from the PAS-HC if a HHC is screen failed with date < IP randomization + 7 days and is screen failed because of influenza test positive result. These cases will be detected from IxRS screen fail reasons (INC02 for protocol v3 and v4, and INC03 for protocol v1 and v2) as the local influenza test data are not collected on the CRF nor samples are analyzed by central laboratory.

3.3 PHARMACOKINETIC-EVALUABLE POPULATION

Not Applicable.

3.4 SAFETY POPULATIONS

The Safety Index Patients Population will include all subjects from the SCR-IP who are randomized and receive any amount of study drug. Patients will be grouped for analysis according to the received treatment. Additional analysis will be conducted in the:

- Pediatric Safety set: composed of the subset of patients in Safety IP set who are less than 12 years old.
- Adolescent Safety set: composed of the subset of patients in Safety IP set with 12 ≤ age < 18 years.
- Adult Safety set: composed of the subset of patients in Safety IP set with age ≥ 18 years.

The Safety Household Contacts Population will include all partial or full study enrolled subjects. Safety data for HHCs will be summarized overall.

4. STATISTICAL ANALYSES

The analyses outlined in this SAP supersede those specified in the protocol. Unless otherwise specified (e.g., in the context of group-sequential analyses, see Section 4.7), statistical hypotheses will be tested at the 5% significance level (alpha = 0.05) against two-sided alternatives. A confirmatory testing strategy for the primary and selected secondary endpoints is described in Section 4.2 and Section 4.3.

All statistical analyses will be performed using SAS statistical software (Version 9.4), unless otherwise specified.

4.1 GENERAL CONSIDERATIONS

All summaries will be produced for the FAS-IP and FAS-HC, unless otherwise specified.

Unless otherwise specified, the Clinical Data Interchange Standards Consortium (CDISC) standard will be used for the analysis and presentation of study days, which correspond to protocol study Day + 1, apart from days prior to randomization, where analysis days and study protocol days are the same, as shown in Table 5.

Table 5 Analysis Periods

	Study Days											
Protocol Days	0[1]	1	2	3	4	5	6	7	8	9	10	11+
Analysis Days	1	2	3	4	5	6	7	8	9	10	11	12+
IP Periods	B[2]	Before Day 5			Day 5			After Day 5				
HHC Periods	B[2]			Up to Day 5				After Day 5				

HHC = household contact; IP = index patient.

Note: [1] Protocol study Day 0 Baloxavir Marboxil (MV40618) SAP v2 = randomization [2] Baseline for a HHC is not later than IP randomization Day + 2 day, to accommodate for late arrival still eligible for inclusion in FASi-HC (Section 3.2).

Summaries over time of data including selected symptoms/signs and viral titer will be produced by visit, and for HHC symptoms, by the post Baseline analysis periods provided in table above.

Baseline values are defined as:

- For IPs: the last assessment prior to study treatment dosing. For baseline influenza status (positive/negative) and type/subtype and any parameters where time is not collected, assessments on or prior to the day of study treatment dosing will be considered as detailed in the DAP-M2. Information on local and central laboratory sample testing related to the baseline definition is provided in Section 5.5.

- For HHCs: the earliest assessment recorded not later than IP randomization day+2 days. Baseline data collected later than IP randomization day+2 will be taken from late screening visits if occurring before any documented post-baseline visit of the HHC

Full details on Baseline derivation will be provided in the DAP-M2 Section 3.1.6.2.

Descriptive statistics (mean, median, SD, range) will be presented for continuous variables, and frequencies and percentages will be presented for categorical variables, unless otherwise specified. The percentages will be calculated based on the number of non-missing values.

4.1.1 Missing Data

The primary focus of this section is on missing influenza assessment data because of their direct impact on key efficacy study endpoints and analysis set derivation.

Missing influenza test data can arise both at the IP/household and HHC level:

- If the IP or at least one household contact have baseline influenza result missing, (unless the late arrival HHC joined the HH after Day 7 - see SAP Section 3.2.1 for details), the whole household will be excluded from the primary analysis set (PAS-HC) because it would be unclear if any HHC infections detected post-baseline would originate from the IP.
- In case of missing influenza type/subtype in positive PCR measurements in HHC (or in IP), the influenza transmission endpoint is considered positive.
- In the absence of PCR measurements post-baseline in HHCs in the relevant time window, the transmission by Day 5 endpoints would be considered non-evaluable. More specifically, the endpoint would be considered as non-evaluable if no post-baseline assessment is taken prior to study Day 7 or if the latest post-baseline assessment is taken prior to study Day 4 and is negative for influenza. Handling of non-evaluable/missing data is described below for the key analyses.
- Where a central laboratory sample result is missing for any reason, the local laboratory sample would be tested at the central laboratory in an attempt to avoid missing data, this is applicable to any index patient and household contact (see Section 5.5, local laboratory sample testing for influenza at the central laboratory, scenario 1). Unless otherwise specified, the point of care (POC) tests or other local laboratory test results for influenza will be used for the analysis, and for analysis sets derivation, if the local assay was a PCR method (i.e., if local test type is PCR or LIAT) when no central lab assessment is available for qualitative PCR (see Chen et al. 2015). These data will be provided in listings. Section 5.5 (scenario 2 and 3) describes the selection of the value for analysis in specific cases where there is a mismatch between the results from the local and central laboratory.

For the symptomatic transmission by Day 5 endpoint, missing data may also arise in case of no symptoms collected post-baseline.

Missing endpoint data or missing scheduled assessments may mainly arise because of:

- Subjects lost to follow-up or discontinued prior to the assessment visit (e.g., prior to Day 5). A conservative 15% drop-out rate was assumed for sample size considerations. Protocol compliance is however expected to be higher because of the short study duration and the possibility to perform study visits at the clinic or by nurse home visits, as per subjects' preference.
- As per protocol, respiratory samples from HHCs only will be tested using the POC system and if any sample is positive, no further respiratory swabs will be taken. In this case, the primary endpoint will be evaluable unless the central assessment is negative. Data will not be collected at Day 5 and/or Day 9 because of prior POC/local laboratory positive test (e.g., at unscheduled visit) not eventually confirmed by central laboratory measurements. A minor impact is expected on primary endpoint, because of the high specificity of POC test and the limited number of unscheduled visits expected by Day 5 visit.
- Missing data due to sample mishandling or other laboratory issues.

For the primary estimand (see Section 4.2), missing/non-evaluable values will be imputed by considering any existing out-of-window post-baseline PCR assessment, or imputing negative if no post-baseline PCR assessment is available.
Other analysis will be conducted on complete cases. This approach was suggested by (Ma et al. 2013) for any clinical trials with a small percentage of missing outcomes ($<$ 15%). Sensitivity analyses will however be conducted for the primary endpoint as specified in Section 4.2.3.

4.2 PRIMARY ENDPOINT ANALYSIS

PAS-HC will be the primary efficacy analysis set. Supportive and secondary analyses will also be conducted on the other analysis sets described in Section 3.2 and Section 3.3.

The primary estimand for the analysis of the primary endpoint and for the first of the secondary endpoints (see 4.3.1.1) is defined as follows (details are provided in the DAP M2):

a) **Treatment regimens for patients to be evaluated**
One dose of oral BXM or placebo given to the IP at baseline.

b) **Populations of patients targeted by the clinical question (e.g., patients meeting inclusion/exclusion criteria, grouped according to randomized treatment assignment)**
PAS-HC, i.e., unvaccinated HHCs who are included in the full study (requirements include that the HHC joins the household no later than on Day 2), who are linked to a baseline PCR positive randomized IP and are from a completely PCR negative HH (except for the IP) at baseline. The PCR test status for the IP and household contacts at baseline need to be confirmed by the central laboratory. HHCs are

grouped for analysis according to the randomized treatment assignment of the respective IP.

c) **Patient-level outcome to be measured**
PCR-confirmed infection (primary endpoint) by Day 5, Symptomatic PCR-confirmed infection by Day 5 (Symptomatic Transmission first secondary endpoint). The window for the Day 5 visit extends to study Day 6. For the definition of symptomatic cases refer to Protocol Section 2.1.2.1.

d) **How intercurrent events will be handled (e.g., regardless of protocol violations)**

- Protocol violations, including BXM dosing issues (e.g. wrong dose), use of prohibited medication, out-of-window assessments (especially after Day 6), or other events impacting the evaluability of the respiratory samples: treatment policy. All available PCR samples will be used for the evaluation of HHCs for the analysis. For HHCs without evaluable post-baseline samples, the status will be defined according to the baseline sample (i.e., negative for the primary analysis). In case no post-baseline symptoms are collected, the HHC will be considered asymptomatic.
- COVID-19 positivity in the HH: while on treatment (outcome measures used until discontinuation)
- Positive PCR in HHC not matching IP's influenza type/subtype: while on treatment strategy (primary endpoint considered negative unless there are subsequent positive assessments that match the IP's influenza type/subtype)
- Death: Death of a HHC will be treated as a transmission event, death of an IP will also result in their associated HHCs being imputed as transmission events (composite strategy).

e) Population-level summary measure (e.g., Proportion of HHCs who become PCR(+) for influenza by Day 5 Visit).
Odds ratio (primary model)

4.2.1 Definition of Primary Endpoint

The primary endpoint will be derived at the HHC level based on central laboratory influenza test results (or local test results if central laboratory results are missing, see Section 4.1.1 and 3.2), as:

- A HHC will be considered positive, if an influenza test is positive within 6 days post IP randomization, with virus type and subtype matching with that of the respective IP
- A HHC will be considered negative, if they do not meet the criteria for positivity and an influenza test is negative at Day 5 ±1
- Missing otherwise

Detailed endpoint derivation specifications in case of missing data on virus type/subtype are provided in DAP M2 Section 3.2.2.

4.2.2 Main Analytical Approach for Primary Endpoint(s)

The primary analysis will be performed on the PAS-HC with imputed missing primary endpoint information as described in Section 4.1.1 and will compare the transmission rate for the primary endpoint between the two treatment arms and express the effect by an adjusted odds ratio (OR). The comparison will be based on a model providing population-averaged estimates using a generalized estimating equations (GEE) approach and accounting for clustering within households and the stratification factors (see Section 1.2.2) based on eCRF data, or IxRS in case of missing eCRF data.

An exchangeable variance–covariance structure will be applied to model the within-household errors, assuming that the correlation between any pair of HHCs within the same household is the same given the absence of a natural order of HHCs within households. If the model doesn't converge, other variance–covariance structures (in the following order: unstructured, independent) or other starting estimates of the model parameters can be considered.

Adjusted point and 100×(1-α*)% CI estimates of the incidence proportions and odds ratio for the treatment effect at Day 5 and the p–value for the treatment effect as well as for each covariate will be provided (see Section 2.1), where α* is the significance level at the final analysis.

The primary test for the treatment effect at the final analysis will be two-sided and the significance level will be derived based on the alpha spend at the interim analysis and the size of the primary analysis population (PAS-HC) at the time of the analysis using the R package rpact (Wassmer and Pahlke 2019), as described in Section 4.7.1. The number of available data for analysis will be tabulated.

Sample SAS code can be found below (SAS code is regarded as "draft" until fully validated at the analysis stage):

```
proc genmod data=PAS_HC;
   class region nbhhc agecat dursymp trt household influenza;
   model influenza = trt region nbhhc agecat dursymp / dist=bin
link=logit type3;
   repeated subject=household / type=exch covb;
   lsmeans trt / ilink exp diff cl alpha=;
run;
```

In addition to adjusted odds ratio, the treatment effect will also be summarized by the relative risk (RR) as a supportive population-level summary. The point estimate of the relative risk reduction (RRR) will be estimated as 1-RR, where RR is the relative risk derived by means of the formula published in Zhang and Yu 1998:

$$RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)}$$

with *OR* and P_0 being the adjusted odds ratio and incidence rate for Placebo obtained from the primary GEE model, respectively. The confidence interval for the RRR will be calculated using the bootstrap method, where the sampling is conducted at the household level to maintain the correlation structure and the RRR for each sample is calculated as described above using the OR and P_0 obtained from the GEE model.

4.2.3 Sensitivity Analyses

The primary analysis will be conducted adjusting for stratification factors derived from eCRF data as described in Section 4.2. If it is found that a large number of mis-stratifications occurred, a sensitivity analysis will be carried out on the primary endpoint, using the IxRS stratification factors in the model.

Sensitivity analysis will be conducted to evaluate the robustness of the results of the primary analysis with respect to different assumptions regarding missing endpoint data.

Sensitivity Analysis Approach	**Rationale**
Complete case analysis (no missing data imputation)	To evaluate the effect under the assumption of data missing completely at random, i.e., assuming that the probability of missing transmission data is unrelated to any other measured variable and also unrelated to the values of the actual transmission itself (National Research Council 2010)
Missing transmission endpoint will be imputed as negative	To evaluate the effect using an approach in line with the FDA guidance “Influenza: Developing Drugs for Treatment and/or Prophylaxis” (April 2011, Section 8) HHCs with insufficient data for the evaluation of the primary endpoint are considered as not having PCR-confirmed influenza
Multiple Missing Data Imputation (MI)	To evaluate the effect imputing missing transmission endpoints using the placebo response rates. Further details are given below the table.
Excluding data from Study sites with health authority identified Good Clinical Practice (GCP) violations	US site 320035 was inspected by the FDA and issued a Form 483, Inspectional Observations and later warning letter for investigator lack of oversight and supervision of the clinical study and failure to ensure that subjects met

Sensitivity Analysis Approach	Rationale
	protocol-required inclusion criteria for a non-Roche study. While no findings identified by the FDA for Study MV40618, data from this site is excluded due to FDA concerns regarding inadequate investigator oversight. A site management organisation Medipharma Co ltd was inspected by the Japanese Ministry of Health, Labor and Welfare (MHLW) and reported to have committed GCP violations, including data falsification. Medipharma supported Japanese sites 325126 and site 323607 in the conduct of study MV40619. Data from the sites supported by medipharma are excluded due to failings in GCP conduct by Medipharma.

FDA = Food and Drug Administration; HHC = house hold contact; PCR = polymerase chain reaction.

Multiple Missing Data Imputation (MI): will consist of the following 3 phases:

Imputation: The missing endpoint data will be filled in with values estimated by a logistic regression model fitted using placebo patients to predict the value of the missing endpoint and thus derive complete data set for analysis. Model terms will include those of the primary analysis model. This imputation process will be repeated 50 times. Example SAS code is provided below (to be validated):

```
proc mi data= PAS_HC nimpute=50 out=mi_PAS_HC seed=&seed; ;
   class region nbhhc agecat dursymp trt influenza;
   fcs logistic (influenza= region nbhhc age_ip age_hhc dursymp
/link=logit);
mnar model(influenza /modelobs=(trt='PLACEBO'))
run;
proc sort; data= mi_PAS_HC; by _imputation_; run;
```

Analysis: each of the complete data sets generated during imputation is then analyzed using the primary analysis model. Example SAS code is provided below (to be validated):

```
proc genmod data=mi_PAS_HC;
   class region nbhhc agecat dursymp influenza trt household;
   model influenza = trt region nbhhc agecat dursymp / dist=bin
link=logit type3;
   repeated subject=household / type=exch  ecovb;
   lsmeans trt / pdiff ;
by _imputation_;
    ods output ParameterEstimates=gmparms ParmInfo=gmpinfo
CovB=gmcovb Diffs = gmlmd Lsmeans=gmlsm;
run;
```

Pooling: The parameter estimates derived from each analyzed data set are combined for the derivation of final results. Log OR will be combined because their distribution can be better approximated by normal distribution. The final point and 100× (1- α*)% CI estimate of the OR, where α* is the significance level at the final analysis, will be obtained by taking anti-log transformation on the corresponding statistics for the log OR, further transformation to relative risk will be done as with the primary analysis. Example SAS code is provided below (to be validated):

```
proc mianalyze parms(classvar=full)= gmlmd alpha=;
   class trt; modeleffects trt;
   ods output ParameterEstimates=es_mi_final,
run;
data es_mi_final; set es_mi_final;
   OR=exp(Estimate);
   OR_upper=exp(UCLMean);
   OR_lower=exp(LCLMean);
run;
```

4.2.4 Supplementary Analyses

A supportive analysis will be conducted on the primary endpoint imputed as per primary estimand:

- Repeating the primary analysis model on the FASi-HC. Rationale is to evaluate to what extent the BXM effect can be generalized to a wider population which also includes HHCs who are vaccinated or not confirmed negative for influenza at Baseline. This analysis will also evaluate if the BXM effect is robust to the inclusion of HHCs from households where the related IP was not treated as per protocol.

Additional sensitivity analysis of the primary estimand for robustness of the primary endpoint analysis will include:

- To evaluate the treatment effect expressed as relative risk by repeating the primary analysis by means of the modified Poisson regression model adjusting for stratification factors (Zou 2004, Yelland et al. 2011) and using the log link-function instead of the logit link-function (dist=poisson link=log in the SAS model statement). Similar to the primary analysis of odds ratio, an exchangeable variance-covariance structure will be applied to model the within-household errors. Convergence and stability of estimates will be evaluated.
- Repeating the primary analysis of the primary endpoint derived using only central lab results, if local PCR influenza test results are used for the derivation of the primary analysis in more than 5% of the PAS-HC (see Section 3.2). Rationale is to evaluate the extent to which the BXM effect is dependent on the inclusion of local PCR results.
- Repeat the primary analysis model on the PAS-HC excluding households where a HHC met the primary endpoint at or before protocol day 2, if more than 10% of the events occur at or before protocol Day 2. Rationale is to evaluate the extent to which the BXM effect is dependent on the inclusion of infections which might represent primary cases rather than secondary transmissions from the IP.

- A two-sided Cochran-Mantel-Haenszel (CMH) test to compare each treatment arm, stratified by the same factors used for the primary analysis. The adjusted proportion ratios (relative risks) and ORs will be presented at Day 5 along with the corresponding 100×(1-α*)% CI, where α* is the significance level at the final analysis.
  Rationale is to evaluate the impact of not accounting for within-household correlation when evaluating the BXM effect and other assumptions behind the validity of the GEE model used for the primary analyses.

  Sample SAS code can be found below (SAS code is regarded as "draft" until fully validated at the analysis stage):

```
proc freq data=PAS-HC;
   tables  region*nbhhc*agecat*dursymp*trt*influenza / cmh
relrisk nocol nopercent;
run;
```

Subgroup Analyses for Primary Endpoint(s)

The generalizability of primary endpoint results when comparing BXM to placebo is investigated by estimating the treatment effect in subgroups as described in Section 4.6.5.

4.3 SECONDARY ENDPOINTS ANALYSIS(SES)

4.3.1 Confirmatory Secondary Endpoints

If the primary outcome reaches significance (see Section 4.2), a hierarchical, sequential testing procedure is planned for the analysis of secondary efficacy endpoints. The secondary efficacy endpoints (see derivation details in DAP M2 Section 3.2.3) will be tested at the same significance level as the primary endpoint as part of a confirmatory analysis in hierarchical order as listed in Table 6. A given endpoint will be tested for confirmatory purposes if and only if all endpoints preceding it have reached significance at the defined level. All tests not part of the confirmatory analysis will be considered exploratory. Confidence intervals for the endpoints in the testing hierarchy will be adjusted based on the significance level at the final analysis in the same way as for the primary endpoint (Section 4.2.2).

Table 6 Confirmatory Hierarchical Order for the Secondary Endpoints

Order	Secondary Analysis
1	Symptomatic Transmission by Day 5 (Primary Endpoint and Symptoms).
2	Virological Transmission at the HH Level by Day 5 (Primary Endpoint at HH Level).
3	Symptomatic Transmission at HH Level by Day 5 (Primary Endpoint and Symptom at HH Level).
4	Virological Transmission by Day 9.
5	Symptomatic Transmission by Day 9.
6	Any Virological Infection by Day 9
7	Any Virological Infection at the HH Level by Day 9.

Order	Secondary Analysis
8	Any Symptomatic Infection by Day 9.
9	Any Symptomatic Infection at the HH Level by Day 9.

HH = household.

4.3.1.1 Symptomatic Transmission by Day 5 (Primary Endpoint and Symptoms)

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rates of symptomatic transmission by Day 5 (with a visit window extending to Day 6). It will be conducted on an endpoint which restricts the definition of the primary endpoint to only transmissions of influenza with specific new signs or symptoms, (or worsened versus baseline) at any time post baseline. This analysis will be conducted for the PAS-HC by means of the same model used for the primary analysis.

Sensitivity analyses detailed in Section 4.2.3 will be conducted also for the symptomatic transmission by Day 5 endpoint.

Subgroup analyses detailed in Section 4.6.5 will be conducted for the symptomatic transmission by Day 5 endpoint.

A supportive analysis will be conducted by repeating this model on the FASi-HC.

4.3.1.2 Virological Transmission at the HH Level by Day 5 (Primary Endpoint for the Primary Estimand at HH Level)

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rates of transmission by Day 5 at the household level. If one HHC in the household meets the primary endpoint for the primary estimand, then the household is counted as having met the endpoint. The analysis will be conducted by means of a CMH approach on the PAS-HH.

If appropriate, the analysis will be repeated as sensitivity for the FASi-HH. Additional sensitivity (not confirmatory) analyses will be conducted on PAS-HH and may also be provided for FASi-HH by using a generalized linear model (GLM) similar to the model used for the primary analysis. Sample SAS code can be found below (SAS code is regarded as “draft” until fully validated at the analysis stage):

```
proc genmod data= PAS_HH;
   class region nbhhc agecat dursymp influenza trt;
    model influenza = trt region nbhhc agecat dursymp /
    dist=bin link=logit type3;
    lsmeans trt / ilink exp diff cl;
run;
```

The model for the analysis conducted on the FASi-HH will include a covariate term indicating whether at least one of the HHC was PCR positive at screening.

4.3.1.3 Symptomatic Transmission at HH Level by Day 5 (Primary Endpoint and Symptom for the Primary Estimand at HH Level)

The analysis defined in Section 4.3.1.2 will be performed as part of the confirmatory strategy on the symptomatic transmission endpoint (see endpoint defined in Section 4.3.1.1) for the Primary Estimand, at the household level for the PAS-HH.

A supportive (not confirmatory) analysis will be conducted by repeating this analysis on the FASi-HH.

4.3.1.4 Virological Transmission by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of influenza transmission by Day 9 visit, including 1 day of tolerance (see DAP-M2). It will be conducted on an endpoint which extends the primary endpoint to include all occurrences of possible tertiary (within household) transmission (i.e., transmission from an influenza infected HHC to a further HHC) and all cases of transmission of virus with reduced susceptibility (i.e., influenza bearing a PA/I38X or T20K (influenza B only) substitution). This composite endpoint (see Table 1) will be analyzed by means of the same model used in the primary analysis for the PAS HC, and summarized by time of transmission and subtype match with IP.

A supportive (not-confirmatory) analysis will be conducted by repeating this model on the FASi-HC.

4.3.1.5 Symptomatic Transmission by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of symptomatic influenza transmissions by Day 9 visit, including 1 day of tolerance (see DAP-M2). The endpoint restricts the definition in Section 4.3.1.4 to symptomatic cases by considering only transmission of influenza with specific new signs or symptoms (or worsened versus baseline) at any time post baseline (see Table 1). This analysis will be conducted for the PAS-HC by means of the same model used for the primary analysis.

A supportive (not-confirmatory) analysis will be conducted by repeating this model on the FASi-HC.

4.3.1.6 Any Virological Infection by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of all influenza infections by Day 9 visit (endpoint is positive following any positive influenza test between baseline and Day 10). This endpoint will be analyzed by means of the same model used in the primary analysis for the PAS-HC, and summarized by time of transmission and subtype match with IP for PAS-HC.

A supportive (not-confirmatory) analysis will be conducted by repeating this model on the FASi-HC.

4.3.1.7 Any Virological Infection at the HH Level by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of all influenza transmissions by Day 9 visit. It will be conducted on the endpoint in Section 4.3.1.6 at the household level. This endpoint will be analyzed using the same confirmatory analysis method used in Section 4.3.1.2 for the PAS-HH.

4.3.1.8 Any Symptomatic Infection by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of all symptomatic influenza transmissions by Day 9 visit at the HHC level. It will be conducted on an endpoint which restricts the endpoint in Section 4.3.1.6 to only transmissions of influenza with specific new signs or symptoms, (or worsened versus baseline) which may depend on HHC age (see Table 1). This analysis will be conducted for the PAS-HC by means of the same model used for the primary analysis.

A supportive (not-confirmatory) analysis will be conducted by repeating this model on the FASi-HC.

4.3.1.9 Any Symptomatic Infection at the HH Level by Day 9

This analysis is part of the confirmatory strategy and evaluates the treatment effect on the rate of all symptomatic influenza transmissions by Day 9 visit. It will be conducted on the endpoint in Section 4.3.1.8 at the household level. This endpoint will be analyzed by means of the same model used in Section 4.3.1.2 for the PAS-HH.

4.3.2 Supportive Secondary Endpoints

4.3.2.1 Health Status Utility Endpoints

4.3.2.1.1 EQ-5D-5L Questionnaire

EQ-5D-5L is the questionnaire used to measure Quality of Life of IPs age 12 and above at Baseline, Day 3 and Day 9 visits (see Van Hout et al. 2012). The EQ-5D-5L consists of a descriptive system of 5 items (i.e., mobility, self-care, usual activity, pain/discomfort and anxiety/depression) and an EQ Visual Analogue Scale (VAS; vertical graduated scale [0-100], with 100 at the top representing “best imaginable health state” and 0 at the bottom representing “worst imaginable health state”). It will be used in this study for informing pharmacoeconomic evaluations. Analyses will be documented separately in a pharmacoeconomic report

4.3.2.1.2 Work Productivity and Activity Impairment Questionnaire

Work Productivity and Activity Impairment Questionnaire plus Classroom Impairment Questions: SHP, Version 2 (WPAI++CIQ: SHP, V2) is the questionnaire used to assess working status and the impact of influenza on absenteeism, presentism, and the ability to perform regular activities. It will be used in this study for informing pharmacoeconomic evaluations. Analyses will be documented separately in a pharmacoeconomic report.

4.3.2.2 Palatability and Acceptability of Baloxavir

IPs who are < 12 years' old who have received the oral suspension will be asked to answer a questionnaire regarding the palatability and acceptability of the study drug. IPs responses will be summarized for the FAS-IP.

4.4 EXPLORATORY ENDPOINTS ANALYSIS

Not Applicable.

4.5 SAFETY ANALYSES

Safety data collected from IPs will be analyzed in the Safety IP Population. Selected safety analyses will be repeated in the Adult, Adolescent and Pediatric Safety set.

Safety data collected from included HHCs will be analyzed in the Safety HHC Population.

4.5.1 Extent of Exposure

Actual dose and dose intensity (% of planned dose administered) will be summarized with descriptive statistics for Safety IP set.

4.5.2 Adverse Events

Verbatim description of adverse events (AEs) will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) thesaurus terms and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events Version 5.0 (NCI CTCAE v5.0).

Adverse Events for index patients

For IPs, all AEs with an observed or imputed date of onset on or after the start date of trial treatment will be considered as treatment-emergent. If the onset date of the AE is prior to the day of single dose, the AE will be considered treatment-emergent if its end date is on or after the date of administration of investigational medicinal product (IMP) and worsened in NCI CTCAE grading (initial grade < most extreme grade, or most extreme intensity is not missing and initial intensity is missing). An AE with a completely missing start date will be assumed to be treatment-emergent unless the AE has a complete non-imputed end date that is prior to study visit Day 0.

All AEs will be provided in listings, but only treatment-emergent AEs will be summarized. Additional listings will be provided for all AEs associated with COVID-19, and for the subset of those leading to study discontinuation.

An AE summary table will be provided which presents by treatment arm the number and percentage of patients in Safety IP set reporting at least one event within the following categories:

- All AEs

- Serious AEs
- AEs suspected to be caused by study medication
- AEs by most extreme NCI CTCAE grade
- AEs of special interest, defined as AEs reported as of special interest in the eCRF
- AEs of COVID-19
- AEs leading to study discontinuation

This table will be repeated in the Pediatric, Adolescent, and Adult Safety sets. This table will also be repeated using the Safety IP set, but excluding data from Study sites with health authority identified GCP violations (US site 320035 and Japanese sites 325126 and 323607), further details of the violations are given in Section 4.2.3.

AEs and SAEs will also be summarized by date of first onset according to the following categories: Day 0 after IMP, Day 1-3, Day 4-5, Day 6-9, After Day 9.

Summaries of AEs will be tabulated by MedDRA term, appropriate MedDRA levels (by system organ class [SOC] and preferred term [PT]), and when specified by NCI CTCAE grade. At each level of summarization (at least one event, SOC and PT), patients reporting more than one AE will be counted only once. For each patient, if multiple incidences of the same adverse events occur, the maximum severity reported will be used in the summaries. In particular, summaries will be presented by SOC and PT for IPs who presented at least one event for the following:

- All AEs
- Serious AEs
- AEs resulting in death
- AE suspected to be caused by study medication
- AE by most extreme NCI CTCAE grade
- AE of special interest
- AE leading to study discontinuation
- AE associated with COVID-19 leading to study discontinuation
- AEs (non-serious only) occurring in $\geq$5% of patients in at least one treatment group
- AEs occurring in $\geq$5% of patients in at least one treatment group
- AEs occurring in $\geq$1% of patients in at least one treatment group
- AE by outcome
- AE by body weight $<$80 kg and $\geq$80 kg
- AE by race
- AE by region
- Overdose

The following AE tables will be presented for the Adult, Adolescent and Pediatric Safety sets by SOC and PT:

- AE
- Serious AEs
- AEs resulting in death
- AE suspected to be caused by study medication
- AE leading to study discontinuation
- AE by most extreme NCI CTCAE grade
- AE of special interest
- AE occurring in n≥1% of patients in at least one treatment group.

Adverse Events for Household Contacts

Only study procedure related AEs will be collected for HHCs. A summary of all AEs will be provided for the household contacts.

4.5.3 Additional Safety Assessments

Vital Signs

Vital signs be assessed at screening in IPs only and will be summarized as described in Section 4.6.2.

4.6 OTHER ANALYSES

4.6.1 Summaries of Conduct of Study

The number of IPs and HHCs who enroll, discontinue, or complete the study will be summarized by treatment arm and overall in FAS-IP and FAS-HC sets. Reason for screening failure will be summarized for the SCR-IP and SCR-HHC.

Reasons for premature study withdrawal will be listed and summarized by treatment arm and overall for both IPs and HHCs.

Major protocol deviations will be listed and summarized by treatment arm and overall to be evaluated for their potential effects on the interpretation of study results (see Section 3.3).

Major Protocol Deviations Related to COVID-19 will be listed and summarized by treatment arm and overall for both IPs and HHCs to evaluate the impact of the COVID-19 pandemic on study conduct.

Listings of influenza and COVID-19 tests will be provided.

The number and percentage of IPs and household contacts at each visit will be provided.

4.6.2 Summaries of Treatment Group Comparability

Demographic and baseline characteristics of IPs (including, but not limited to, age, sex, region, race, ethnicity, number of household contacts, duration of symptoms) will be summarized overall and by treatment arm in PAS-IP and may also be presented in the FAS-IP and Safety Populations. Medical history will be summarized in FAS-IP. Descriptive summaries of baseline vital signs will be provided by treatment arm for the PAS-IP set and will comprise measurement of temperature, blood pressure and heart rate. Demographics and baseline characteristics will be produced also for the Pediatric, Adolescent, and Adult Safety sets.

Demographic and baseline characteristics of HHCs (including age of the HHC, sex of the HHC, number of household contacts of the respective IP, duration of symptoms of the respective IP) will be summarized overall and by treatment arm in the PAS-HC (see Section 3.3) and may also be presented in the FASi-HC. Medical history will be summarized in FAS-HC.

Prior and Concomitant medication

All medications (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by the IPs and HHCs will be recorded.

Prior and concomitant medications will be summarized by treatment group for the FAS-IP and FAS-HC and show the number (%) of subjects who used the medication at least once by Anatomic Therapeutic Chemical (ATC; level 2) and preferred term (PT) classification system using the WHO Drug Dictionary.

4.6.3 Virological Analyses

Effect of IP baseline viral titer: the rate of direct transmission of either influenza A or B to HHCs in relation to viral titer for the IPs at Baseline will be analyzed on the primary endpoint by including terms for the logarithm of IP baseline virus load in the main analysis model (see Section 4.2). Only the amount of virus titer of the transmitted virus matching with IP will be used for analysis, summing the viral titers of all transmitted viruses in case of co-infection in both IP and HHC subjects. The analysis will be conducted by subgroups defined on quartiles of the distribution of IP baseline virus titer across the overall PAS-HC. Point and 95% interval estimates of rates of transmission by treatment arm and the treatment effects will be presented. The analysis will also be conducted by including the interaction between baseline virus load and treatment arm in the model and results will also present the p-value for interaction term. The analysis will be performed for the PAS-HC, with baseline viral titer assessed using vp/mL values from viral RNA PCR. The analysis will be repeated for the baseline viral titer measured using 50% tissue culture infectious dose (TCID50) values. The effect of IP symptoms at Baseline will be explored in supportive models.

Similar analyses will be performed using symptomatic transmission as the target variable.

Proportion of HHCs with virus subtype consistent with IP (both treatment arms): the number (%) of HHCs with laboratory assessed PCR+ influenza and virus subtype consistent with IP in at least one post-baseline time point, will be tabulated by treatment arm and baseline IP influenza type/subtype for the PAS-HC.

Measurement of virus in IPs:

- Change from Baseline in influenza virus titer (TCID50) and in amount of virus RNA (RT-PCR) at each timepoint:

Viral titer assessments of IPs will be summarized in the logarithmic scale together with the change from baseline, by treatment arm at all time-points (Day 0, 3, 5, and 9 visits) for the PAS-IP for virology assessments expressed in vp/mL (amount of virus RNA) and $TCID_{50}$/mL. If a patient is infected with multiple virus types, the sum of those amounts of virus RNA will be used for analysis (sum performed in the original scale and result transformed in the logarithmic scale). Change from baseline will be presented in graphs and analyzed by means of analysis of covariance (ANCOVA) model adjusted for baseline titer and stratification factors. All analyses will be performed in the logarithmic scale.

Data below the limit of quantification (BLOQ) will be included in the analysis with the imputed value equal to the [$\log_{10}$ LLOQ] – 0.001, where the LLOQ is the lower limit of quantification reported in the laboratory record in vp/mL or $TCID_{50}$/mL.

For RT-PCR (i.e., values in vp/mL), in case of influenza co-infection:

- If only one titer is BLOQ, then the overall titer is equal to the non BLOQ titer
- If both are BLOQ with limit of quantification $LLOQ_1$ and $LLOQ_2$, the imputed value will be $\log_{10}(10\text{^}([\log_{10} LLOQ_1] - 0.001) + 10\text{^}([\log_{10} LLOQ_2] - 0.001))$.

Virology data expressed in Ct unit will be provided only in listings.

- Proportion of IPs with positive influenza virus titer according to TCID50 assessment at each timepoint.

Proportion of IPs positive for influenza virus titer using TCID50/mL assessments will be presented by treatment arm at Baseline and Days 3,5 and 9 visits, and is defined as the percentage of patients whose influenza virus titer is above the limit of quantification. Titer values BLOQ ("<0.75") will be considered as negative value titer. This analysis will use the CMH test stratified by the stratification factors and performed on PAS-IP, with no imputed data.

- Proportion of IPs with positive influenza virus titer according to RT-PCR assessment at each timepoint.

Proportion of IPs positive as per qualitative result (positive/negative) by RT-PCR central lab assessment (or local test results if central laboratory results are missing, see Section 4.1.1) will be presented by treatment arm over time. This analysis will use the CMH test stratified by the stratification factors and performed on PAS-IP, with no imputed data.

Measurement of influenza virus in HHCs:

- Amount of virus RNA (RT-PCR) at first time point a HHCs is positive by RT-PCR central lab assessment.

Amount of virus RNA will be summarized overall, and by IP age group (<12, ≥ 12) and by influenza virus type, in the logarithmic scale by treatment arm in HHCs positive for primary endpoint (see Section 4.2.1) for PAS-HC. Difference between treatment arms will be evaluated by means of the Van Elteren test, stratified by IP age group and influenza virus type/subtype. The analysis will also be presented in HHCs positive for symptomatic transmission by Day 5 endpoint (see Section 4.3.1.1) for PAS-HC.

Example SAS code is provided below (to be validated):

```
proc npar1way data=PASHC;
    strata age_virus_type;
    class trt;
    var log_titer;
run;
```

- Exploration of infectivity and of symptoms in relation to viral titers in HHCs

The association of viral titers and the development of symptoms (as per symptomatic endpoint definition or for other combinations the individual symptoms) will be explored by means of logistic regression or similar techniques in HHCs who become PCR positive post-baseline.

In an attempt to understand if BXM can reduce transmission beyond its effect on the secondary transmission, the statistical transmission model developed in IPs (see above: Effect of IP baseline viral titer) will be evaluated for its applicability to infected HHCs. This evaluation will consider factors such as viral titers or types and timing of symptoms in IPs (used to develop the model) vs HHCs. If the transmission model can be applied to HHCs, the probability of transmitting the virus once infected that will be derived from the model will be compared for the two treatment arms.

4.6.4 Baloxavir Resistance Analysis

Polymorphic and treatment-emergent amino acid substitutions in PA will be analysed by Sanger sequencing for all BXM treated IPs and associated HHCs. In addition, sequencing of PB1 and PB2 may be performed in IPs or HHCs if a PA/I38X, PA/T20K (influenza B only) or United States Package Insert (USPI) defined resistance associated

substitution (RAS) (see DAP M2 Section 3.2.1.7) is identified in PA of the IP and the associated HHC.

PA data will be listed for all HHCs living in households with at least one contact and/or IP with RAS. If additional sequencing in PB1/PB2 is performed, a listing is produced in households with at least one contact and/or IP with any mutation in PB1/PB2, and reported in a CSR addendum.

All USPI defined RAS, including I38X and T20K, will be listed by time point in RES-IP. Additional listings by time points will be produced for:

- IPs in RES-IP with USPI defined RAS and/or any additional treatment emergent mutations
- IPs in RES-IP with USPI defined RAS and having only post-baseline PA assessments
- HHCs in FASi-HC with USPI defined RAS

Proportion of IPs with influenza bearing I38X / T20K and USPI defined RAS substitutions (BXM arm only)

- The number (%) of IPs in the BXM arm with baseline and/or treatment-emergent I38X substitutions or a T20K substitution (influenza B only) (see DAP M2 Section 3.2.1.7) will be tabulated for the RES-IP, including only IPs having paired sequencing, i.e., having both baseline and at least one post-baseline sequencing result. I38X and T20K summaries will also be presented by age group (<12, ≥ 12 years old) for the RES-IP, having paired sequencing.
- The analysis will be repeated for the number (%) of IPs in the BXM arm with baseline and / or treatment emergent USPI defined RAS substitutions for the RES-IP (see DAP M2 Section 3.2.1.7)
- Summaries will be provided for any additional PA substitution identified only in post-baseline time points (treatment-emergent) for the RES-IP having paired sequencing.

The I38X/T20K and USPI defined RAS substitutions tables will be repeated to summarize the proportion of IPs with at least one transmission with matching I38X/T20K /USPI defined RAS in the household.

Proportion of HHCs with influenza bearing I38X / T20K and USPI defined RAS substitutions (BXM arm only)

The following analysis will be conducted in the FASi-HC population limited to those HHCs whose IPs were exposed to BXM:

- The number (%) of HHCs with an I38X, or T20K (influenza B only) substitution will be presented for all HHCs in the BXM arm and for HHCs in the BXM arm who tested positive for influenza with at least one PA sequence.
  Summaries will be also provided for HHCs in the BXM arm meeting the virological

and symptomatic transmission by Day 9 endpoint (see Section 4.3.1.4) with at least one PA sequence.
The HHC number (%) will be tabulated by HHCs having an IP in BXM arm and by HHCs having an IP with matching I38X or T20K substitution.

- The number (%) of HHCs with a USPI defined RAS substitution will be presented for all HHCs in the BXM arm, HHCs in the BXM arm who tested positive for influenza with at least one PA sequence.
  The HHC number (%) will be tabulated by HHCs having an IP in BXM arm and by HHCs having an IP with matching USPI defined RAS substitution.

Summaries will also be provided in the FASi-HH.

Phenotypic Analysis in IPs (BXM arm only)

Selected samples from BXM-treated IPs who experience a rebound in viral titers and/or show novel treatment-emergent amino acid substitutions (different from the USPI defined RAS) in PA identified by sequencing analysis will be assayed for phenotypic reduced susceptibility.

As assessment of the drug susceptibility of the influenza virus, the 50% effective concentration (EC_{50}) of BXM will be measured by the ViroSpot™ assay using baseline and selected post-baseline swab samples. EC_{50} values from post-baseline samples will be compared with baseline values from the same patient, and all EC_{50} values will be compared with values of reference strains. Respective ratios (EC_{50} post-baseline / EC_{50} baseline, and EC_{50} / EC_{50} reference) will be reported.

The detailed definition of viral rebound and the reference influenza virus strains will be documented in the DAP-M2 Section 3.2.1.8 and study CSR. Mean EC_{50} value of the two B reference strain EC_{50} values will be used to calculate EC_{50} ratio.

Using EC_{50} reference as denominator for the EC_{50} ratio, a summary of drug susceptibility will be provided for FAS-IP, including only IPs with study drug exposure, randomized to BXM, and with at least one post-baseline EC_{50} result. All EC_{50} data will be provided in listing for FAS-IP, indicating if a rebound was observed for the IP.

The phenotypic analysis will be documented in a CSR addendum.

4.6.5 Analyses of Subgroups of Interest

Subgroup analysis for the primary/secondary endpoints will be conducted in patients based on accurate matching of IPs to HHCs influenza virus to assess if the set of viral sequences within a household are more similar to each other than to sequences in the community (or between households) (see Protocol Section 3.4.6.5). The analysis will be based on whole genome sequencing on influenza A and B virus positive clinical specimens and can be accomplished either phylogenetically (in locations with a sufficient number of sequences) or using a measurement of genetic distance between

paired populations. Full details will be described in a separate analysis plan and reported separately.

Other subgroups of interest are defined by:

- The stratification factors used for randomization (as per main analysis approach).
- IP age group (Pediatric [Age $<$ 12], Non-Pediatric [Age $\geq$ 12])
- Influenza type in IPs (A/H1N1 versus A/H3N2 versus B – coinfections in IPs will be assessed as a separate category)
- Influenza season (2019/2020 vs later seasons grouped according to enrollment)

The analysis will be performed on the primary and the first secondary endpoints for the primary estimand using the main analysis model for PAS-HC (see Section 4.2). Subgroup analyses will be primarily descriptive and treatment effect will be presented in forest plots. Exploratory treatment by subgroup interaction tests (or treatment by continuous covariate interaction, if applicable) will be evaluated to investigate the heterogeneity of the treatment effects and results will be included in the study report. Only subgroups with at least 10 positive and 10 negative cases will be considered. If the limited sample size available within a subgroup prevents model convergence, the smallest subgroup will be excluded (if at least two will remain) and this issue will be reported in the study report.

4.7 INTERIM ANALYSES

Unblinded safety analyses will be performed regularly by an external statistical group (an independent data coordinating center) and reviewed by iDMC. In the context of these safety reviews, efficacy data will only be provided if requested by the iDMC. Further details on the function and logistics (e.g., frequency of the meetings) of the iDMC will be provided in the iDMC Charter. Note that an interim analysis was conducted on July 27, 2023 by the iDMC using the methodology and process described in Section 4.7.1 below and in the iDMC charter. The outcome is described in Section 4.7.2 below.

4.7.1 Planned Interim Analysis

An interim analysis of efficacy data will be performed at the end of the 2022/2023 Northern hemisphere influenza season, when recruitment rates are likely to be very low. The aim of this interim analysis is to review, in addition to safety data, the primary (see 4.2) and the first secondary endpoint (see 4.3.1.1) in order to stop the study in case of sufficient evidence of efficacy or for futility for a study with a large uncertainty regarding the true effect size. The analysis will include IPs randomized up to 07 May 2023 and their HHCs. Recruitment will continue while the analysis is being conducted. In case of a decision to stop the double blind phase as a result of the interim analysis, the study will be unblinded, otherwise other efficacy interim analyses may be performed at later stages of the study based on the same methodology.

The efficacy analysis with an early stop option will follow a group-sequential design. A Lan-DeMets alpha-spending function with O'Brien-Fleming boundaries will be used to define the null-hypothesis rejection region for the primary and the first secondary endpoint. Tests at the interim analysis will be one-sided (alternative hypothesis for the interim analysis: OR <1 for the comparison of BXM vs. placebo).. (Note that the test at the final analysis will be two-sided and the overall two-sided significance level for the study will be controlled at 5%.)

The iDMC will review interim analysis results and provide a recommendation concerning the continuation of the study. A recommendation to stop the study for efficacy with rejection of the null-hypothesis should be considered if the rejection boundaries are crossed for both the primary and the first secondary endpoint (based on their respective primary estimand). The hierarchical testing procedure will ensure control of the overall type I error rate at the pre-defined level (Hung et al. 2007; Glimm et al. 2010).

The boundaries for the interim and the final analysis will be derived based on the size of the primary analysis population (PAS-HC) at the time of the analysis using the R package rpact (Wassmer and Pahlke 2019).

If 65% of the planned HHC sample is available for the IA, it is projected that an observed OR of 0.52 or lower at this analysis will result in a statistically significant difference (based on the specific boundary p-value) between treatment arms in the PAS-HC population (i.e., an OR of 0.52 will be the MDD for the analysis; this approximately corresponds to a relative risk reduction 44% in the BXM arm compared to placebo).

For the calculations, we used an effective total sample size of 1698 HHCs, using a design effect of 1.196, resulting from a mean number of HHCs per household of 2.20 and an intra-household correlation of 0.163: The effective final sample size was calculated as 2030 / design effect, with the design effect being defined as $1+$((HH size – 1)*intra-household correlation) (Hemming et al. 2011).

Table 7 illustrates the p-value boundary for early stopping based on the projected number of evaluable HHCs at the time of the interim analysis, considering, however, that the analysis will be conducted even if the sample size will deviate from this projected number.

Table 7 Projected Interim and Final Analyses Characteristics

Analysis	No. of evaluable HHCs	% of the planned final sample size	Projected Cutoff Date [a]	Projected MDD [b]	Projected Boundary (one-sided p-value) [c]
Interim	1320	65%	May 2023	0.52	$p<0.0054$
Final	2030	100%	Apr 2024	0.74	$p<0.0233$

HHC = house hold contact, MDD = minimally detectable difference.

[a] Study month at which the last IP and the associated HHCs included in the interim / final analysis are enrolled. Analysis results will be available after data cleaning.

[b] The largest observed odds ratio that is projected to be statistically significant. For the interim analysis the first secondary endpoint determines the MDD. For the final analysis, the primary endpoint.

[c] The projected one-sided boundary for statistical significance for the number of evaluable HHCs (actual boundary to be calculated at time of analysis based on actual number of HHCs). The boundary for the final analysis assumes that no additional interim analysis is conducted. Note that the two-sided equivalent will be used for the final analysis (2 x one-sided boundary), since the tests at the final analysis will be two-sided.

The assessment of futility will be based on the outcome of the analysis of the primary endpoint as described in Section 4.2.1 as well as the secondary endpoint: continuation of the study will be considered futile for virological efficacy < 15% and symptomatic efficacy < 30% at the interim analysis. This futility stopping rule is non-binding and the iDMC may recommend continuation of the trial accounting for all information at their disposal.

The futility cutoffs for the virological endpoint were computed based on simulations using study parameters obtained from a blinded data look performed in May 2023. Simulations evaluated the probability of a binary endpoint (a set of Bernoulli trials per arms assuming 1:1 randomization in 2 arms with a success probability in the two arms reflecting the true efficacy) to meet the futility criterion at the IA and statistical significance at study end using Fisher's exact test.

Based on simulations of 100,000 trials per condition, the chance of stopping for futility if the true efficacy is 0%, 10% and 20% is 82%, 63% and 36%, respectively. For 30% true efficacy, the chance of stopping studies that would achieve a significant effect at study end would be ~3%.

4.7.2 Changes to Protocol-planned Analyses

The list of PA substitutions considered relevant for the definition of "Virological transmission by Day 9" and related endpoints, including those listed in Protocol Section 2.1.3.1, will be limited to I38X or T20K (influenza B only). Other substitutions listed in Protocol Section 3.4.6.4 will be included in additional analyses presenting the incidence of USPI defined RAS in IPs and in HHCs from the BXM arm.

PA/I38 mutations demonstrate ≥ 10-fold change in the EC_{50} of baloxavir for the recombinant virus harboring the amino acid substitution to that of the wild-type strain for type A virus and a ≥ 5-fold change for type B virus. The other listed mutations display smaller EC_{50} changes.

An efficacy interim analysis (Section 4.7.1), with a methodology compatible with the optional interim analysis described in Protocol Section 6.9.2 was conducted at the end of the 2022/2023 season.

Interim Outcome-The interim was conducted by the iDMC as described in Section 4.7.1 above and in the iDMC charter on 27 July 2023. The recommendation from the iDMC was not to stop for efficacy nor futility and to continue the trial as planned. The significance level at the final analysis will be derived using the methodology described in Section 4.7.1, which accounts for the alpha spend at the interim analysis.

5. SUPPORTING DOCUMENTATION

5.1 SYMPTOMS ASSESSED IN THE FULL STUDY HOUSEHOLD CONTACTS

Starting at Day 0, full study household contacts (HHCs; or the responsible adult of the children < 12 years of age) will be instructed in the symptoms of early influenza. Qualifying symptoms for household contacts are shown below:

Symptoms assessed in household contacts aged ≥ 12 years of age

Symptoms assessed in household contacts aged ≥ 12 years of age	
Qualifying respiratory symptoms	Cough Sore throat Nasal congestion
Qualifying general systemic symptoms	Headache Feverishness or chills Muscle or joint pain Fatigue
Symptoms assessed in household contacts aged ≥ 2 to < 12 years of age	
Qualifying upper respiratory tract infection symptoms	Cough Nasal congestion or rhinorrhea

Assessment of influenza symptoms for full study HHC ≥12 years old

DAY 0 prior to enrollment

HCP asks HHC ► about symptom (x) ▼	Have you had (x) in the last 7 days?	Is your (x) mild or is it worse than mild?	What is causing your (x)? (record the cause or write "not known")	Action
Cough	Y or N	Mild or Worse	Cause:	Do **NOT** enroll if any symptom is (1) **worse than mild** or (2) cause is **not known** or likely **infection**.
Sore throat	Y or N	Mild or Worse	Cause:	
Nasal congestion	Y or N	Mild or Worse	Cause:	
Headache	Y or N	Mild or Worse	Cause:	
Feverishness or chills	Y or N	Mild or Worse	Cause:	
Muscle or joint pain	Y or N	Mild or Worse	Cause:	
Fatigue	Y or N	Mild or Worse	Cause:	

Record Day 0 temperature.	__ __ . __°C Site:	Do **NOT** enroll if ≥ **38.0 °C**

HCP obtains swab	Is local influenza A/B, or local SARS-CoV-2 positive? Y or N	Do **NOT** enroll if **Y**

DAYS 5, 9, AND UNSCHEDULED for full study HHC ≥12 years old

HCP asks HHC ► about symptom (x) ▼		For symptoms present at baseline: is your (x) worse since you enrolled?	For symptoms not present at baseline: do you have (x)?
Respiratory	Cough	Y or N	Y or N
	Sore throat	Y or N	Y or N
	Nasal congestion	Y or N	Y or N
General Systemic	Headache	Y or N	Y or N
	Feverishness or chills	Y or N	Y or N
	Muscle or joint pain	Y or N	Y or N
	Fatigue	Y or N	Y or N

Review HHC temperature log. HHC to obtain temperature if not yet done.* Record highest temperature here ►	__.__ °C	Date:
	Site:	Time:

HCP obtains swab	Is local influenza A/B positive? Y or N **	If **Y**, refer HHC to investigator for standard of care (including anti-viral medicine). No more samples for this HHC needed. Continue to monitor for symptoms at subsequent visit if symptoms endpoint not yet met.
		If **N**, continue sampling at next visit. However, investigator may initiate standard of care if suspicion for influenza is high due to symptoms.

* Temperature obtained from tympanic thermometers provided to households (or exceptionally from other thermometers/locations in case of tympanic thermometer failure for any reason).

** Test SARS-CoV-2 via local test if HHC develops symptoms or fever versus baseline. If positive, withdraw all study subjects in household from the study.

Assessment of influenza symptoms for full study HHC ages ≥2 to <12 years old

DAY 0 prior to enrollment

HCP asks adult ► about symptom (x) ▼	Has your child had (x) in the last 7 days?	Is (x) mild or is it worse than mild?	What is causing (x)? (record the cause or write "not known")	Decision
Cough	Y or N	Mild or Worse	Cause:	Do **NOT** enroll if any symptom is (1) **worse than mild** or (2) cause is **not known** or likely **infection**.
Nasal congestion or rhinorrhea (stuffy or runny nose)	Y or N	Mild or Worse	Cause:	

Record Day 0 temperature.	__.__ °C	Do **NOT** enroll if **≥38.0 °C**
	Site:	
HCP obtains swab	Is local influenza A/B, or local SARS-CoV-2 positive? Y or N	Do **NOT** enroll if **Y**

DAYS 5, 9, AND UNSCHEDULED for full study HHC ages ≥2 to <12 years old

HCP asks adult ► about symptom (x) ▼	For symptoms present at baseline: is your child's (x) worse since your child enrolled?	For symptoms not present at baseline: does your child have (x)?
Cough	Y or N	Y or N
Nasal congestion or rhinorrhea (stuffy or runny nose)	Y or N	Y or N

Review HHC temperature log. Obtain child's temperature if not yet done.* Record highest temperature here ►	__.__ °C	Date:
	Site:	Time:

HCP obtains swab	Is local influenza A/B positive? Y or N **	If **Y**, refer HHC to investigator for standard of care (including anti-viral medicine). No more samples for this HHC needed. Continue to monitor for symptoms at subsequent visit if symptoms endpoint not yet met.
		If **N**, continue sampling at next visit. However, investigator may initiate standard of care if suspicion for influenza is high due to symptoms.

* Temperature obtained from tympanic thermometers provided to households (or exceptionally from other thermometers/locations in case of tympanic thermometer failure for any reason).

** Test SARS-CoV-2 via local test if HHC develops symptoms or fever versus baseline. If positive, withdraw all study subjects in household from the study.

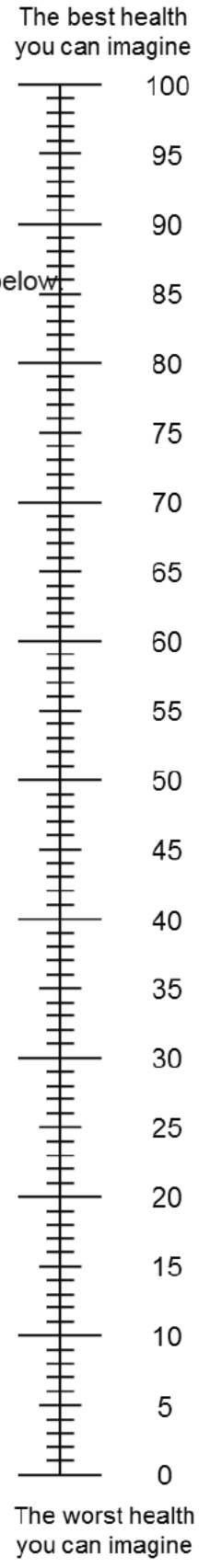
5.2 EQ-5D-5L TO MEASURE QUALITY OF LIFE

Under each heading, please check the ONE box that best describes your health TODAY.	
MOBILITY	
I have no problems walking	❑
I have slight problems walking	❑
I have moderate problems walking	❑
I have severe problems walking	❑
I am unable to walk	❑
SELF-CARE	
I have no problems washing or dressing myself	❑
I have slight problems washing or dressing myself	❑
I have moderate problems washing or dressing myself	❑
I have severe problems washing or dressing myself	❑
I am unable to wash or dress myself	❑

USUAL ACTIVITIES (e.g., work, study, housework, family or leisure activities)	
I have no problems doing my usual activities	❑
I have slight problems doing my usual activities	❑
I have moderate problems doing my usual activities	❑
I have severe problems doing my usual activities	❑
I am unable to do my usual activities	❑
PAIN / DISCOMFORT	
I have no pain or discomfort	❑
I have slight pain or discomfort	❑
I have moderate pain or discomfort	❑
I have severe pain or discomfort	❑
I have extreme pain or discomfort	❑
ANXIETY / DEPRESSION	
I am not anxious or depressed	❑
I am slightly anxious or depressed	❑
I am moderately anxious or depressed	❑
I am severely anxious or depressed	❑
I am extremely anxious or depressed	❑

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



5.3 WORK PRODUCTIVITY AND ACTIVITY IMPAIRMENT QUESTIONNAIRE PLUS CLASSROOM IMPAIRMENT QUESTIONS: SHP, VERSION 2 (WPAI+CIQ:SHP, V2)

The following questions ask about the effect of your influenza/flu on your ability to work, attend classes, and perform regular daily activities. Please fill in the blanks or circle a number, as indicated.

1. Are you currently employed (working for pay)?

____NO ___YES (If NO, check "NO" and skip to question 6.)

The next questions are about the past seven days, not including today.

2. During the past seven days, how many hours did you miss from work because of problems associated with your influenza/flu? Include hours you missed on sick days, times you went in late, left early, etc. because of influenza/flu. Do not include time you missed to participate in this study.

_____HOURS

3. During the past seven days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?

_____HOURS

4. During the past seven days, how many hours did you actually work?

_____HOURS (If "0", skip to question 6)

5. During the past seven days, how much did influenza/flu affect your productivity while you were working?

Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If influenza/flu affected your work only a little, choose a low number. Choose a high number if influenza/flu affected your work a great deal.

Consider only how much influenza/flu affected productivity <u>while you were working</u>.

Influenza/flu had no effect on my work												**Influenza/flu** completely prevented me from working
	0	1	2	3	4	5	6	7	8	9	10	

Circle a number

6. Do you currently attend classes in an academic setting (middle school, high school, college, graduate school, additional course work, etc.)?

___NO ___YES (If NO, check “NO” and skip to question 10.)

7. During the past seven days, how many hours did you miss from class or school because of problems associated with your influenza/flu? Do not include time you missed to participate in this study.

_____ HOURS

8. During the past seven days, how many hours did you actually attend class or school?

_____ HOURS (If “0”, skip to question 10.)

9. During the past seven days, how much did your influenza/flu affect your productivity while in school or attending classes in an academic setting?

Think about days your attention span was limited, you had trouble with comprehension or days in which you could not take tests as effectively as usual. If influenza/flu affected your productivity at school or in class only a little, choose a low number. Choose a high number if influenza/flu affected your productivity at school or in class a great deal.

Consider only how much influenza/flu affected productivity while in school or attending classes.

Influenza/flu had no effect on my class work												**Influenza/flu** completely prevented me from doing my class work
	0	1	2	3	4	5	6	7	8	9	10	

Circle a number

10. During the past seven days, how much did influenza/flu affect your ability to do your regular daily activities, other than work at a job or attending classes?

By regular activities, we mean the usual activities you do, such as work around the house, shopping, child care, exercising, studying, etc. Think about times you were limited in the amount or kind of activities you could do and times you accomplished less than you would like. If influenza/flu affected your activities only a little, choose a low number. Choose a high number if influenza/flu affected your activities a great deal.

Consider only how much influenza/flu affected your ability to do your regular daily activities, other than work at a job or attending classes.

Influenza/flu had no effect on my daily												**Influenza/flu** completely prevented me
Activities	0	1	2	3	4	5	6	7	8	9	10	from doing my daily activities

Circle a number

Adapted from: Reilly et al. 1996.

5.4 PALATABILITY AND ACCEPTABILITY QUESTIONNAIRE OF STUDY DRUG (INDEX PATIENTS AGED 5 YEARS OLD TO <12 YEARS OLD)

Instructions: Please answer the following questions to help us understand your experience with the study medicine. For the question with a face, indicate which of the faces best matches how you felt about the medicine.

For children who cannot read yet, parents or caregiver, please help us understand your child's experience with the study medicine. Invite your child to look at the cartoon, ask him or her the questions, and record his or her answer.

If your child cannot answer the question, please skip the question. The questions below are to be answered as soon as possible after swallowing the medicine.

1. How was the taste of the medicine?

□	□	□	□	□
Like very much	Like a little	Not sure	Dislike a little	Dislike very much

2. Would you be happy to take the medicine again?
   - □ Yes
   - □ No
   - □ Not sure/no answer

5.5 LOCAL LABORATORY SAMPLE TESTING FOR INFLUENZA AT THE CENTRAL LABORATORY

There are 3 scenarios where the local laboratory samples will be tested for influenza at the central laboratory, including rationale for the testing:

Scenario 1. Where a central laboratory sample result is missing for any reason, applicable to any index patient and household contact.

A sample collected for the central laboratory testing has not arrived and is classed as missing or is determined by the central laboratory vendor to not be appropriate for testing, such as due to sample handling issues. The local laboratory sample would be tested at the central laboratory in an attempt to avoid missing data. This would be applicable to IP and HHC samples.

For Index patient Day 0 (screening): In the rare circumstance where the local and central sample cannot be tested at the central laboratory (i.e., both judged to be contaminated by the central laboratory or both missing) and the local test at conducted at screening was a rapid influenza diagnostic tests (RIDT), the index patient post baseline sample at or before Day 3 will be evaluated to determine whether the screening sample is influenza positive or negative and the virus type/subtype if available. This decision is supported by (a) the timing of the index patient symptom onset and (b) the known time to cessation of viral shedding by RT-PCR for otherwise healthy patients treated with baloxavir marboxil and placebo from CAPSTONE-1 (median [95% CI]: 216hrs [216, 240 hrs] vs. 240hrs [240, 336 hrs] respectively).

Scenario 2. Where the index patient screening local laboratory sample result is influenza positive, but the central laboratory sample result is influenza negative.

The development of rapid molecular point-of-care diagnostics provide a fast and accurate diagnosis to help inform the clinical management of patients. The cobas Liat is the preferred system for local analysis of respiratory samples and will be provided to sites. When the preferred system is not available, other point of care tests for influenza tests such as RIDTs or other molecular/PCR based assays may be used for local laboratory sample testing for influenza in CENTERSTONE.

Local laboratory samples that are positive for influenza, provides evidence that an index patient has an influenza infection. However, when a separate sample collected for PCR testing at the central laboratory returns a negative influenza result, the index patient is determined to not have an influenza infection, even though there is evidence of infection for local laboratory testing.

Mismatches between a positive local sample result and negative central laboratory sample result may lead to an index patient (and household) not being eligible for the study, impacting the primary analysis sample size and statistical power. Therefore, in this scenario, testing the local laboratory sample at the central laboratory is warranted.

Scenario 3. Where the household contact post-baseline local laboratory sample result is influenza positive, but the central laboratory sample result is influenza negative.

Rationale regarding point-of-care diagnostics and mismatched influenza results between local laboratory and central laboratory samples is provided in scenario 2, which is broadly applicable to household contacts.

In the scenario that a household contacts post-baseline time-point has a positive local laboratory sample result and negative central laboratory sample result, a household contact will not meet efficacy endpoints, even though there is evidence that the local sample met the primary endpoint.

Algorithms for determining the sample result for analysis in scenario 2 and 3 are described in below Table 8:

Table 8 Outcomes Following Local Sample Testing at Central Laboratory for Scenario 2 and 3.

Sample collected	Local sample test result by Liat or other PCR	Central sample test result at Central lab	Local sample test result at Central lab	**Sample result for analysis**
IP – Screening	Positive	Negative	Negative	**NEGATIVE**
IP – Screening	Positive	Negative	Positive	**POSITIVE**
HHC - post baseline	Positive	Negative	Negative	**NEGATIVE**
HHC - post baseline	Positive	Negative	Positive	**POSITIVE**

HHC = household contact; IP = index patient; PCR = polymerase chain reaction.

The identification of local samples for testing at the central laboratory will be made by the Sponsor.

Note in all scenarios, retesting of identified local samples will be performed if a local sample was received by the central laboratory and the received sample was compatible with the validated central laboratory assays i.e., local sample was received in the study provided universal transport media tubes and of sufficient volume for analysis.

The local laboratory sample testing at the central laboratory and the transfer of results will occur prior to interim analyses or database lock. For a database lock, all scenarios will be completed and results data transferred to the Sponsor prior to treatment assignment information unblinding. The process has been reviewed and approved by the CENTERSTONE Steering Committee.

6. REFERENCES

Chen L, Tian Y, Chen S, et al. Performance of the cobas(®) Influenza A/B assay for rapid PCR-based detection of influenza compared to Prodesse ProFlu+ and viral culture. Eur J Microbiol Immunol (Bp). 2015;5(4):236-245.

Glimm E, Maurer W, and Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. 2010; 29(2):219-228.

Hemming K, Girling AJ, Sitch AJ, et al. Sample size calculations for cluster randomised controlled trials with a fixed number of clusters. BMC medical research methodology. 2011;11(1):1-1.

Hung HMJ, Wang SJ, Robert O'Neill R. Statistical Considerations for Testing Multiple Endpoints in Group Sequential or Adaptive Clinical Trials, Journal of Biopharmaceutical Statistics, 2007; 17:6, 1201-1210, DOI: 10.1080/10543400701645405.

Ma J, Raina J, Beyene J, Thabane L. Comparison of population-averaged and cluster-specific models for the analysis of cluster randomized trials with missing binary outcomes: a simulation study. BMC Medical Research Methodology. 2013;9(13). Available at: https://doi.org/10.1186/1471-2288-13-9.

Panel on Handling Missing Data in Clinical Trials, National Research Council. 2010. The Prevention and Treatment of Missing Data in Clinical Trials. 2010; ISBN 978-0-309-15814-5. Available at: https://www.nap.edu/initiative/panel-on-handling-missing-data-in-clinical-trials.

Reilly MC, Tanner A, Meltzer EO. Work, classroom and activity impairment instruments: validation studies in allergic rhinitis. Clinical drug investigation. 1996;11:278-288.

Van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health 2012;15:708-715.

Wassmer G, Pahlke F. Rpact: Confirmatory adaptive clinical trial design and analysis. R package version 3.0.3. 2020. Available at: https://CRAN.R-project.org/package=rpact

Welliver R, Monto AS, Carewicz O, et al. Effectiveness of oseltamivir in preventing influenza in household contacts: a randomized controlled trial. JAMA. 2001;285:748–754.

Yelland LN, Salter AB and Ryan P et al. Performance of the Modified Poisson Regression Approach for Estimating Relative Risks From Clustered Prospective Data, *American Journal of Epidemiology*, Volume 174, Issue 8, 15 October 2011, Pages 984–992, https://doi.org/10.1093/aje/kwr183.

Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA. 1998;280(19):1690-1691.

Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004 Apr 1;159(7):702-6. doi: 10.1093/aje/kwh090. PMID: 15033648.

Signature Page for MV40618 (Xofluza/Baloxavir Marboxil/RO7191686/138558) SAP V4
System identifier: RIM-CLIN-617672

Approval Task	[redacted] Company Signatory 06-Aug-2024 08:51:11 GMT+0000